Risk and Benefit Evaluation in Development of Pharmaceutical Products

by C. S. Aaron, P. R. Harbach, S. S. Mattano, J. K. Mayo, Y. Wang, R. L. Yu, and D. M. Zimmer

Pharmaceutical products are intended to cure disease, reduce pain and suffering, prolong life, and correct metabolic deficits in patients. However, the potential patient population is intrinsically genetically heterogeneous, and this factor complicates the evaluation of data on all aspects of safety evaluation of new drugs. Often the genetic heterogeneity is related to drug metabolizing capacity, but recent evidence suggests that heterogeneity in repair capacity as well as structural integrity of the chromatin (fragile X) have been shown to be relevant. Because drugs are biologically active and may have more than one type of effect, the evaluation of a large number of parameters is necessary in arriving at a rational estimate of potential risk. In this paper, several specific examples of risk assessments and some generic genotoxicity questions that are recurrent, including the question of the relevance of in vitro chromosomal aberration induction at high dose/sampling time, are raised. Other examples of the kinds of concerns from the safety evaluation of U-48753E, U-54461, and U-68553B are discussed. The drug U-48753E was discovered to be slightly mutagenic in the AS82 assay, and significant efforts were expended in evaluation of the metabolism-based generation of a reactive intermediate. The drug U-54,461 was shown to be capable of breaking chromosomes in vitro but extensive in vivo data as well as a variety of other studies served to reduce the level of concern substantially. The induction of apparent chromosome breaks and gaps by U-68553B was shown to be an artifact that was restricted to a single cell line of rodent cells. These diverse examples of risk assessments illustrate the complexity of risk/benefit analysis in drug development.

Introduction

The focus of this conference, as suggested by its title “Environmental Mutagenesis in Human Populations at Risk,” is a very important aspect of environmental mutagenesis. This paper deals with some of the considerations relating to this topic from the perspective of the development of human pharmaceutical products. Mutagens are present in the natural environment (1) and in the “synthetic” environment with which we have surrounded ourselves. For example, the simple act of cooking food produces potent mutagens. Even uncooked food and other environmental constituents can contain mutagens. Many pesticides are known to be mutagenic in short-term tests, and the care taken in handling these compounds should reflect the degree of hazard. Thus, the task of fully comprehending the problem of environmental mutagenesis in human populations at risk is not trivial. Part of the effort we devote to this work is in identifying mutagens and potential mutagens. Another part must be devoted to study of the populations exposed to the potential mutagens, and still a large part of the job is estimating, no matter how primitively, the degree of impact of mutagen exposure. The daunting difficulty of the problem is clearly delineated (2) by the case of ionizing radiation, where despite the devotation of significant resources over many years, only limited estimates are possible. Carrying out such extensive scientific investigation for any other potential mutagenic agent is highly unlikely. The case for chemicals and drugs is even more difficult because of the complications of absorption and metabolism.

In the pharmaceutical industry, “lead” (or prototype) compounds are often tested for mutagenic properties at an early stage, usually before a material is introduced to humans for the first time. Here we focus on efforts to identify and evaluate mutagens using short-term tests and to determine whether these findings are relevant in vivo in humans. This area presents the most direct application of environmental mutagenesis because these materials are known to be biologically active and are intended to be administered to humans, sometimes in relatively large quantities. In many countries (3–6), the relevance of mutagenicity to the overall hazard evaluation of drugs is well established, although the details of the process vary greatly from one country to the other. Of course, for companies and other institutions (like the World Health Organization) intending to produce and distribute drugs worldwide, safety to the consuming public is paramount. We begin by considering some of the regulatory require-
ments and the need for further harmonization of guidelines. We then give several examples of materials that have shown positive responses in short-term tests and discuss some of the factors that were considered in developing a risk assessment for these materials. Finally, we briefly discuss several techniques we are using to decrease the cost of evaluation and to improve our ability to understand the significance of positive findings.

General Considerations

Regulatory authorities in Europe, Japan, and the United States have developed extensive guidelines for the mutagenicity data needed to support registration of new drugs and chemicals (3–7). Virtually all these countries have acknowledged that the need for such a data is driven by both the potential risk from somatic mutation as well as from germ cell mutation. However, the difficulty of observing germinal mutation in the offspring of exposed population (2) due to the small populations that can be observed as well as the ethical restraint imposed on continued, voluntary exposure of humans to such materials has generally led to a lack of efforts to regulate on that basis (8). Thus, the mutagenicity data currently being produced in the countries previously mentioned are generally used to evaluate the carcinogenic potential of a compound.

Human drugs are used for a variety of purposes, from the straightforward relief of pain to the elimination of parasites and cure of microbial disease. These conditions vary enormously in severity and in the consequences of not using the drugs, and therefore the mutagenicity of the drugs (particularly in vitro mutagenicity) must be carefully weighed in light of the benefits. For example, cancer is a dread disease that takes many lives in all countries, and in the United States it is the contributing cause of death in about one-fourth of all deaths. AIDS is a particularly significant disease with no chance of survival. Diseases such as cancer and AIDS are often treated with particularly aggressive therapy. In the case of cancer, almost all of the routinely used, effective chemotherapeutic agents are mutagens and may be carcinogenic. AZT used in the treatment of AIDS as well as the newer compounds used are generally mutagenic in one or more routine assays that evaluate potential carcinogenicity. Other diseases from the common cold to insomnia are less serious, and therefore the kinds of side effects that are tolerable do not include mutagenicity. The problem of risk assessment for drugs must be solved using a holistic approach that may well provide a model for the assessment of human populations at risk from other potential environmental mutagens. This means that all aspects of the risk assessment must be evaluated, e.g., exposure, whether the risk is voluntary, whether the population at risk is the population exposed, and whether the risk is real or only perceived.

The Upjohn Test Battery

The evaluation of the safety of a proposed new drug begins with subjecting it to several short-term tests. The timing of the assays is dictated in some measure by the availability of the drug because at very early stages of a drug’s development, very little of the drug is available and therefore few safety tests can be performed. Usually the first assay is a microbial assay for mutagenesis such as the Ames test with Salmonella typhimurium. This assay is simple, cheap, and reasonably good at identifying bacterial mutagens (9). Our experience with drug candidates has been that about 5–10% of compounds chosen randomly are positive in this test (if homologous series of compounds are treated as single entities), and this level is also reflected in the Japanese experience (3) with a variety of materials; clearly the high prevalence of mutagens in the data set of the National Toxicology Program (10) is not reflective of the universe of chemicals.

The Ames test is followed or accompanied by tests of the ability of the compound to reach and damage mammalian DNA. This may take the form of the in vitro unscheduled DNA synthesis (UDS) assay or other assays of DNA-damaging ability. Such an assay is needed because the bacterial chromosome is significantly different from that of the eukaryote. In the UDS assay, rat hepatocytes are exposed to drug, and the incorporation of DNA precursors during repair is followed with autoradiographic techniques.

The Ames assay and UDS are carried out at relatively early stages of the development process and are followed by more elaborate tests such as in vitro cytogenetic assays, mammalian cell mutation assays and in vivo cytogenetic assays when sufficient drug becomes available. Our experience has been that approximately 5–10% of these potential drug candidates produce positive responses. This of course excludes anticancer agents which are routinely positive. This figure (5–10%) bears emphasizing because it is significantly lower than the frequency of positives among the National Toxicology Program (NTP) database of compounds and reflects the selected nature of those compounds. The importance of this discrepancy between the NTP data and that of scientists in the pharmaceutical industry (many colleagues in the pharmaeutical industry and the contract testing laboratory industry would probably confirm the frequency) is that the prevalence of positives in the data set determines to some extent the reliability of the estimates of sensitivity and selectivity at lower and therefore more realistic prevalence rates.

Positive responses in these early tests, i.e., compounds which raise a “red flag,” are given significantly greater scrutiny. Unlike the process used in the NTP evaluation of mutagenic potential of a nonrandom set of carenogens (10), further study of the potential for metabolism, complications caused by high-dose toxicity, and species specificity are often aggressively evaluated. Our experience has been that mechanistic follow-up of these screening results often reveals them to be due to species or system-specific effects. Even when the follow-up reveals confirmatory information, such follow-up is warranted because better decisions are possible concerning alternatives of structure, for example.
Other Information Used in Risk Assessments

Agents that are capable of reaching and damaging the DNA of human chromosomes represent a real risk and must therefore be taken seriously. However, a positive response in a particular short-term test is not definitive proof that the compound is capable of reaching or damaging human DNA. One of the key factors leading to a discrepancy between the result in a bacterial or cell culture assay and the expected outcome in humans is a difference in the metabolism of the compound in vivo. Also, these in vitro assays do not reflect the human body's absorption, distribution, and elimination of the compound. Thus, the probability that the active ingredient in the drug is the active principle in the mutational end point must be high before significance can be attached to the outcome of an in vitro short-term test. Furthermore, the amount of relevant chemical species must be produced in vivo. The in vivo (whole animal) system is much more complicated in terms of distribution and excretion than an in vitro bacterial or cell culture assay because many time-dependent factors come into play. Even a factor like mutagenicity, which would not be expected to have a threshold in vitro, may have a physiologically mediated practical threshold in vivo. Unfortunately, the kind of data necessary to understand these factors often depends on availability of radiolabeled drugs and sophisticated analytical techniques that are available only in later stages of drug development.

Another critical factor that is often difficult to gauge is the likely exposure of humans to the drug. Clearly, human consumption of a drug will be less than kilograms per day, but the disease entity and factors related to bioavailability (the fraction that will reach the circulation following oral exposure), potency (plasma concentration required to produce the desired therapeutic effect), and biological persistence (half-life in the body) often determine the exposure. All these factors will similarly impact the risk assessments for nondrug entities but may be even harder to evaluate because of the lack of radiolabeled compound, complex nature of the material, or other reasons.

One factor that is considered in the case of drugs, which may not have an analogy in assessment of other environmental mutagens, is the side effect profile of the competitor drugs. For example, AZT causes severe bone marrow suppression and a constellation of other effects including mutagenesis in some short-term tests, which limit the time a patient can take the drug even if he or she can afford it. The combination of the immunosuppression (which may be the reason for the appearance of the enhanced frequency of Kaposi's sarcoma in AIDS) with the mutagenic properties of AZT would tend to support caution even in the earlier stages of AIDS therapy if truly life-saving AIDS therapy becomes a reality. An alternative drug with equal or greater potency (efficacy) would in all likelihood replace AZT if it had significantly reduced side effects. In fact, a potent anti-AIDS drug with reduced genotoxic potential compared to nucleotide analogs would be preferable because the characteristic of this disease is immunosuppression. Another trade-off totally unrelated to AIDS is the case of U-68553B, a potential antischizophrenic drug. In this case the value judgment relates the potential mutagenic risks suggested by in vitro results to the existing therapies, which are so onerous that one of the biggest problems is patient compliance (assuring that the patients take the drug). Several specific examples will illustrate some of the kinds of considerations in the pharmaceutical risk assessment setting.

**U-48753E.** U-48753E was a promising antidepressant drug of a novel class. Depression afflicts many millions of people throughout the world. In the early development of this drug (2), a range of short-term genotox assays was carried out including the Ames test and the in vitro UDS assay, which were both negative. After the compound was in clinical trials, it was found to produce mutations in the Chinese hamster ovary/HPRT assay, one assay for assessing the ability to induce gene mutations in eukaryotic cells. Mutagenesis was only observed in the presence of S9 metabolic activation, and the increases were sporadic. More consistent mutagenicity was observed when the genetically engineered cell line AS52 was used as the tester strain, again, only with an added metabolic activation system. This finding led Upjohn to stop clinical trials and initiate an investigation of the mutagenic risk associated with the drug (2).

The drug U-48753E is like many drugs that contain methylamine substituents. These kinds of compounds are commonly demethylated by mammalian enzymes to yield formaldehyde equivalents. In fact, this compound was found to produce formaldehyde in the presence of rat liver enzymes. Furthermore, the addition of formaldehyde dehydrogenase to the incubation mixture ameliorated the mutagenicity, reinforcing the hypothesis that the observed mutations were probably induced by liberated formaldehyde in vitro. Formaldehyde dehydrogenase is a potent and ubiquitous mammalian enzyme and is capable of rendering low concentrations of formaldehyde harmless in vivo. Although there is some evidence of induction of cancer by formaldehyde at the site of inhalation, these studies all involve high exposure and may be explained as a consequence of irritation rather than direct genotoxicity. Clearly, formaldehyde is mutagenic, but it is also rapidly metabolized and reacts with a variety of electrophilic sites to significantly decrease the potential for direct reaction in the nucleus with DNA. Nevertheless, the formaldehyde which could theoretically be produced by U-48753E reveals that liberated formaldehyde from this source would yield less than 0.01% of the normally circulating levels of formaldehyde equivalents. Because this was not viewed as a significant perturbation of the formaldehyde pools, clinical trials were resumed.

**U-68553B.** U-68553B was an antischizophrenic drug candidate. Early in the development process, the compounds in the analog series revealed a tendency to produce mutations in the Ames test. Interestingly, the effect was only observed in the presence of S9 metabolic activation. Several experiments were carried out that supported the proposition that the mutagenicity was mediated by reactive oxygen or other radicals. However, the compound was
unable to induce mutations in mammalian cells in vitro, nor was any evidence of mutagenicity seen in vivo.

**U-54461.** Bropirimine (U-54461) is a biological response modifier that is thought to act through the induction of lymphokines. It was originally developed as a potential anticancer agent to be used after tumor debulking through surgery. A decided advantage of this kind of therapy (immune stimulation) is the lack of side effects generally associated with more traditional chemotherapy.

Several mutagenicity studies with bropirimine failed to produce enhanced mutagenesis. However, the compound was shown to induce chromosome aberrations in vitro in the absence of activation at very high doses (12). Testing failed to offer a consistent explanation for the positive in vitro results (13), and no evidence of in vivo mutagenicity was found (14). Although a detailed mechanistic explanation of such observations has seldom been possible, many workers are coming to the conclusion that isolated high-dose positives, particularly when accompanied by toxicity, are irrelevant to the assessment of risk for mutation.

**U-73975.** The final example is that of the anticancer agent U-73975, and we combine this discussion with mention of some recent results on the detection of mutagenesis in the primate. This drug is, like many cytostatic anticancer agents, a very potent mutagen. By very potent, we mean that it produces mutations at vanishingly small doses (15,16). U-73975 is very specific in its binding properties, preferring adenines in the DNA to the exclusion of guanines. The binding data were obtained using naked DNA in vitro, and for a long time little information was available in vivo. However, the result of testing this material in the Ames test with Salmonella typhimurium strain TA102 revealed that it was a particularly potent mutagen in this strain at doses that were negative in the other common strains; strain TA102 was genetically engineered to detect A-T base pair mutagenesis, the only Ames test strain with this property. The extreme specificity of the binding combined with the high potency suggest that the compound might be highly selective, and clinical trials are now in progress.

This is a class of compounds which has a very high specificity for adenine sites in DNA, and we are in the process of evaluating the mutational spectra in systems that have eukaryotic chromosomes. Specifically, we are collaborating (15) with Ken Tindall of the National Institute of Environmental Health Sciences in evaluating the spectrum of U-73975 induced mutants in the AS52 cell line. Preliminary results suggest that about half of the mutants are functional deletions, but a sizeable number of intragenic changes appear to have been induced. We are currently sequencing these mutants to determine where the mutations have occurred, but there are several copies of the consensus sequence for binding within the structural gene.

For a variety of reasons, we also studied the induction of mutations in the cynomolgus monkey by U-73975, and the results may indicate a potential problem with the use of HPRT assays for monitoring. We treated three monkeys with U-73975 and monitored the appearance of HPRT mutations in the peripheral blood using an adaption of the technique of Albertini (16). In this experiment (unpublished data), we also monitored the appearance of chromosome aberrations in the peripheral blood using standard cytogenetic techniques. Our previous experience (16,17) had shown that mutants appeared in the periphery only after a period of several weeks. In these experiments the frequency of cells with gross rearrangements clearly showed induction of complex rearrangements, but we failed to find mutant T-cells even though we followed the animals for a significant period beyond that found to be necessary in earlier experiments. We do not yet understand this result because in mammalian cells in culture, induction of mutations was observable at very low levels of U-73975. Thus it seems prudent to add the note of caution that a negative result in an in vivo HPRT assay with T-cells may significantly underestimate the potential for mutation induction.

**Alternative Methods of Evaluation**

Returning for just a moment to the technologies for evaluating drugs, we want to mention two techniques that are particularly exciting in the context of genetic toxicology for potential drugs. We should point out the fact that in the early stages of drug development, very little compound is typically available, and this factor severely limits the amount of testing that is possible. Particularly in those cases in which a series of analogs are available, structural analysis often provides insight into the mechanism through which the mutagenicity is manifest. We have found the Spiral Salmonella assay, originally introduced by Claxton and co-workers (18,19) to be particularly useful in identifying potent mutagens and potentially useful for structure-toxicity evaluation of a series of analogs. The chief benefit of the system is that relatively small quantities of drug [or complex environmental mixture (19)] are required for a test. However, this system suffers from the fact that compounds that are water soluble diffuse into the agar and hence appear less mutagenic than might be the case in the standard plate incorporation or preincubation assay. Nevertheless, we have found this method useful in structure—activity evaluation (20) and recommend its consideration in situations in which limited quantity of test substance precludes more extensive testing.

Another technique that has proven interesting in the recent past is computerized structure activity programs in toxicology. The program that we use is called TOPKAT and was written by HeathDesigns Incorporated. Using this system, it is possible to estimate the toxicity for several end points including the oral LD50 in rats, eye irritation, dermal irritation, and mutagenicity and carcinogenicity. This system suffers from the same shortcomings as all computerized systems of this kind, namely, it is limited by the database used in “teaching” the program what aspects of molecules are toxic. The difficulty is that most new drugs are novel chemical entities and therefore have moieties that may be unusual for the database and thus hard to estimate. However, by careful dissection of the molecules into component parts, it is often possible to arrive at likely estimates of the toxicity. This is not the only
system available for the prediction of carcinogenesis of compounds based principally on structural features (9,21–24), however, TOPKAT has the advantage of being compact enough to fit on a personal computer, and it is commercially available.

Summary

In summary, anticipating the effect of environmental mutagenesis in human populations at risk is difficult. Identifying the potential mutagenic material is only the first step in the process, which must include significant exposure estimates and study of metabolism. Competing risks must be weighed. Dose response must be ascertained, and a holistic approach to the risk assessment must be adopted. Particularly, in those cases of low-level environmental exposure, creative methods must be developed to evaluate the potential hazard using computerized techniques and miniaturized and automated procedures. Perhaps the approaches used in pharmaceutical research can be adapted to the problem and thereby improve the throughput and reliability of the process. Structures of the molecules mentioned in this manuscript are given elsewhere (15,25).

REFERENCES

1. Ames, B. N., Magaw, R., and Gold, L. S. Ranking possible carcinogenic hazards. Science 236: 271–280 (1987).
2. Sankaranarayanan, K. Ionizing radiation and genetic risks (special issue). Mutat. Res. 258: 1–122 (1991).
3. Sato, S. Regulation of food related carcinogens in Japan. Regul. Toxicol. Pharmacol. 11: 149–157 (1990).
4. Kramers, P. G. N., Knaap, A. G. A. C., van der Heijden, C. A., Thalmann, R. D. F. M., and Mohn, G. R. Role of genotoxic assays in the regulation of chemicals in The Netherlands: considerations and experiences. Mutagenesis 6: 487–493 (1991).
5. Carere, A., and Benigni, R. Strategies and governmental regulations. Teratog. Carcinog. Mutagen. 10: 199–208 (1990).
6. Muller, L., Kasper, P., and Madle, S. The quality of genotoxic testing of drugs. Experiences of a regulatory agency with new and old compounds. Mutagenesis 6: 143–149 (1991).
7. Kirkland, D., Ed. Basic Mutagenicity Tests: UKEMS Recommended Procedures. Cambridge University Press, Cambridge, (1990).
8. Delarco, V. L., Rhomberg L., and Shelby, M. D. Perspectives and future directions for genetic risk assessments. Environ. Mol. Mutagen. 16: 132–134 (1990).
9. Ashby, J., and Tennant, R. W. Definitive relationship among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the US NTP. Mutat. Res. 257: 229–305 (1991).
10. Tennant, R. W., Margolin, B. H., Shelby, M. D., Zeiger, E., Haseman J. K., Spalding, J., Caspary, W., Resnick, M., Stasiewicz, S., Anderson, B., and Minor, R. Prediction of chemical carcinogenicity in rodents from in vitro genotoxicity assays. Science 236: 933–941 (1987).
11. Aaron, C. S., Stankowski, L. F., Harbach, P. R., Valencia, R., Mayo, J. K., Mirsalis, J., Mazurek, J. K., Steinmetz, K. L., Wiser, S. K., Zimmer, D. M., and Trzos, R. J. Comparative mutagenicity testing of a drug candidate, U48758E: mechanism of induction of gene mutations in mammalian cells and quantitation of potential hazard. Mutat. Res. 223: 111–120 (1989).
12. Aaron, C. S., Petry, T. W., Thilagar, A., Kumaroo, P. V., and Kirby, P. Comparative mutagenicity testing of bropinirime, 1. Induction of chromosome aberrations in CHO cells is not reflected in induction of mutation at the TK locus of L5178Y cells. Mutat. Res. 252: 221–228 (1991).
13. Petry, T. W., Aaron, C. S., Ulrich, R. G., Donarski, W. J., Thilagar, A., and Kumaroo, P. V. Comparative mutagenicity testing of bropinirime, 2. Further characterization and mechanistic investigation of clastogenesis. Mutat. Res. 252: 229–238 (1991).
14. Aaron, C. S., Petry, T. W., Branstetter, D. G., Wickrema-sinha, W. J., Yu, R. L., Song, R., Thilagar, A., Kumaroo, P. V., Sams, J. P., Thornton, B. A., Brown, F. K., and Trzos, R. J. Comparative mutagenicity testing of bropinirime, 3. Bropinirime does not induce cyto genetic damage in vivo in the rat but does produce micronuclei in the mouse. Mutat. Res. 252: 239–246 (1991).
15. Hhook, G. J., Stankowski, L. F. Jr., Aaron, C. S., and Tindall, K. R. Characterization of gpt mutations in AS52 cells induced by the antitumor drug U73975. Environ. Mol. Mutagen. 19(Suppl. 20): 26 (1992).
16. Zimmer, D. M., Aaron, C. S., O’Neill, J. P., and Albertini, R. J. Enumeration of 6-thioguanine-resistant T-lymphocytes in the peripheral blood of nonhuman primates (cynomolgus monkeys). Environ. Mol. Mutagen. 18: 161–167 (1991).
17. Harbach, P. R., Filipunas, A. L., Wang, Y., and Aaron, C. S. DNA sequence analysis of spontaneous and N-ethyl-N-nitrosourea-induced HPRT mutations arising in vivo in cynomolgus monkey T-lymphocytes. Environ. Mol. Mutagen. 20: 96–105 (1992).
18. Houk, V. S., Schalkowsky, S., and Claxton, L. D. Development of and validation of the Spiral Salmonella assay: an automated approach to bacterial mutagenicity testing. Mutat. Res. 223: 49–64 (1989).
19. Houk, V. S., Early, G., and Claxton, L. D. Use of the Spiral Salmonella assay to detect the mutagenicity of complex environmental mixtures. Environ. Mol. Mutagen. 17: 112–121 (1991).
20. Mayo, J. K., Wang, Y., and Aaron, C. S. Further characterization of the Spiral Salmonella assay in mutagenicity testing of drug candidates (abstract). Environ. Mol. Mutagen. 19(Suppl. 20): 41 (1992).
21. Enslein, K. B., Blake, W., and Borgstedt, H. H. Prediction of probability of carcinogenicity for a set of ongoing NTP bioassays. Mutagenesis 5: 305–306 (1990).
22. Lewis, D. F. V., Ioannides, C., and Parke, D. V. A prospective toxicity evaluation (COMPACT) on 40 chemicals currently being tested by the National Toxicology Program. Mutagenesis 5: 433–435 (1990).
23. Rosenkranz, H. S., and Klopman, G. Prediction of the carcinogenicity in rodents of chemicals currently being tested by the US National Toxicology Program: structure-activity correlation. Mutagenesis 5: 425–432 (1990).
24. Doi, O. Guidelines for Toxicity Studies of Drugs Manual (Japan). Yakuji Nippo, Ltd. Tokyo, 1990.
25. Aaron, C. S. Evaluation of drug candidates in a battery of short-term genetic toxicity assays: overview. Mutat. Res. 223: 106–109 (1989).