2 DNA Damage

2.1 MOLECULAR DOSIMETRY

Carcinogen-induced DNA damage involves the formation of specific adducts with DNA and protein. Molecular dosimetry provides the tools to elucidate dose–response relationships for these adducts in carcinogen-exposed humans and laboratory animals.

Compared with measurement of external dose, molecular dosimetry offers the distinct advantage of integrating dose-dependent differences in absorption, distribution, biotransformation, and DNA repair to determine carcinogen dose more accurately. Moreover, molecular dosimetry is capable of determining which events play a critical role in eliciting cancer by examining, for instance, the dose–response relationships between DNA-adduct levels and the incidences of tumours. For example, cell proliferation is required to convert promutagenic DNA adducts to mutations. Since the extent of cell proliferation is known to be dependent on age, tissue, and dose, the impact of each factor on carcinogenesis in a specific population should be examined at the molecular level to improve the certainty in extrapolating a wide array of dose-response relationships for chemical carcinogens among organs and across species.

2.1.1 ABSORPTION AND DISTRIBUTION

Humans are exposed to chemical carcinogens most often by ingestion or inhalation. Consequently, differences in toxicokinetics are usable to anticipate within and among species major differences in the extent to which organs and tissues are exposed directly to a carcinogen. For oral intake, species differences in stomach pH or in metabolism of intestinal flora may affect the availability of a compound to the systemic target organ. Likewise, the formation of carcinogenic metabolites may depend on these conditions, as exemplified by the formation of nitrosamines from nitrite and secondary amines in the acid medium of the stomach. For exposure by inhalation, alveolar ventilation varies greatly among species, as illustrated by the fact that, on a body weight basis, the amount of a chemical inhaled by humans is only 10% of that inhaled by mice, so that for brief times, systemic inhaled doses are usually much less in humans than in mice. By contrast, the rate of elimination in humans is generally slower than in rodents, so that

Methods to Assess DNA Damage and Repair: Interspecies Comparisons
Edited by R. G. Tardiff, P. H. M. Lohman and G. N. Wogan

accumulation may occur more readily in humans than in mice. Physiologically based pharmacokinetic models can simulate such toxicokinetic behaviour of a compound in various species. Allometric scaling, a mathematical interpolation procedure, enables estimation of physiological parameters such as blood flow in a specified organ of a test species when such a parameter has been characterized for the same organ in some other species. Physiologically based pharmacokinetic models are already proving useful in estimating organ load, internal exposure, and covalent binding of reactive intermediates.

2.1.2 BIOACTIVATION AND DETOXIFICATION

To form DNA adducts, most carcinogens have to be converted by the body into reactive metabolites, usually more reactive than the parent compound. This process, bioactivation, is accomplished by enzymes (sometimes called "drug metabolizing" enzymes) present to varying degrees in numerous tissues of the body. These enzymes are families of isoenzymes having varying composition according to cell types and within and among the same and different species. The result is a difficulty in extrapolating the formation of metabolic by-products from one organ to another in the same species or from one species to another for the same organ. Such extrapolations are further complicated when results are obtained from one component of a cell (e.g., the microsomes), yet other cell components have similar capabilities. Moreover, some of these enzymes require the addition of co-factors to function properly outside the body; failure to do so in *in vitro* preparations results in erroneous conclusions about biotransformation reactions and toxic potential. The cytochrome P_{450} isoenzyme families are frequently recognized for their important role in bioactivation, leading often to overlooking the relatively important role of conjugating enzymes. Therefore, to obtain relevant findings from the measurement of biotransformation reactions, "metabolically competent cells" that contain not only P_{450} but also the full complement of xenobiotic transferases are needed.

These same enzymes can also detoxify carcinogens. Certain low molecular factors in the cell, like glutathione or methionine, may trap reactive intermediates before they can form appreciable levels of DNA adducts. Clearly the extent of adduct formation in a cell depends on the efficiency of detoxification and trapping, relative to bioactivation. The differences in every cell type and in every species of the composition of "drug metabolizing" enzymes are major determinants in tissue or cell specificity for carcinogenicity as well as of species differences. In genetically heterogeneous individuals of the same species, a wide variation in these enzyme activities exist, which may explain much of the interindividual variation in sensitivity to chronic toxicity of chemicals. Widely different DNA-adduct levels in the same tissue between individuals or between species may be correlated to differences in the balance between activation and detoxification. The possibility exists that compounds may be activated not only to DNA-adduct-forming

metabolites but also to metabolites able to function as cancer promoters. Therefore, the prediction is very difficult of the extent of formation of any given reactive metabolite when formed through metabolic pathways in which major cell, tissue, and species heterogeneity is present with respect to enzymatic competence for the reactions involved.

2.1.3 METHODS FOR MOLECULAR DOSIMETRY

Many chemical carcinogens or their metabolites form covalent adducts with nucleophilic sites on nucleic acids. Because formation of such DNA adducts is regarded as a key step in mutation and in the initiation of cancer, their measurements could be useful in estimating cancer risk. Such measurements should take into account inter-individual differences in dose, duration, absorption into susceptible tissues, and biotransformation.

Adduct levels vary among organs and cells. Furthermore, the spectrum of adducts also varies widely among carcinogens, yet only a few adducts—each subject to removal by chemical or enzymatic means—are likely to be critical to the induction of disease. The magnitude of genotoxic injuries are held to be correlated in some way to the persistent residues of critical adducts in target cells. For several carcinogens including aflatoxin B1, a linear relationship exists between administered dose and DNA-adduct concentration in experimental animals, regardless of whether the dosage was singular or repeated. Some non-linear correlations have been obtained in some cases, due perhaps to saturation of activation or detoxification pathways, thereby requiring the exercise of caution when extrapolating from high to low doses.

For such analyses, samples of human target tissues cannot be routinely collected; surrogates are needed. With parallelism between adduct levels in tissues of target and surrogate species, measurements in laboratory animals can be used to estimate those in humans. A linear relationship between adduct levels and external or internal doses of a carcinogen facilitates such extrapolations. Although such parallelisms and linear relationships have been observed between experimental animals and humans, generalizations are not yet possible.

The analysis of DNA adducts *in vivo* is complex because the adducts are formed in relatively small quantities (e.g., typically <100 fmol/μg DNA or <3 adducts per 10^5 nucleotides). Two techniques, immunoassay and ^{32}P-postlabelling, can detect DNA adducts in either experimental animals or humans. The immunological approach requires that monoclonal or polyclonal antibodies be produced against either carcinogen-modified DNA or carcinogen-nucleoside adducts coupled to a protein carrier. The resulting antibodies are then used to quantify specific adducts by either of two methods: radioimmunoassay (RIA) or enzyme-linked immunosorbent assay (ELISA). The quantity of adduct detectable by immunoassays depends on the specific method, with ELISA generally giving greater sensitivities (~1 adduct per 10^7 nucleotides) than RIA. Both methods are

relatively inexpensive, and can be used to process many samples. However, their sensitivity is generally limited by the amount of DNA that can be analyzed. Furthermore, because an antibody may cross-react with a broad spectrum of adducts, the findings of DNA adducts may be complicated for individuals with unknown exposures to alkylating substances. These limitations can be overcome partially by combining chromatographic purifications (e.g., HPLC) with the immunoassays.

Under appropriate conditions, the ^{32}P-postlabeling method can detect DNA adducts at nearly one adduct in 10^{10} nucleotides, but is limited in its ability to identify adducts. Where carcinogen exposures are known, adducts can be identified by comparing their chromatographic mobilities to standards—an impossibility for unknown substances. Quantification of adducts can be performed reliably by comparison to standards when the adducts have been identified, but not for unknown entities. The shortcomings with immunoassays and ^{32}P-postlabelling emphasize the need for much improved methods to detect DNA adducts, such as the recently developed mass and fluorescent spectrometric techniques.

As an alternative to measuring adducts to DNA, adducts to haemoglobin and serum albumin have been used for molecular dosimetry. Gas chromatographic/mass spectrometric and immunological methods have been developed and applied to these human tissues. Protein adducts are not repaired, and have a relatively slow turnover rate, permitting their accumulation and facilitating their measurement. A decided advantage of protein is its ready availability in relatively larger quantities than for DNA, hence increasing appreciably the sensitivity of such measurements.

2.1.4. EXAMPLES OF CARCINOGEN DOSIMETRY IN HUMANS

Illustrations of the use of molecular dosimetry for four carcinogens are presented below.

Vinyl chloride monomer (VCM) VCM, a recognized human carcinogen, offers an excellent opportunity for cross-species comparisons because: (a) reliable exposure and epidemiological data on VCM are available; (b) it is the only known organic compound that produces liver angiosarcomas in humans; and (c) its mechanism of carcinogenic action and its dosimetry have been investigated in great detail.

Studies in rats have clearly established a dose dependency for VCMs rate of metabolism. A greater percentage of administered dose is metabolized at low versus high doses, leading to a supralinear dose–response for VCM metabolism and carcinogenesis. The carcinogenic effects of VCM appear entirely due to its ultimate metabolite, chloroethylene oxide.

Several DNA adducts are formed in tissues of rats exposed to VCM. The major adduct, 7-(2″oxoethyl)guanine (Oxel-G) causes no miscoding upon replication. Three etheno adducts are formed in tissues exposed to VCM, and each causes base-

pair mismatch on replication. Because the known adducts of VCM and nucleic acid bases involve only N atoms, the genotoxic effects of VCM may result from *N*-alkylation of DNA bases rather than from *O*-alkylation. After cessation of exposure, Oxel-G is removed rapidly from DNA with a half-life in liver of approximately 62 hours; by contrast, the three etheno adducts were stable or poorly repaired, and their levels in the liver were unchanged for up to 14 days.

In neonatal and adult animals, the levels of etheno adducts resulting from exposure to VCM were found to be reliable indicators of tissues at risk of tumour development. The miscoding potential and persistence in DNA of these etheno adducts provide evidence that they may play an important role in VCM-induced carcinogenesis through the production of point mutations. These DNA-etheno adducts, therefore, should be sensitive and reliable markers of molecular dosimetry of VCM-exposed humans and animals. In the same way that the carcinogenic potency of VCM in humans can be estimated from the incidence of liver angiosarcoma in VCM-exposed individuals occupationally, interspecies comparisons made with rats, mice, and hamsters demonstrate that the TD_{50} of VCM was similar, although no correction for differences in metabolism were made.

Dimethylnitrosamine (DMN) Simple methylating agents such as DMN are among the best studied carcinogens for their ability to alkylate DNA. Early investigations demonstrated no correlation between target and non-target tissues for the formation of either 7-methyl-deoxyguanine (7-met-dG), the major adduct formed by DMN, or simple methylating agents such as 4-(*N*-methyl-*N*-nitrosamino)-1-(3-pyridyl)-1-butanone (NNK). Subsequently, attention was focused on O^6-methylguanine, an adduct predicted to cause base-pair mismatch and which persists in brain, not liver, of exposed rats. The presence of this adduct correlated with the target site for tumor induction by these agents. The dose–response for 7-met-dG was different from that of O^6-met-dG in liver DNA. The greater efficiency for repair of O^6 methylguanine (O^6MetG) at low doses is the result of a saturable DNA repair protein, O^6-metG-DNA-methyltransferase. This highly efficient repair system effectively restores the alkylated base to its normal counterpart, and inactivates the protein preventing repair. For molecular dosimetry, when the DNA repair protein is saturated, the amount of O^6-met-dG present in DNA increases per unit dose of carcinogen, causing a nonlinearity in the dose–response curve, i.e., the slope of the dose–response for O^6-metG increases dramatically above a specified dose range. Such data together with those on cell proliferation appear to be vital to the shape of the dose–response curves for cancer induced by these agents. These data imply that the damage caused by methylating agent-mediated DNA alkylation in target tissues of either humans or laboratory animals is modulated by factors such as tissue distribution, metabolism, DNA repair, and cell replication. Thus, the levels of DNA adducts in a human tissue is likely to be different among individuals exposed to the same dose of methylating agent.

The presence of O^6-met-dG and O^7-metG in human liver DNA was described for the first time in two persons who were poisoned by DMN (estimated exposure of

more than 20 mg DMN/kg bw). More recently, O^6-met-dG was detected in surgical specimens of oesophageal and stomach mucosa taken from patients belonging to populations at very different risk of cancer for these anatomical sites. These patients came from the People's Republic of China or France; the nature of methylating agents to which they were exposed is unknown. O^6-Met-dG was also found in placenta and peripheral lung DNA of smoking and non-smoking individuals and in the oral mucosa of DNA from cigarette smokers.

DMN, a multispecies and multi-organ carcinogen, has been studied extensively for its tissue distribution, metabolism, and DNA repair in rodent and human tissues and cells. The preponderance of evidence indicates that dosimetry data obtained in humans were comparable with those measured in organs of species in which DMN produced tumours, implying that DMN may under high-dose circumstances is a carcinogenic risk to humans. To establish causality, DNA adduct levels in individuals need to be correlated with cancer incidence obtained in molecular epidemiology studies, which to date have not been undertaken.

Aflatoxin B_1 (AFB) Exposure to AFB has been associated to the induction of liver tumours in a wide variety of species including humans. Among experimental animals, the relative sensitivity to AFB in causing hepatocarcinogenesis decreases in the following order: trout > rat >> hamster ≈ mouse ≈ salmon. A comparable dose–response relationship holds for the extent of binding of AFB to liver DNA, as illustrated by the observation that hepatic DNA binding is linear over a 10^5-fold dose range for rats treated with a single dose of AFB. A linear correlation has also been observed for the binding of AFB to plasma proteins. In rats and trout, steady-state hepatic DNA adduct levels were observed at approximately 1 month of continuous administration of AFB. In both species, steady-state adduct concentrations is linearly related to the concentration of AFB administered chronically. Furthermore, if the steady-state adduct levels are compared with the incidence of hepatic tumours, a nearly identical linear relationship is obtained for both species. In humans, a highly significant correlation has been observed between the daily consumption of AFB, the concentrations of serum AFB albumin adducts, and urinary AFB nucleic acid adducts.

Ethene and ethylene oxide Ethylene oxide (EtO) is carcinogenic in rats, and has been linked to an increased incidence of leukaemia in exposed workers. EtO exposure can be monitored by analysis of haemoglobin adducts, and this procedure has been applied to workers whose increased exposure to EtO was found to be mirrored by corresponding increases in adduct concentrations.

Ethene is metabolized to EtO, and thereby gives rise to the same adducts in DNA and haemoglobin. Haemoglobin adducts have been increased in cigarette smokers and in occupationally exposed workers; however, low levels of the same adducts have been detected in presumably unexposed individuals. The level of ethene in urban air and environmental tobacco smoke is considered to be too low to account for the adduct levels detected. The designation of "unexposed" individuals is

complicated by ethene's production metabolically in humans and in experimental animals through lipid peroxidation. This endogenous production of ethene is the most likely source of adducts detected in unexposed individuals.

2.2 CROSS-SPECIES SENSITIVITY IN CARCINOGENIC RESPONSE

Carcinogenicity data for 770 compounds administered by various routes of exposure in approximately 3000 rodent bioassays allow comparisons of cancer rates among animal species. For a carcinogen administered by the same route, either the TD_{50} (in mg/kg per day) or the ratio of minimum TD_{50}s can be used to estimate the relative cancer potency between two species (from among rats, mice, and hamsters). The geometric means of the ratios of minimum TD_{50} for rat/mouse was estimated to be 1:2.2 and 1:3.3 for dietary and gavage exposures, respectively. For many compounds, the minimum TD_{50} for each of the three rodent species is generally within a factor of 100 of one another. Such overall agreement provides justification to extrapolate tumor rates across species including from rodents to humans. Consistently high correlations have also been obtained for the potencies of carcinogens in human and laboratory species, and the potencies are similar to those demonstrated between mouse and rat. Interspecies sensitivities appeared to be less than a ratio of 5:1 for both human/mouse and human/rat. Relative sensitivity is equal to $(\text{potency})^{-1}$, with potency equal to 1 in the mouse.

Correlations among cancer potencies of 23 substances (mostly genotoxic) known to cause cancer in animals and humans were examined using epidemiological data and animal bioassays. Potency was expressed as TD_{25} (in mg/kg bw/day). A log-log plot yielded a statistically significant correlation of $r \approx 0.890$ ($p<0.001$). Adjustment for metabolic rate was made by assuming the human–animal equivalent daily dose to be proportional to dose absorbed per square meter of surface area. Adjustment is performed by multiplying animal dose by $(Wt_a \text{ and } Wt_h)^{1/3}$, whereby W_a and W_h represent the average body weight of a species of laboratory animal and humans, respectively). Using this adjustment, humans are expected to be six and 14 times more sensitive than rats and mice, respectively. The available data, although limited by the small number of chemicals studied and by the large errors for human risk estimates, justify cautious extrapolations across species of carcinogenic potency of chemicals; however, measurements of DNA or protein adducts are likely to improve such risk estimates for humans. A major source of uncertainty in the epidemiological analyses on which cancer potency estimates are based stems from poor exposure data, providing an incentive for the future application of molecular dosimetry methods in clinical and epidemiological investigations.

2.3 APPROACHES TO PREDICT CANCER ACTIVITY AND POTENCY

2.3.1 STRUCTURE–ACTIVITY RELATIONSHIPS

Some chemicals induce carcinogenic effects in mammals, while others are inactive. Some aspects of chemical structure, therefore, may be causally associated with chemical carcinogenesis, thereby providing an opportunity for the study of structure–activity relationships (SAR). Models for SAR rely on defined toxicity caused by certain molecular fragments or structures to predict carcinogenicity among chemicals for which no carcinogenicity data exist. To reliably estimate which compounds are more likely to cause cancer, the amount of data that must be analyzed is large; therefore, these data must be analyzed using computerized systems, particularly those that rely on artificial intelligence. Meaningful correlations between chemical structure and carcinogenic potency, however, do not imply causality, and, therefore, extrapolations from such correlations are not always possible.

Given those basic premises, the next step is differentiation of information at the chemical and biological levels. Relationships between chemical reactivity (as a surrogate of carcinogenicity) and structure should be established for the most likely chemical classes. This foundation should then be used to synthesize a flexible multiclass generic SAR model.

2.3.1.1 Fragmentation of the chemistry

A chemical is defined by its structure. Yet when a chemical elicits a response in a biological system, one or more of its many physico-chemical attributes is responsible. Understanding the systematic way in which these attributes function in biological systems is the foundation for deriving a causal relationship between structure and activity. Such attributes include:

1. Substructural fragments. The simplest and most often used parameter, key substructures such as $ArNO_2$, CH_2Cl, and ring oxide, are recognized for their high correlation with biological activity. These methods are capable of being self-learning and self-correcting. For example, *p*-chloronitrobenzene will initially be classed both as ArCl and as $ArNO_2$; but if the database is self-learning, it will soon associate mutagenicity with $ArNO_2$ and remove an association for ArCl.
2. Molecular shape. Current computerized correlative SARs methods represent structure in two dimensions; however, for some activities (e.g., a receptor interaction), 3-dimensional molecular shape may be the primary determinant. Furthermore, steric effects can dramatically modulate reactivity; e.g., neopentyl bromide is unreactive due to crowding of its bromomethyl group, and, therefore, is not a simple analog of ethyl bromide.
3. Solubility and partition properties. Absolute solubility and log P, which are essential to determine bioavailability, can be either estimated or measured.
4. Chemical stability. Rates of hydrolysis and degradation of a parent

compound or its metabolite(s) can be a major determinant of biological activity (e.g., local versus systemic effects; effects seen *in vitro* versus *in vivo*).

5. Chemical classes. Chemical classes can be designated by the capacity for certain types of reactions, (electrophiles, alkylating agents, or Michael reactivity); however, not all electrophile–nucleophile pairs are equally reactive. Thus, the Swain–Scott equation provides data on relative O, N, S reactivity via *s* values. Extension to aromatic amine nitrenium ions can be achieved. Acrylamide, a Michael reactive agent, is reactive primarily with protein.
6. Metabolic activation/deactivation. The potential for metabolic activation and deactivation can be critical to SARs, but are difficult to predict.

2.3.1.2 Fragmentation of biological parameters

All tissues of all species are not equally susceptible to carcinogenesis. Some of the attributes include:

1. Species, sex, tissue specific metabolism. As data on these variables accrue, they will enhance databases. Furthermore, differences between these and those from S9 fractions will resolve many inconsistencies and explain the absence of expected correlations.
2. Detoxification. This parameter may have a greater impact *in vivo* than *in vitro*.
3. Response to DNA adducts. Not all DNA adducts lead to the same biological changes. Existing data could refine biological data, such as differences in: (a) key enzymes between tissues or species; (b) DNA repair pathways; (c) rates of repair between tissues or species; and (d) DNA sequence specificity for adduct formation and repair.
4. Tumor classification. If tumors are listed by site, species, and sex, more refined SARs are likely to result.
5. Non-genetic toxicity. Such data are usually available, and may contribute key information to improve or refine correlations. Thus, liver or thyroid toxicants predispose to cancer at those respective sites. The hormonal properties of a chemical or its mitogenicity are also potentially important parameters.
6. Non-genotoxic carcinogenesis. A generally accepted concept worthy of separate treatment from electrophilic carcinogens, non-genotoxic carcinogens (e.g., non-mutagenic mouse liver specific carcinogens) have proven difficult to predict, thereby requiring a different SAR approach than that used for genotoxic carcinogens (e.g., mutagenic epoxide skin carcinogens).
7. Genetic activity profiles. Such profiles can provide a valuable means of biological recognition of unexpected chemical classes or biological patterns.

Consensus correlations Methods exist by which a large volume of data can be

merged so that a chemical with divergent activities can be categorized by consensus to either a probable genotoxicant or a non-genotoxicant, and likewise to a probable carcinogen or non-carcinogen. However, such SARs are less precise than those derived within a specific chemical class and having a documented biological endpoint.

2.4 GENOTOXICITY PROFILES IN EUKARYOTES

Progress in understanding how genotoxicants exert their biological effects in diverse species has been impeded by the lack of a concept by which physicochemical properties are linked to mutation spectra. The "single-endpoint" approach (i.e., relating induced genetic damage to exposure levels) is of limited value to predict relative carcinogenic potency of chemicals in mammals, despite its wide application in more than 150 short-term tests for mutagenicity. For example, attempts to correlate relative carcinogenic potency of chemicals with qualitative mutagenicity data obtained from a single test was an oversimplification of a complex matter, as demonstrated repeatedly for a wide array of genotoxic carcinogens also tested in the *Salmonella* assay.

Multi-endpoint analysis in eukaryotes combined with SAR considerations provides a more useful tool to categorize genotoxicants with known mechanism(s) of adversely affecting DNA. Fundamentally, this approach compares various endpoints with each other (e.g., gene mutation induction with clastogenic effects) and evaluates genotoxicity profiles against the following parameters: (a) nucleophilic selectivity as expressed by Swain–Scott's *s* values and/or N^7/O^6-alkyl dG ratios; and (b) relative carcinogenic potency (e.g., TD_{50} estimates in mg/kg bw) against nucleophilic selectivity. In addition, either modifying DNA repair conditions or assessing enhanced genotoxicity (e.g., hypermutability) in relation to SAR parameters can provide insight as to the role of various promutagenic lesions and to the significance of repair in mutation fixation.

In particular the Swain–Scott empirical linear relationship has proven valuable in not only understanding but also predicting the great diversity of genotoxicity profiles exhibited by numerous alkylating agents. Two sets of data emerged from multi-endpoint analysis of 60 genotoxic chemicals, mostly alkylating carcinogenic, in rodents and *Drosophila*: (a) carcinogenic potency and hypermutability were correlated linearly with the *s* value; (b) monofunctional arachidonic acid that can alkylate oxygen atoms had lowest TD_{50} values; and (c) genotoxic agents capable of cross-linking DNA could be separated clearly from those giving the monoadduct. Thus, these multi-endpoint analyses should assist in the quantitative evaluation of risk from genotoxic agents.