META. 2. A Dictionary Model of Mammalian Xenobiotic Metabolism

Joseph Talafous,† Lawrence M. Sayre,† John J. Mieyal,‡ and Gilles Klopman*,†

Chemistry Department, Case Western Reserve University, Cleveland, Ohio 44106-7078, and Department of Pharmacology, Case Western Reserve University, Cleveland, Ohio 44106-4965

Received April 14, 1994⊗

META is a new knowledge-based expert system that provides computer simulation of the biotransformation of chemicals. The program is based on the recognition of key functional groups within the complete chemical structure and therefore can predict the metabolites of new xenobiotics. Here, we describe a comprehensive knowledge base built for the purposes of modeling mammalian metabolism with META methodology.

INTRODUCTION

In the preceding paper in this issue,[1] the underlying methodology of the META program was described as well as the concept of a dictionary which acts as the knowledge base. Used with the appropriate dictionary, the META program applies chemical transformation rules, called *transforms*, which recognize and substitute chemical functional groups. The inputs to META are line-coded chemical structures, and the output is a graphical display of possible metabolites together with an overall map of the metabolic pathways traversed. This paper describes the construction of a dictionary that accurately models the pathways known to be involved in the metabolism of xenobiotics in mammals.

Many pharmacologic effects of foreign compounds, especially drug efficacy and toxicity, are intimately related to metabolic biotransformations. The identification of the structures of the metabolites is among the first steps in delineating the molecular basis for pharmacological effect. The use of META methodology coupled to a comprehensive dictionary provides the needed collection of possible metabolites. META may predict metabolites that are overlooked or even unexpected by the human expert, especially if finer-detailed and less-encountered metabolic knowledge is included in the dictionary. An accurate apportionment of pharmacological efficacy can be accomplished only after all the metabolic species are defined.

Mammalian metabolism lends itself to modeling by META methodology for two reasons. First, examples of metabolic activation of xenobiotics in mammals are abundant and provide a great amount of knowledge from which to design transforms. Second, the enzymes involved in xenobiotic metabolism must be fairly nonspecific toward their substrates, otherwise the number of specific enzymes required would be enormous and solvent capacity[2] would be exceeded. This broad substrate specificity allows META transforms to extrapolate accurately, providing the user with a powerfully predictive tool.

METHOD

Since the utility of META is largely determined by the contents of the dictionary, the philosophy of its design was carefully considered. Comprehensiveness was a primary design objective, in contrast to other published programs in this field[3−5] that model a limited and specific knowledge domain. Our dictionary incorporated only well-established metabolic data from review, textbooks,[6−12] and monographs.[13−15] Available metabolic literature pertains to any of several mammalian species. Rodent-specific pharmacological data was preferred, and rodent liver metabolism was the most common reference, but data from any mammalian source was substituted if the rodent data was unobtainable. Hence, we do not claim that the resulting dictionary is an exhaustive collection of biotransformation rules for any particular mammal. Rather, our goal was to model a theoretical "average" mammal, in which the resulting dictionary would provide an excellent starting point for the development of more specific dictionaries that pertain to various species and organs.

An initial dictionary was carefully constructed that closely parallels the present consensus of xenobiotic metabolic knowledge by the authors (J.J.M. and L.M.S.) who are experts in the field of medicinal chemistry. The META transforms focused on chemical substructures and were organized around pharmacological *reaction type*, i.e., O-dealkylation, N-hydroxylation, etc. The reaction types were associated with *enzyme systems* that are known to catalyze such reactions. In many cases, more than one type of enzyme can carry out the same reaction type, and each enzyme system may include several different reaction types. This method of classification allows, but does not require, parameterization of enzyme−substrate lipophilicity relationships, pharmacological compartment location, etc. It continues to serve as an adaptable framework for the addition of new metabolic data. Each enzyme system belongs either to phase 1 (functionalization) or phase 2 (conjugation) metabolism. The transforms in the initial dictionary were then intuitively prioritized.

The consensus knowledge provides general guidelines of possible metabolic pathways to follow, but the actual pathways traveled are far more specific to the structure of the substrates. Accordingly, the initial dictionary was subsequently improved by assimilation of specific biotransformations that had not been predicted accurately. An iterative refinement of the dictionary using test xenobiotics was performed. The initial dictionary was tested with 50 xenobiotics that were selected for their well-defined metabolism, which served as a training set for guiding dictionary

* To whom all correspondence should be addressed.
† Chemistry Department.
‡ Department of Pharmacology.
⊗ Abstract published in *Advance ACS Abstracts,* September 15, 1994.

0095-2338/94/1634-1326$04.50/0

refinement. For each xenobiotic, the performance of individual transforms was determined by comparing the actual experimental metabolic status (observed vs not observed) against the META predictive status for each transform (hit vs miss).

Depending upon performance, transforms were debugged as necessary so as to increase the predictive accuracy of the latest dictionary version. Since generalizability of the transforms is the source of the predictive power of META, transforms were constructed to recognize the smallest substrate fragment, yet still generate products that are consistent with experimental observations. This struck a good balance in the context of predicting chemical toxicity and drug efficacy, where a significantly higher cost would result for not predicting an actual metabolite as compared to suggesting a nonexistent metabolite. In addition, reducing the size of transforms increases processing speed and lowers memory requirements. Transform prioritization was adjusted as necessary on the basis of observed pharmacokinetics of several test molecules. For enzymes whose transform efficiency changes with substrate lipophilicity (e.g., cytochrome P-450[12]), we used a log *P* function to prioritize the transforms.

THE DICTIONARY

Both phase 1 and phase 2 of xenobiotic metabolism were modeled with META methodology. A given xenobiotic usually passes through both phases, but xenobiotics may undergo either phase or neither phase before elimination. Given the breadth of both phases of mammalian metabolism, our comprehensive META dictionary was required to be extensive, so a complete listing here is beyond the scope of this paper. Instead, we present the essential activity of each reaction type that has been encoded into META transforms, providing xenobiotic archetypes. A representative transform for each reaction type is provided. Biotransformation reaction types for each phase of xenobiotic metabolism are listed in Tables 1 and 2. Presently, we have constructed over 750 fundamental transforms based on pharmacological data for more than 150 xenobiotics.

We use the reaction type of aliphatic hydroxylation (a phase 1 biotransformation, see below) to demonstrate the mechanics of META (Figure 1). The principal metabolite of the sedative pentobarbital is that obtained by hydroxylation of its *sec*-pentyl side chain at the ω-1 position. The solvent hexane undergoes an analogous reaction which is modeled by the same transform. Note that a single META transform models both biotransformations.

Phase 1 Biotransformations. Phase 1 biotransformations usually metabolize the xenobiotic first, generating polar functional groups (OH, NH_2, COOH) that are further derivatized in phase 2 reactions (see below). However, the metabolites of phase 1 biotransformations are frequently not excreted, even though they may be less lipophilic. Phase 1 biotransformations include oxidative, hydrolytic, and reductive reactions and generally diminish lipophilicity. A product of phase 1 biotransformations may have higher or lower pharmacological effect than the substrate, or it may exhibit an altogether different effect such as toxicity. Table 1 lists phase 1 reaction types that have been modeled by META transforms.

The cytochromes P-450 monooxygenase isozymes[16,17] are central to phase 1 xenobiotic metabolism, catalyzing several

Table 1. Phase 1 Biotransformations[a]

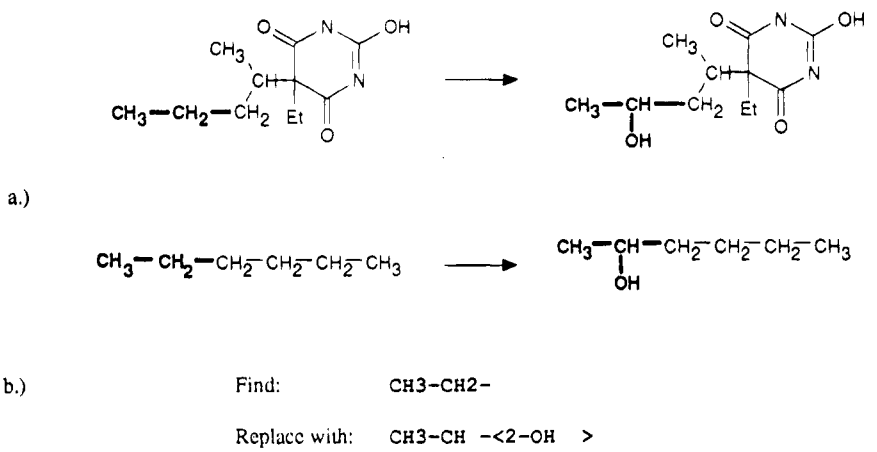
Reaction type	Example of META transform	Example substrate	Example enzyme system
Aromatic hydroxylation	CH =CH - C =CH -<1-OH >	benzene	P-450
Aliphatic hydroxylation	CH3-CH2- CH3-CH -<2-OH >	hexane	P-450
Epoxidation	CH =CH - CH -CH -<1-O -2>	benzo[a]pyrene	P-450
N-oxygenation	CH2-N -CH3-<2-C -> CH2-NO -CH3-<2-C ->	imipramine	P-450 and flavin monooxygenase
N-hydroxylation	NH2-CO - NH -CO -<1-OH >	urethane	P-450
O-dealkylation	CH"-O -CH2- CH"-O -CH -<3-OH >	phenacetin	P-450
S-dealkylation	CH"-S -CH3 CH"-S -CH2 <3-OH >	6-methylthiopurine	P-450 and flavin monoxygenase
N-dealkylation	CH3-N -CH3 CH3-N -CH2 <3-OH >	dimethylformamide	P-450
Deamination	NH2-CH - NH =C -	amphetamine	monoamine oxidase
Sulfoxidation	S -S - SO -S -	disulfiram	flavin monooxygenase
Desulfurization	CH2-O -P =S <2-O -> CH2-O -P =O <2-O ->	parathion	P-450
Oxidative dehalogenation	X -C -X X _C =O	halothane	P-450
β-oxidation	CH2-CH2-CO -OH CO CH3-CO -OH <1-OH >	phenylalkyl acids	mitochondrial β-oxidation
Ester hydrolysis	CO -O -CH2- CO -OH CH2-<3-OH >	procaine	esterases
Amide hydrolysis	NH -CO - NH2 CO -<2-OH >	phenacetin	amidases
Epoxide hydrolysis	CH -O^ -CH - CH -OH CH -<3-OH >	naphthalene 1,2-oxide	epoxide hydratase
Alcohol oxidation	OH -CH2- O =CH -	ethanol	alcohol dehydrogenase
Aldehyde oxidation	CH3-COH CH3-CO <2-OH >	acetaldehyde	aldehyde oxidase
Purine oxidation	N =CH -NH - N =C -NH -<2-OH >	theophylline	xanthine oxidase
Carbonyl reduction	COH-C - CH2-C -<1-OH >	trichloro-acetaldehyde	carbonyl reductase
Azoreduction	C =C -N =N -C = C =C -NH2 NH2-C =	prontosil	flavin reductase
Nitroreduction	NO2-C = NO -C =	chloramphenicol	flavin reductase

[a] Special characters in the transforms are as follows. The caret symbol (^) represents an atom in a three-membered ring, the double quote (") represents an sp^2 carbon, and X represents a halogen.

reaction types that are represented by a total of 197 transforms. They are perhaps the most versatile biological catalysts known and are the most powerful *in vivo* oxidizing agents. Cytochromes P-450, which are often the first line of attack on lipophilic, nonpolar substances, have extremely broad substrate specificity. There are thousands of chemical compounds that are known to be metabolized by cytochrome P-450. The overall behavior of cytochrome P-450 is well-defined,[13] allowing for simple coding by META transforms. All the reactions of cytochrome P-450 insert a single atom of oxygen into the xenobiotic, presumably by a free radical mechanism, to form a hydroxy or epoxide moiety.[18]

Aliphatic hydroxylation[15] by P-450 is represented by 43 transforms and one type (ω-1) was discussed above (Figure 1). Our dictionary contains transforms for hydroxylation of methyl, methylene, and methine sp^3 carbons in an aliphatic chain. Aliphatic substituents on aromatic rings (e.g., toluene, tolbutamide, and methaqualone) are generally hydroxylated at the methylene adjacent to the aromatic ring because the phenyl ring activates this position.

Aromatic hydroxylation generally proceeds via an epoxide intermediate, but direct hydroxylation via an electrophilic aromatic substitution mechanism occurs for some aromatic substrates.[19] The epoxide intermediate may either be relatively stable or spontaneously rearrange to the phenol, as indicated by the NIH shift. The electrophilic epoxide may be hydrolyzed, or it may be adducted with glutathione, or it may bind covalently to macromolecules via reaction with nucleophilic sites on proteins or DNA. The latter is one type of covalent binding to DNA that has been implicated in chemical carcinogenesis.[20] Specific transforms for the generation of transannular oxygenation, such as in the



a.)

b.) Find: CH3-CH2-

Replace with: CH3-CH -<2-OH >

Figure 1. (a) Biotransformations of pentobarbital and hexane, in this case aliphatic hydroxylation. (b) A META transform that models the biotransformation. It recognizes the ethyl substrate fragment and replaces it with the corresponding hydroxylated product fragment. The substrate fragment is kept as small as possible so that the transform will apply to xenobiotics outside the training set. (Branching groups are located between angle brackets. The first numerical inside indicates the atom of the main chain to which the branching group is connected.)

generation of naphthalene 1,4-oxide, were also written. The analgesic phenacetin, the antidepressant imipramine, benzene, and particular molecular positions on polycyclic aromatic hydrocarbons[15] (e.g., benzo[*a*]pyrene) are epoxidized by P-450. The sensitivity of aromatic hydroxylation/epoxidation to substituent effects prompted us to extend modeling to use quantum mechanical indices and graph indices, which will be described fully in a forthcoming publication.

For aliphatic ethers, sulfides, halides, and amines, hydroxylation by P-450 generally occurs at the carbon adjacent to the heteroatom which is usually followed by spontaneous dissociation of the hydroxylation product to release the substituted heteroatom and generate an aldehyde or ketone as the second product. The net biotransformation is oxidative cleavage of C—X bonds, where X can be oxygen, sulfur, halogen, or nitrogen. A hydrogen α to the heteroatom in the substrate is essential. Twenty-eight transforms model *O-dealkylation* of aromatic and aliphatic ethers (e.g., the conversion of codeine to morphine and phenacetin to acetominophen). Sixteen transforms model *S-dealkylation* with the thyroid inhibitor 6-methylthiouracil being an example substrate. *Dehalogenation*[21] of insecticides and commercial solvents is a common metabolic pathway which is modeled by 30 transforms. Halogenated alkanes[15] such as the anaesthetic halothane are dehalogenated by P-450. Eighty transforms were constructed to model *N-dealkylation*, a frequently encountered biotransformation which acts upon substrates such as ethylmorphine, dimethylformamide, and imipramine.

Deamination is the same as N-dealkylation except that the former term pertains to primary amines and can occur through at least three enzyme systems: microsomal cytochrome P-450, mitochondrial flavin-dependent monoamine oxidase (MAO),[22] and the copper amine oxidases.[22] The P-450 reaction involves α-hydroxylation, whereas the MAO reaction proceeds through dehydrogenation to an imine which is subsequently hydrolyzed.[13] Since there must be a hydrogen α to the nitrogen, anilines are poor substrates. Catecholamine neurotransmitters are deactivated via deamination by MAO. MAO mainly dehydrogenates aliphatic primary amines, but secondary and tertiary allylic amines are also metabolized. Fifteen META transforms modeled deamination by MAO. Two other transforms represent the copper amine oxidases,[23] an enzyme system including plasma amine oxidase, semicarbazide-sensitive amine oxidase, and diamine oxidase whose substrate specificity overlaps with MAO. The copper amine oxidases contain a quinone cofactor and metabolize strictly unbranched alkyl and aralkyl primary amines. Histamine, putrescine, and cadaverine undergo deamination by diamine oxidases.

***N*-Hydroxylation and *N*-oxygenation**[24] are catalyzed by flavin monooxygenase (FMO) and/or cytochrome P-450, with overlapping substrate specificities. Xenobiotics with no α-hydrogen, such as tert-alkylamines (e.g., phentermine), primary arylamines (e.g., aniline, benzidine), primary arylamides (e.g., phenacetin), and hydrazines (e.g., isopropylhydrazine, procarbazine), are principally N-hydroxylated by P-450. Aromatic heterocyclic amines[24] (e.g., 3-methylpyridine) are also N-oxygenated by P-450. FMO N-oxidizes strictly aliphatic amines. Virtually all tertiary amines (e.g., trimethylamine, cocaine, nicotine, the antihypertensive pargyline, and tranquilizer perazine) are potential substrates for N-oxygenation by FMO, generating amine *N*-oxides.[25] Secondary amines are N-oxidized by both enzymes.[9] The hydroxylamine products can be subjected to repeated FMO N-oxygenation to give nitrones (from secondary amines) and oximes (from primary amines). 2-Acetylaminofluorene undergoes N-hydroxylation which is required for activation to the ultimate carcinogen.[26] Sixty-two META transforms were constructed which model N-hydroxylation and N-oxygenation.

Sulfoxidation is almost exclusively[9] catalyzed by FMO, which oxidizes a broad variety of divalent sulfur compounds. Sulfoxidation[27] is more common than S-dealkylation. Thirty-six transforms modeled the biotransformation of alkyl sulfides (e.g., chlorpromazine, cimetidine); disulfides (e.g., disulfiram); thiols (e.g, thiophenol); sulfenic, sulfinic, and sulfonic acids;[25] and thioamides and thiocarbamides (e.g., thionicotinamide, methimazole[28]). Thiocarbamides and thioamides (e.g., thioacetamide) undergo sequential oxidation by FMO, the products of which are subject to tautomerization. Several pesticides (e.g., disulfoton) are excellent substrates for S-oxidation by FMO in vitro. *Desulfuration*, which occurs when a sulfur atom is replaced by oxygen, is catalyzed by both P-450 and FMO. The insecticide parathion and the solvent carbon disulfide[15] undergo desulfuration.

Alcohol oxidation is catalyzed mainly by the enzyme alcohol dehydrogenase,[13] which is modeled with four META transforms. Alcohol dehydrogenase has broad specificity and readily oxidizes most primary and secondary aliphatic alcohols and aralkyl alcohols,[29] while tertiary and sterically hindered secondary alcohols are poor substrates. In particular, ethanol is oxidized to acetaldehyde.[29] Alcohol dehydrogenase is also known to reduce aldehydes and aromatic ketones.[9]

Aldehyde oxidation of aliphatic and aromatic aldehydes is catalyzed by aldehyde dehydrogenases.[30] Metabolites of MAO and alcohol dehydrogenase usually undergo aldehyde oxidation. Seven META transforms were needed to model the broad specificity of aldehyde dehydrogenase. Acetaldehyde and 5-hydroxyindole acetaldehyde are examples of substrates. Many pesticides (e.g., aldicarb) are oxidized to acids by aldehyde dehydrogenases.[13]

Purine oxidation is catalyzed by xanthine oxidase and aldehyde oxidase.[13] These two activities actually represent several different species of enzymes that incorporate an atom of oxygen into the product from water rather than from molecular oxygen, as do P-450 and FMO. This is a characteristic of molybdenum flavoheme enzymes that use FAD. Both enzymes oxidize nitrogen-containing heterocyclic compounds as well as aliphatic and aromatic aldehydes. Four META transforms were required to model their broad and overlapping substrate specificities. The antineoplastic agent 6-mercaptopurine and the antiasthmatic theophylline undergo purine oxidation to their uric acid derivatives. In a more unusual example, the nicotine $\Delta^{1'(5')}$-iminium ion is converted to the lactam cotinine by aldehyde oxidase.[31]

***β*-Oxidation** of xenobiotic straight-chained alkyl carboxylic acids as well as endogenous fatty acids occurs through the same metabolic pathway. One META transform was required to model the oxidative cleavage of two carbon units from the substrate. Phenylalkyl acids undergo *β*-oxidation.[9]

Ester hydrolysis[32] is catalyzed by a wide range of enzymes (e.g., acetylcholinesterase, carboxylesterase) with different substrate specificities. Twenty META transforms represent hydrolysis of aromatic and aliphatic esters. The pharmacological effects of procaine and atropine are terminated upon hydrolysis of their ester bond. Thioesters (e.g., phenyl thioacetate) are also subject to hydrolysis.

Amide hydrolysis[32] is similarly catalyzed by numerous enzymes with broad specificity. Amides are generally more stable to hydrolysis than their analogous esters.[9] Two META transforms represent the hydrolysis of all amides. The amide bond in the antiepileptic phenytoin and the anesthetic lidocaine are subject to hydrolysis by amidases. Hydrolysis of the amide bond in phenacetin generates products that may induce methemoglobinemia and renal toxicity.[33]

Epoxide hydrolysis, catalyzed by epoxide hydratases,[34] generates 1,2-diols. Eleven META transforms were needed to model this highly nonspecific reaction type. The substrates are likely to be the epoxide products of cytochrome P-450 monooxygenase. Hydrolysis of benzo[*a*]pyrene 7,8-oxide produces a precursor for a mutagen/carcinogen.[35]

Carbonyl reduction of aldehyde and ketones was simulated with 20 META transforms. Carbonyl reductases include both aldehyde[36] and ketone reductases[36] which do not overlap in substrate specificity. Carbonyl-containing xenobiotics which are not oxidized are reduced to alcohols and thereby supply a site for conjugation (see phase 2 reactions below). Aldehyde reductases do not appear to be involved in endogenous metabolism and show broad specificity for aromatic, aliphatic, and sugar aldehydes. Ketone reductases reduce aliphatic, unsaturated, and aromatic ketones. Common substrates are quinones, aromatic aldehydes, and ketoaldehydes.

Table 2. Phase 2 Biotransformations in Mammals[a]

Phase 2 reaction type	Example of META transform	Example substrate	Example enzyme system
N-glucuronide conjugation	NH2-C = NH -C =<1-Z4 >	aniline	N-glucuronyl transferase
O-glucuronide conjugation	OH -C = O -C =<1-Z4 >	salicylamide	O-glucuronyl transferase
O-sulfate conjugation	OH -C = O -C =<1-SO2-OH >	acetaminophen	phenol sulfotransferase
N-sulfate conjugation	NH2-C = NH -C =<1-SO2-OH >	2-naphthylamine	arylamine sulfotransferase
N-acetylation	NH2-NH - NH -NH -<1-CO -CH3 >	isoniazide	N-acyltransferase
N,O-transacetylation	OH -NH -CH - O -NH -CH -<1-CO -CH3 >	N-hydroxy-2-acetylaminofluorene	N,O-acyltransferase
Glutamine conjugation	OH -CO -C = NH -CO -C =<1-Z2 >	benzoic acid	glutamyl transferase
Glycine conjugation	OH -CO -C = NH -CO -C =<1-CH2-CO -OH>	salicylic acid	glycyl transferase
N-methylation	NH2-CH2- NH -CH2-<1-CH3 >	phenylethanolamine	phenethylamine N-methyl
O-methylation	C" -OH C" -O <2-CH3 >	catechol	COMT
S-methylation	SH -C =CH - S -C =CH -<1-CH3 >	2-thiouracil	S-methyl transferase
Glutathione conjugation	CH -O^ -CH - CH -Z3 CH -<3-OH >	styrene oxide	glutathione S-transferase

[a] Z2 represents glutamine, Z3 represents glutathione, and Z4 represents glucuronic acid. See Table 1 for explanation of other special symbols.

Nitroreduction is catalyzed by flavin-containing reductases, with cytochrome P-450 reductase being an example. Cytochrome P-450 reductase reduces nitro compounds (e.g., 2-nitrofluoranthene, chloramphenicol, and nitrobenzene) to their respective nitroso form. META transforms have also been included that model the further reduction of nitroso to hydroxylamine and then to amine.

A META transform representing *azoreduction* of the dye prontosil and sulfasalazine exists in our dictionary. The biotransformation of the azo functionality to aromatic primary amines occurs by both the intestinal microflora and by cytochrome *c* (P-450) reductase.[37] Another reductase, DT diaphorase,[38] reduces the red food dye amaranth.

Phase 2 Biotransformations. Also termed conjugation reactions[39] or synthetic reactions, phase 2 biotransformations are listed in Table 2. Phase 2 biotransformations attach hydrophilic carbohydrate, amino acid, peptide, or sulfate groups which ordinarily remove bioactivity and convert the xenobiotic to more excretable forms. However, methylation and acetylation of heteroatom-containing substrates also occur, which do not necessarily enhance water solubility or excretability. The limited number of conjugation centers (-COOH, -OH, $-NH_2$, -SH, epoxides, and other electrophilic centers) makes prediction of phase 2 reactions by META reliable. At least one phase 2 reaction will probably occur if a conjugation center is present.

***N*- and O-Glucuronic acid conjugation**[40] is the most common phase 2 detoxication pathway. Seventeen transforms were written for phenol, alcohol, and carboxylic acid substrates. Salicylamide is excreted mainly as the glucuronide. The resulting adduct is quite hydrophilic and is usually excreted in the urine and bile immediately, unless first deconjugated by intestinal flora with subsequent enterohepatic cycling.

If not glucuronidated, the hydroxyl group usually undergoes *O-sulfation*.[13] Sixteen transforms were written for O-sulfation of aliphatic and aromatic alcohols[15] and hy-

droxylamines. The *N,O*-sulfate product of the carcinogen *N*-hydroxy-2-(acetylamino)fluorene is much more toxic than its substrate. Many substituted phenols such as 3-hydroxycoumarin[41] undergo O-sulfation. One transform models aromatic amine *N-sulfation* with aniline and sulfanilamide being example substrates. Sulfation reactions are catalyzed by several sulfotransferases.

***N*-Acetylation** is the major route of aromatic amine (e.g., procainamide, sulfisoxazole) metabolism for most species. N-acetylation of the sulfonamido group in sulfanilamide results in a compound that is less soluble in both water and lipid, which may precipitate in the kidney as it is concentrated in the urine. Fourteen transforms were used to model N-acetylation of amines. *N*-acetyltransferase[42] has a broad specificity toward compounds with an amino group connected to unsaturated rings or carbonyl groups. Secondary amines are not acetylated.[9] Hydrazines[15] such as isoniazid undergo N-acetylation. Six transforms represent *N,O-transacetylation*,[13] which is catalyzed by *N,O*-acyltransferase. Aromatic amines are thought to induce tumors as a consequence of their N-oxygenation followed by O-acetylation. Hydroxyamines and hydroxamides are the primary substrates.

Glycine conjugation[13] forms an amide bond involving the α-amino group of glycine and an aliphatic or aromatic carboxylic acid that has been activated by conversion to the thioester with coenzyme A. Five transforms were encoded for biotransformation of primary, secondary, and tertiary aliphatic acids and aromatic acids. Benzoic acid, salicylic acid, nicotinic acid as well as medium chain alkyl carboxylic acids are conjugated with glycine.

Glutamine conjugation[13] forms an amide bond with the α-amino group of L-glutamine and seems to occur exclusively in primate species.[43] In humans, phenylacetic acid is conjugated with L-glutamine instead of glycine as in other species. Eight META transforms were used to model conjugation of aliphatic and aromatic carboxylic acids. Mescaline is known to conjugate with glutamine.

Glutathione (GSH) conjugation, catalyzed by glutathione *S*-transferase, is the principal route of phase 2 metabolism of reactive electrophiles (e.g., halides, Michael acceptors, and epoxides). Following GSH conjugation, the glutamic acid and glycine termini of the GSH tripeptide are hydrolytically removed, and the resulting cysteine conjugate is N-acetylated to afford mercapturic acids that are excreted. Although glutathione conjugation is usually not quantitatively significant in xenobiotic elimination, it protects proteins and nucleic acids from electrophilic addition. Reactive epoxides that are generated by cytochrome P-450 are removed in this way. Seventy-one META transforms were used to model GSH conjugation of epoxides (e.g., styrene epoxide), α,β-unsaturated compounds (e.g., iminoquinone), alkyl halides (e.g., methyl iodide) vinyl halides (e.g., tetrachloroethylene), and isocyanates (e.g., methyl isocyanate). Glutathione conjugate formation is an important bioactivation mechanism for several groups of compounds.[44] Two META transforms were written for cysteine conjugate β-lyase,[13] which liberates acetic acid from the cysteine conjugate. Two META transforms were written for glyoxylase,[13] which conjugates GSH as an intermediate in reduction of dicarbonyl compounds.

Methylations[13] are important because they often significantly change the pharmacological activity of the substrate. Most *N-methylations* are catalyzed by a family of specific *N*-methyltransferases, such as indoleethylamine *N*-methyltransferase. N-Methylation occurs mainly in the biosynthesis or metabolism of naturally occurring endogenous compounds such as histamine, estradiol, thyroxine, norepinephrine, dopamine, and serotonin.[45] Several META transforms were written to model N-methylation for aliphatic and aromatic primary, secondary, and tertiary amines, including cyclic amines such as normorphine. Pyridines (e.g., nicotinamide, nicotinic acid) are N-methylated to quaternary *N*-methylpyridinium salts.

Six META transforms represent *O-methylation* of phenols catalyzed by a family of monohydric, dihydric, and trihydric phenol *O*-methyltransferases, which have broad specificity with regard to o,p-substitution on the aromatic ring.[46] Caffeic acid is O-methylated. O-methylations of endogenous catechols such as epinephrine and other neurotransmitters occur primarily via catechol *O*-methyltransferase (COMT).

EXERCISING THE DICTIONARY

After its development and refinement, we needed to test the dictionary under conditions consistent with actual pharmacological use. The chief use of META is anticipated to be the prediction of all significant metabolites of a given xenobiotic. The first exercise demonstrated that META could accurately predict the metabolism of the relatively simple drug molecule phenacetin in mammals[33] (Figure 2). Its major primary metabolite is the therapeutically active product acetaminophen formed via O-dealkylation by the P-450 enzyme system. Deacetylation via hydrolysis by amidases is usually second in prominence. The other two P-450-mediated reactions, epoxidation and N-hydroxylation, are minor pathways leading to potentially toxic metabolites.

The metabolic profile for phenacetin expands broadly as secondary biotransformation pathways are considered. Examples of enzyme-mediated reactions as well as nonenzymatic reactions are illustrated in Figure 2. Acetaminophen and *N*-hydroxyphenacetin are hydrolyzed by amidases to *N*-hydroxyphenetidine and *p*-aminophenol, respectively. Acetaminophen and *p*-aminophenol undergo autoxidation to their iminoquinone form, which is subsequently conjugated with glutathione. Phenacetin-2,3-oxide spontaneously rearranges to 2-hydroxyphenacetin, which is modeled by a separate set of spontaneous reaction transforms. (This set contains over 400 fundamental transforms and will be described fully in a subsequent publication.) These two reactions illustrate the potential coupling between spontaneous reactions (ring opening, autoxidation) and enzyme-mediated reactions (glutathione conjugation). Both the primary and secondary metabolites of phenacetin were correctly predicted by META.

The second pharmacological exercise had a larger focus, insofar as we were interested to determine whether META could generate (among others) the most prominent metabolites for a large set of "classic" xenobiotics which are still commonly encountered or in use today. Table 3 lists the compounds arbitrarily selected from R. T. Williams' book[47] and review,[48] which were mostly outside the training data. The META program was run with our dictionary and correctly predicted the major biotransformation for *all* selected chemicals with one exception. Ethyl ether is excreted unchanged via the lungs, yet this was not predicted

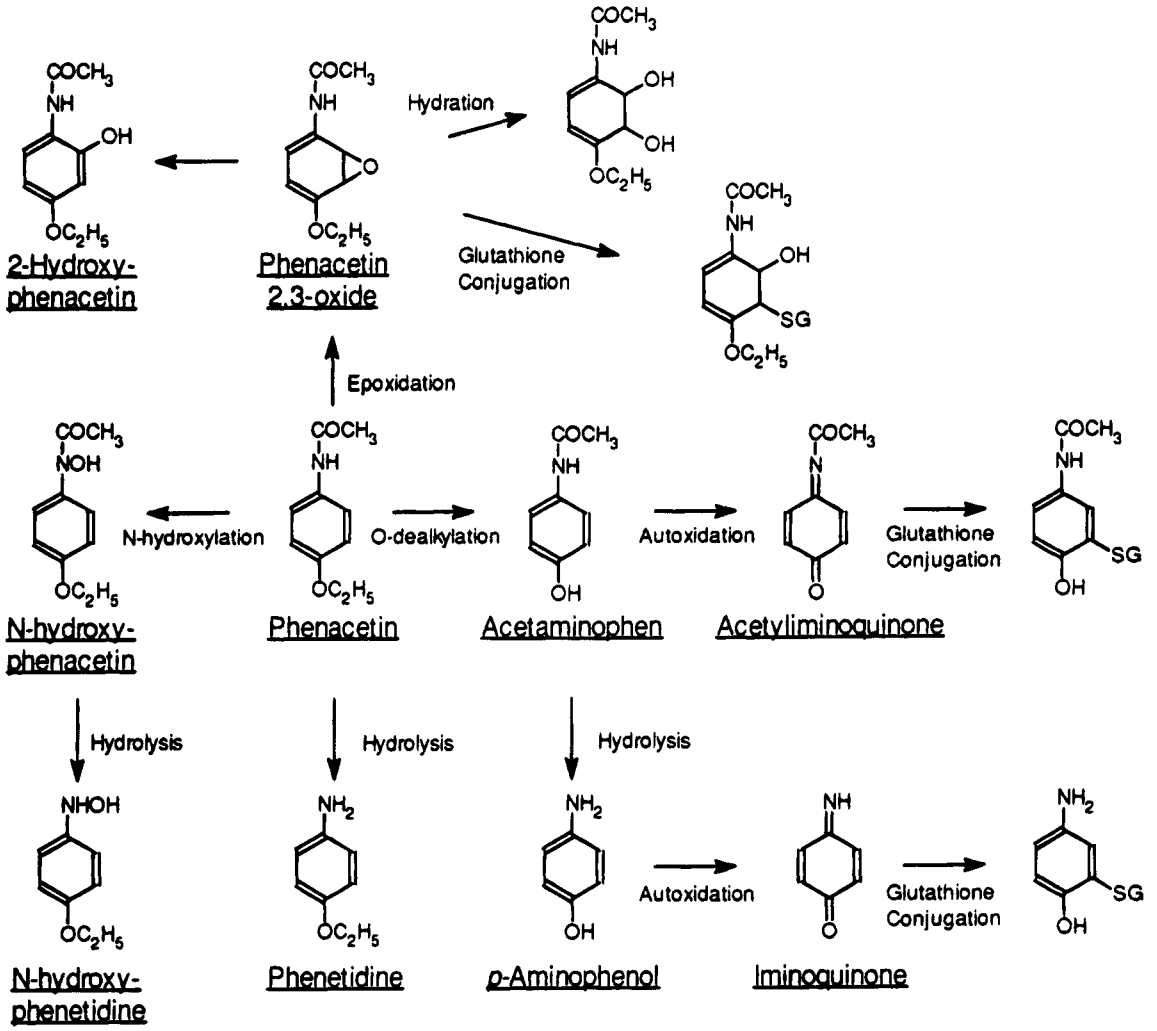


Figure 2. Primary and secondary metabolites of phenacetin that have been successfully predicted by META consulting the dictionary described in the text.

Table 3. Major Biotransformations of Test Compounds That Were Correctly Predicted by META

compound	biotransformation	compound	biotransformation
acetaldehyde	oxidation	mephobarbital	demethylation
acetoxime	hydrolysis and reduction	meprobamate	aliphatic hydroxylation
aniline	N-acetylation and aromatic hydroxylation	naphthalene	epoxidation
anisole	O-demethylation	nicotinic acid	glycination
aspirin	hydrolysis	nitrobenzene	reduction and aromatic hydroxylation
benzamide	hydrolysis	oxalic acid	excreted unchanged
benzanthracene	aromatic epoxidation	phenacetin	O-dealkylation
benzylamine	oxidative deamination	phenobarbital	aromatic hydroxylation
bromobenzene	epoxidation	phenol	sulfation
cadaverine	oxidative deamination	phenylacetamide	hydrolysis
camphor	aliphatic hydroxylation	phenylbutazone	aromatic and aliphatic hydroxylation
chloral hydrate	reduction	phenylhydroxylamine	reduction
cyclohexane	aliphatic hydroxylation	procaine	hydrolysis
erythritol	excreted unchanged	prontosil	reduction
ethanol	oxidation	quinone	glutathione conjugation
ethyl ether	excreted unchanged	salicylamide	glucuronidation
ethylene glycol	oxidation	sulfanilamide	N-acetylation
isoniazid	N-acetylation	thiopental	aliphatic hydroxylation
meperidine	hydrolysis		

due to its volatility, an attribute which is not handled by META methodology (for now).

In the third pharmacological exercise, we checked whether the major metabolic route to elimination could be accurately predicted. Levine[49] characterizes the major multistep pathways given for the four common functionalities listed in Table 4. META followed the observed pathways for test structures containing each functionality.

Prodrugs[50] can be used to increase duration of action and/or reduce side effects of particular pharmacological agents. Esterification is often used in the design of prodrugs. The prodrug enalapril is hydrolyzed by esterases to the active parent enalaprilate, which is an inhibitor of the angiotensin converting enzyme.[10] META correctly predicted this de-esterification with no previous specific knowledge of enalapril.

Table 4. Major Routes of Biotransformation Predicted Correctly by META

functionality	major route to elimination	predicted?
hydroxyl		
aliphatic	oxidation, glucuronide conjugation	yes
aromatic	glucuronide, sulfate conjugation, or methylation	yes
carboxyl		
aliphatic	oxidation, glucuronide conjugation	yes
aromatic	glycine or glucuronide conjugation	yes
amino		
aliphatic	deamination, glucuronide conjugation	yes
aromatic	acetylation, glucuronide conjugation, methylation	yes
aromatic rings	hydroxylation, epoxidation	yes

CONCLUSIONS

We have successfully modeled the majority of mammalian xenobiotic biotransformations with a dictionary of META transforms and demonstrated that the dictionary is useful and predictive. The next level of sophistication requires three aspects of metabolism to be addressed: improvement of metabolite prioritization, species differences, and organ distribution.

Our dictionary has proven to rarely miss the prediction of possible metabolites. It is important to realize that the accuracy of some predictions is indeterminate because the pharmacological data needed to verify the prediction is not available. Pharmacological studies seldom determine that a possible metabolite is *not* produced in any quantity, making rigorous evaluation of predictive performance[5] problematic. Indeed, the prediction of these minor, possibly unrecognized metabolites is one reason why META was developed. In contrast, most experimental studies report the *extent* to which detected metabolites occur. Consequently, the metabolites can be ordered in terms of their pharmacological significance. Examples are plentiful. Cytochrome P-450 oxidizes methylene groups preferentially over methyl groups if both are present in the substrate; alkyl side chains are usually hydroxylated on the penultimate carbon atom. Lipophilicity affects whether an epoxide substrate is conjugated with glutathione or hydrolyzed by epoxide hydratase. Esters that are not sterically hindered are more likely to be hydrolyzed.[9] Carboxylic acids are subjected to competing reactions of glycine and glutamine conjugation. The META methodology contains a parameterization which permits relative adjustment of transform priorities, such as to reflect a substrate's biotransformation preference. Experimental observations such as the above will continue to be incorporated into the dictionary by refining our initial transform prioritization.

Our dictionary is a starting point for prospective species-specific or individual-specific dictionaries. Biological factors (e.g., species,[51] sex, genetics, xenobiotic exposure history, and age) are known to affect metabolism and disposition. Malathion is hydrolyzed in mammals to produce readily excretable compounds, but it is oxidized in insects to produce a cholinesterase inhibitor that confers its insecticidal effects. The degree of N-deacetylation is species-dependent; whereas humans experience the side effect of methemoglobinemia associated with the deacetylated products of phenacetin, rabbits do not show this side effect because their amide hydrolysis is relatively low. Furthermore, the ratio of N-hydroxylation to O-dealkylation in hamsters is much larger than the corresponding ratio in mice, whereas guinea pigs have a very low *N*-hydroxylase activity.[52] Regarding individual differences, the distribution and induction of P-450 isozymes depends on the xenobiotic exposure and genetic background of the individual, which accounts for interindividual differences in response to some therapeutic drugs.[17] There are differences in carcinogenicity and pharmacokinetics of aromatic amines and hydrazine drugs because *N*-acetyltransferases are products of multiple genes.[9]

A primary value of META application will be the prediction of metabolite toxicity, and there are many important cases of organ-specific toxicity. This results from the selective accumulation of a particular metabolite, an unusually high activity of a certain enzyme, and/or an unusual sensitivity of the tissue to the toxic effect of the metabolite. For example, the hepatic toxicity of trichloroethylene is probably due to its increased metabolism to trichloroacetic acid, dichloroacetic acid, and trichloroethylene epoxide.[53] Also, 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) is metabolized by monoamine oxidase B in the brain to the pyridinium derivative MPP^+ which accumulates to toxic levels in dopaminergic neurons.[54] Prioritization of transforms could be used to develop tissue-specific META dictionaries. Such separate dictionaries are a prerequisite for modeling the distribution and transport of xenobiotics and their metabolites among various compartments.

Further refinement of the dictionary will proceed through increasingly challenging pharmacological exercises. Development of the interface of META with the Computer Automated Structure Evaluation program[55] is currently underway which will allow quantitative structure−activity relationships to extend across biotransformations. The dictionary is available from Case Western Reserve University c/o Professor Gilles Klopman.

ACKNOWLEDGMENT

This project was supported by a grant from the Edison Biotechnology Center of Ohio and National Institutes of Health Grant NS22688 (L.M.S.).

REFERENCES AND NOTES

(1) Klopman, G.; Dimayuga, M.; Talafous, J. *META*. 1. A Program for the Evaluation of Metabolic Transformation of Chemicals, preceding paper in this issue.
(2) Atkinson, D. E. *Cellular Energy Metabolism and Its Regulation*; Academic Press: New York, 1977.
(3) Tinker, J. F.; Gelernter, H. Computer Simulation of Metabolic Transformation. *J. Comput. Chem.* **1986**, *7*, 657−665.
(4) Darvas, F. Predicting metabolic pathways by logic programming. *J. Mol. Graphics* **1988**, *6*, 80−86.
(5) Gifford, E.; Johnson, M.; Tsai, C.-c. A graph-theoretic approach to modeling metabolic pathways. *J. Comput. Aided Mol. Des.* **1991**, *5*, 303−322.
(6) Hathaway, D. E. *Foreign Compound Metabolism in Mammals*; The Chemical Society: London, 1975; Vol. 3.
(7) Testa, B.; Jenner, P. *Drug Metabolism: Chemical and Biochemical Aspects*; Marcel Dekker: New York, 1976.
(8) *Xenobiotic Conjugation Chemistry*; Paulson, G. D., Caldwell, J., Hutson, D., Menn, J. J., Eds.; American Chemical Society: Washington, DC, 1986; Vol. 299.
(9) Williams, D. A. In *Principles of Medicinal Chemistry*, 3rd ed.; Foye, W. O., Lea & Febiger: Philadelphia, 1989; pp 79−117.
(10) Gilman, A. G.; Rall, T. W.; Nies, A. S.; Taylor, P. *The Pharmacological Basis of Therapeutics*; 8th ed.; MacMillian Publishing Co.: New York, 1990.
(11) Alvares, A. P.; Pratt, W. B. In *Principles of Drug Action*, 3rd ed.; Pratt, W. B., Taylor, P.; Eds.; Churchill Livingstone: New York, 1990.
(12) Timbrell, J. A. *Principles of Biochemical Toxicology*; 2nd ed.; Taylor & Francis: London, 1991.

(13) Jakoby, W. B. *Enzymatic Basis of Detoxification*; Academic Press: New York, 1980; Vol. 1.
(14) Aizawa, H. *Metabolic Maps of Pesticides*; Academic Press: New York, 1982.
(15) *Bioactivation of Foreign Compounds*; Anders, M. W., Ed.; Academic Press: New York, 1985.
(16) Porter, T. D.; Coon, M. J. Cytochrome P-450. Multiplicity of Isoforms, Substrates, and Catalytic and Regulatory Mechanisms. *J. Biol. Chem.* **1991**, *266*, 13469−13472.
(17) Guengerich In *Cytochrome P450*; Schenkman, J. B., Greim, H. Eds.; Springer-Verlag: Berlin, 1993; Vol. 105, pp 89−103.
(18) Koymans, L.; Donne-op, G. M.; Kelder, D.; Te, J. M. K.; Vermeulen, N. P. E. Cytochromes P450: Their Active-Site Structure and Mechanism of Oxidation. *Drug Metab. Rev.* **1993**, *25*, 325−387.
(19) Rietjens, I. M. C.; Soffers, A. E. M. F.; Veejer, C.; Vervoort, J. Regioselectivity of Cytochrome P-450 Catalyzed Hydroxylation of Fluorobenzenes Predicted by Calculated Frontier Orbital Substrate Characteristics. *Biochemistry* **1993**, *32*, 4801−4812.
(20) Miller, E. C.; Miller, J. A. Searches for Ultimate Chemical Carcinogens and Their Reactions with Cellular Macromolecules. *Cancer* **1981**, *47*, 2327−2345.
(21) MacDonald, T. L. Chemical mechanism of halocarbon metabolism. *CRC Crit. Rev. Toxicol.* **1983**, *11*, 85−120.
(22) *The Enzymes*; Boyer, P. D., Ed.; Academic Press: New York, 1975; Vol. XIIb.
(23) Zeller, E. A. Diamine Oxidase: Specificity and Mechanism. *Adv. Biochem. Psychopharmacol.* **1972**, *5*, 167−171.
(24) *Biological Oxidation of Nitrogen in Organic Molecules*; Bridges, J. W., Gorrod, J. W., Parke, D. V., Ed.; Taylor & Francis: London, 1972.
(25) Ziegler, D. M. Flavin-containing Monooxygenases: Catalytic Mechanism and Substrate Specificities. *Drug Metab. Rev.* **1988**, *19*, 1−32.
(26) Frederick, C. B.; Mays, J. B.; Ziegler, D. M.; Guengerich, F. P.; Kadlubar, F. F. Cytochrome P-450- and Flavin-Containing Monooxygenase-catalyzed Formation of the Carcinogen *N*-hydroxy-2-aminofluorene and Its Covalent Binding to Nuclear DNA. *Cancer Res.* **1982**, *42*, 2671−2676.
(27) Mitchell, S. C.; Waring, R. H. The Early History of Xenobiotic Sulfoxidation. *Drug Metab. Rev.* **1986**, *16*, 255−284.
(28) Sitar, D. S. *Metabolism* of Thioamide Antithyroid Drugs. *Drug Metab. Rev.* **1990**, *22*, 477−502.
(29) *Biochemistry and Pharmacology of Ethanol*; Majchrowicz, E. Noble, E. I., Ed.; Plenum Press: New York, 1979.
(30) Sladek, N. E.; Manthey, C. L.; Maki, P. A.; Zhang, Z.; Landkamer, G. J. *Xenobiotic* Oxidation Catalyzed by Aldehyde Dehydrogenases. *Drug. Metab. Rev.* **1989**, *20*, 697−720.
(31) Kyerematen, G. A.; Vessel, E. S. *Metabolism* of Nicotine. *Drug Metab. Rev.* **1991**, *23*, 3−41.
(32) Junge, W.; Krisch, K. The Carboxylesterases/Amidases of Mammalian Liver and Their Possible Significance. *CRC Crit. Rev. Toxicol.* **1974**, *3*, 371−434.
(33) Estus, G. S.; Mieyal, J. J. Structure-activity Relationships for the Deacetylation of a Homologous Series of Phenacetin Analogs and Their *N*-hydroxy Derivatives. *Drug. Metab. Dispos.* **1983**, *11*, 471−476.
(34) *Molecular Aspects of Monooxygenases and Bioactivation of Toxic Compounds*; Arinc, E., Schenkman, J. B., Hodgon, E., Ed.; Plenum Press: New York, 1989.
(35) Conney, A. H. Induction of Microsomal Enzymes by Foreign Chemicals and Carcinogenesis by Polycyclic Aromatic Hydrocarbons: G. H. A. Clowes Memorial Lecture. *Cancer Res.* **1982**, *42*, 4875−4917.
(36) *Enzymology of Carbonyl Metabolism 2: Aldehyde Dehydrogenase, Aldo-Ketoreductase, and Alcohol Dehydrogenase*; Flynn, T. G., Weiner, H., Ed.; Alan, R. Liss, Inc.: New York, 1985.
(37) Levine, W. G. Metabolism of Azo Dyes: Implication for Detoxification and Activation. *Drug Metab. Rev.* **1991**, *23*, 253−309.
(38) McLane, K. E.; Fisher, J.; Ramakrishnan, K. Reductive Drug Metabolism. *Drug. Metab. Rev.* **1983**, *14*, 741−799.
(39) Caldwell, J. Conjugation Reactions in Foreign-Compound Metabolism: Definition, Consequences, and Species Variations. *Drug Metab. Rev.* **1982**, *13*, 745−777.
(40) Dutton, G. J. *Glucuronidation of Drugs and Other Compounds*; CRC Press, Inc.: Boca Raton, FL, 1980.
(41) Egan, D.; O'Kennedy, R.; Moran, E.; Cox, D.; Prosser, E.; Thornes, R. D. The Pharmacology, Metabolism, Analysis, and Applications of Coumarin and Coumarin-related compounds. *Drugs Metab. Rev.* **1990**, *22*, 503−529.
(42) Price-Evans, D. A. *N*-Acetyltransferase. *Pharmacol. Ther.* **1989**, *42*, 157−234.
(43) Webster, L. T., Jr.; Sibbiqui, U. A.; Lucas, S. V.; Strong, J. M.; Mieyal, J. J. Identification of Separate Acyl-Co-A: Glycine and Acyl-CoA: L-Acyltransferase Activities in Mitochondrial Fractions from Liver of Rhesus Monkey and Man. *J. Biol. Chem.* **1976**, *251*, 3352−3358.
(44) Anders, M. W.; Dekant, W.; Vamvakas, S. Glutathione-dependent toxicity. *Xenobiotica* **1992**, *22*, 1135−1145.
(45) Usdin, E.; Borchardt, R. T.; Creveling, C. R. *Transmethylations*; Elsevier: New York, 1979.
(46) Creveling, C. R.; Morris, N.; Shimizu, H.; Ong, H. H.; Daly, J. Catechol O-methyltransferase IV. Factors Affecting *m*- and *p*-Methylation of Substituted Catechols. *Mol. Pharmacol.* **1972**, 8, 398−409.
(47) Williams, R. T. *Detoxication Mechanisms*, 2nd ed.; Wiley: New York, 1959.
(48) Williams, R .T. Detoxication mechanisms in man. *Clin. Pharmacol. Ther.* **1962**, *4*, 234−254.
(49) Levine, R. R. *Pharmacology*: *Drug Actions and Reactions*; Little, Brown and Company: Boston, 1983.
(50) Stella, V. J.; Himmelstein, K. J. Prodrugs and Site-Specific Drug Delivery. *J. Med. Chem.* **1980**, *23*, 1275−1282.
(51) Caldwell, J. The Current Status of Attempts to Predict Species Differences in Drug Metabolism. *Drug Metab. Rev.* **1981**, *12*, 221−237.
(52) Miller, E. C.; Miller, J. A.; Hartman, H. A. The Comparative Carcinogenicities of 2-Acetylaminofluorene and Its *N*-Hydroxy Metabolite in Mice, Hamsters, and Guinea Pigs. *Cancer Res.* **1963**, *24*, 2018−2032.
(53) Davidson, I. W. F.; Beliles, R. P. Consideration of the Target Organ Toxicity in Terms of Metabolite Toxicity and Pharmacokinetics. *Drugs Metab. Rev.* **1991**, *23*, 493−599.
(54) Sayre, L. M. Biochemical mechanism of action of the dopaminergic neurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). *Toxicol. Lett.* **1989**, *48*, 121−149.
(55) Klopman, G. Artificial Intelligence Approach to Structure−Activity Studies. Computer Automated Structure Evaluation of Biological Activity of Organic Molecules. *J. Am. Chem. Soc.* **1984**, *106*, 7315−7320.