Gauging Reactive Metabolites in Drug-Induced Toxicity

Marsha R. Eno and Michael D. Cameron*

Scripps Research Institute – Scripps Florida, Department of Molecular Therapeutics, 130 Scripps Way, Jupiter, FL 33458, USA

Abstract: Over the past decades, it has become abundantly clear that enzymes evolved to detoxify and eliminate foreign chemicals from the body, occasionally generate highly reactive metabolites which have toxicological implications. To decrease the probability of late clinical failure or market withdrawal, there has been an increased prioritization on understanding key metabolic processes that might cause drug interactions or toxicities. Significant advances have been made in the detection of reactive metabolites and in understanding the structure activity relationship. It is now widely accepted that compounds with certain functional groups such as anilines, quinones, hydrazines, thiophenes, furans, acylpropionic acids, and alkynes have a much greater associated risk towards formation of reactive metabolites than compounds that do not contain such “structural alerts”. Detection of reactive metabolites is usually done with in vitro assays, which have become more sensitive with advances in mass spectrometry. As an increasingly large number of compounds that form reactive metabolites have been identified, much of the focus has shifted from detection to evaluation of toxicological implication. While there is a disproportionate number of compounds metabolized to reactive metabolites that are associated with drug-induced hepatotoxicity and serious skin toxicities such as toxic endothelial necrolysis and Steven’s Johnson syndrome, attempts to predict toxicity based on in vitro testing have been discouraging. In this review we attempt to summarize the experimental options available to evaluate reactive metabolites.

Keywords: Bioactivation, cytochrome P450, hepatotoxicity, reactive metabolites.

INTRODUCTION

Experiments designed to detect the formation of reactive metabolites have historically been done retrospectively to understand the toxicological mechanisms of compounds shown to have associated human toxicities. More recently, reactive metabolite screening has worked its way into the metabolic and safety profiling of compounds in the drug discovery process [1-5]. This brings us to the central focus of this review. Upon being informed that a compound or a series of compounds is metabolized to a reactive metabolite, what additional in vitro or in vivo experiments might be performed to predict the toxicological risk while making decisions to advance or deprioritize compounds?

Hepatic toxicity is the leading cause for market withdrawal of approved drugs [6]. Additionally, the typical rate of acute hepatic failure (hepatic toxicity leading to death or liver transplant) for withdrawn drugs was one case per ten-thousand patients [7]. Most compounds that fail in clinical development, obtain black box warnings, or are withdrawn due to hepatic toxicity may have had no cases of hepatic failure during the clinical development and the majority of clinical trial patients would have shown no markers indicating even mild hepatic toxicity. While the decision making trees for individual companies are living documents with continual updates, several companies have published manuscripts that discuss relevant factors or retrospectively evaluate data for use in making risk-based assessments [1, 8-10].

What are Reactive Metabolites?

The term reactive metabolite is not universally defined, but generally indicates a short lived metabolite capable of chemically modifying biomolecules formed through the metabolism of a “non-reactive” parent compound. Cases where the parent compound is inherently reactive or is converted to a reactive form by facile chemical means such as encountering the low pH environment of the stomach are not usually termed reactive metabolites even though they might be associated with similar toxicities.

Why are Reactive Metabolites Generated and from an Evolutionary Standpoint Why are they Tolerated?

Because of the excessive number of exogenous chemicals that we come into contact with, it is not possible to have specialized enzymes to safely eliminate each foreign compound. One option to deal with the array of threats would be with an arrangement similar to the immune system where multiple layers of defense combine to make the innate and adaptive immune system. This provides protection against common pathogens and permits tailored responses to new threats through the incorporation of a hyper variable domain that allows the binding epitope of antibodies to rapidly mutate in response to newly encountered pathogens.

Whereas biological pathogens are built upon twenty standard amino acids with similar stereochemistry, only
slightly expanded by non-standard and modified amino acids, the plethora of xenobiotics encountered are structurally diverse and make an adaptive response implausible. A fraction of compounds can be excreted unchanged via filtration in the kidney or export into bile whereas others, particularly lipophilic compounds would otherwise accumulate in biological membranes and must be modified prior to elimination. While there are several families of enzymes responsible for “chemical decontamination”, we will focus on the cytochrome P450 enzymes as they are the most prominent of these systems.

The human genome encodes for fifty-seven P450 proteins. P450s are most prevalent in the liver where their heme group is responsible for the distinctive color of liver. A majority of the P450 enzymes are minimally expressed or catalyze reactions involved in such cellular processes as steroid and prostaglandin biosynthesis; however, approximately a dozen P450s are exceptionally non-specific and are capable of oxidizing thousands of substrates [11-14]. This high degree of non-specificity is absolutely essential to deal with the wide range of xenobiotics encountered in the natural environment. The substrate promiscuity of P450s is achieved through harnessing the oxidative potential of molecular oxygen to generate a powerful oxidant, the porphyrin iron (IV) oxo species, which has sufficient potential to oxidize most organic molecules [15-17]. These same organic molecules could have been degraded directly through the three electron reduction of molecular oxygen to form the hydroxyl radical, an exceptionally strong oxidant; however, this mechanism cannot be controlled and would lead to excessive cellular toxicity. P450s “sequester” and control access to the oxidative species, thus decreasing the oxidation of unintended biomolecules. Because the P450 oxidative species is so strong, substrates can often be oxidized at more than one position and the metabolite profile has as much to do with the presentation of the molecule to the activated oxygen as it does the “easiest” site of oxidation based on redox potential [18-21]. This often leads to the catalysis of multiple metabolites from a single P450 enzyme or sequential metabolism giving rise to secondary and tertiary metabolites.

Depending on the nature of the substrate, oxidations catalyzed by P450 can generate electron deficient metabolites which in turn oxidize cellular components. There are certain chemical moieties that are more prone to the generation of reactive metabolites and there have been several excellent reviews that have discussed “structural alerts” and correlated these to the formation of reactive metabolites and/or hepatotoxicity [22-25].

**Are Reactive Metabolites Always Toxic?**

Clinically severe toxicity is observed in only a portion of compounds that have been shown to generate reactive metabolites *in vitro.* This is one of the key shortcomings in the ability to understand the importance of reactive metabolites. It should be pointed out that these studies are built upon biased data sets as they do not comprise a random sampling of molecules positive or negative for reactive metabolite formation. Literature references where the extent of reactive metabolite formation was indexed to toxicity were done primarily with drugs or drug candidates. This biases the analysis to a test set of compounds comprised exclusively of molecules that were deemed “drug-like” and had sufficiently clean preclinical toxicity profiles to progress into human trials. Compounds that can be clearly identified as toxins from *in vitro* or animal studies, e.g. aflatoxin b1 which is bioactivated by P450 to form a reactive epoxide leading to necrosis, cirrhosis, and carcinoma of the liver with lethal doses in most species below 1 mg/kg would not be included in the analysis [26, 27]. Compounds identified as toxic in preclinical species would either be excluded from progressing to subsequent human clinical trials, or the dose would be reduced to levels where significant toxicity was not expected. This leads to the possibility that drugs known to form reactive metabolites but not showing hepatic toxicity may comprise a sub-group of compounds with reduced toxicity profiles.

**What Protection Does the Liver have from Reactive Metabolites?**

The liver resides between the gut and general circulation. Foreign substances, including toxins and bacteria, that are ingested are typically absorbed in the small intestine where they enter the portal vein and are directed to the liver before proceeding to the rest of the body. Because of its location and function, the liver is exposed to higher levels of toxins than most other organs. One of the key reasons that serious hepatic events are rare is that the liver has well-orchestrated safeguards.

One of the key regulators in response to electrophiles is Nrf2. In its “resting” state, Nrf2 is retained in the cytoplasm by binding to Keap1. When bound to Keap1, Nrf2 is ubiquininated and rapidly turned over with a cellular half-life of approximately 15 minutes [28]. In the presence of elevated levels of electrophiles, the sulfhydryl groups of the exposed cysteines on Keap1 are covalently modified resulting in a decrease in the binding affinity for Nrf2 [29, 30]. Nrf2 translocates to the nucleus and binds to the antioxidant response element and promotes the transcription of numerous cytoprotective genes [31, 32]. These include but are not limited to: NAD(P)H dehydrogenase which can reduce quinones, quinone-imines, and semi-quinones; aldo-keto reductase, which reduces reactive ketones and aldehydes; sulfortransferases, glutathione S-transferases and glutathione reductase, proteins to increase glutathione synthesis, heme oxygenase, ferritin and metallothioneins to reduce heavy metal toxicity and redox cycling, and efflux transporters to increase the flow of xenobiotics into the bile [33-35].

Endogenous levels of proteins that function to protect against electrophiles are elevated in the liver over their concentration in other tissues. The most abundant proteins in the liver, the organ with the highest concentration of P450s, are glutathione-S-transferase enzymes, which have a detoxification role by “quenching” electrophiles generated by P450s through coupling them with glutathione [36-39].

The liver has the largest population of tissue resident immune cells with approximately 25% of the non-hepatocytes in the liver being macrophages or lymphocytes [40]. In addition to its resident macrophages, liver has the ability to rapidly recruit additional neutrophils, monocytes, and lymphocytes in response to tissue injury. There is significant literature dealing with the role of immune cells in
drug-induced liver injury. Earlier literature reported contradictory conclusions, but more recent reports have elucidated a dual role of many immune cells. While they can release cytokines and pro-inflammatory mediators which can lead to increased tissue damage, they also release factors involved in cell survival and tissue repair [41-43]. For additional reading on the function of immune cells in drug-induced hepatotoxicity see the indicated references [44-48]. The high hepatic macrophage population allows damaged cells to be safely cleared, reducing necrotic cell death and the accumulation of scar tissue/cirrhosis.

The liver is an immunologically privileged organ and the default response to antigens is immune tolerance [49-52]. Despite the propensity for hepatic immune tolerance, there are several compounds where toxicity has clear genetic links implicating an immune reaction. Immune-mediated hepatotoxicity will be briefly discussed later in this review.

Are Reactive Metabolites Associated with Dose-Dependent or Idiosyncratic Toxicity?

One of the challenges of evaluating metabolism-induced toxicity is that it may exhibit features of acute, chronic or mixed toxicities. Observed toxicity is a function of the properties of the compound combined with the susceptibility of the individual [53, 54]. Compounds that are acutely toxic typically exhibit a dose-dependent increase in toxicity. Reactive metabolites have been implicated in the toxicity of numerous drugs or environmental pollutants and may manifest through multiple mechanisms including endoplasmic reticulum [55] or mitochondrial stress [56, 57], chromosomal damage [58-62] or change in cellular redox potential [63-65].

Reactive metabolite formation has also been implicated in idiosyncratic hepatic or dermal drug-induced toxicity. By definition, an idiosyncratic toxin demonstrates toxicity that is peculiar to the individual. Idiosyncratic toxins display rare and unpredictable toxicity which is not strictly dose-dependent. Information explaining why the majority of patients taking these drugs are subjected to similar levels of reactive metabolite, yet only a small percentage of patients develop significant toxicity is a source of active investigation. The best supported hypothesis involves an immune response and this is discussed in greater detail later in this review.

The Acetaminophen Paradox – Safe Drug or Bad Actor

The most studied compound to form a reactive metabolite is acetaminophen (paracetamol). Acetaminophen generally exhibits dose-dependent toxicity, it is toxic across animal species, rapidly causes high levels of hepatic damage, and the toxic metabolite is known [66-70]. Acetaminophen is oxidized to an electrophilic metabolite N-acetyl-p-benzoquinone imine (NAPQI) by multiple hepatic P450 [67, 71-77]. Acetaminophen toxicity is the leading cause of acute liver failure in the Western World and factors associated with acute acetaminophen-induced toxicity are well known and universal. While acetaminophen is widely regarded as being safe, a 2006 study by the FDA – Office of Drug Safety estimated that in the United States alone there were 56,000 emergency room visits and 458 deaths from acute liver failure caused by acetaminophen [78]. Without N-acetylcysteine treatment to minimize cellular damage caused by NAPQI [79, 80], acetaminophen-induced deaths would be even higher. Many of the overdoses are associated with coformulations, particularly with prescription pain killers where acetaminophen is mixed with oxycodone (Percocet) or hydrocodone (Vicodin) [78, 81, 82]. The FDA has recently restricted the amount of acetaminophen that can be used in these pain management co-formulations. Vicodin, the most prescribed medication in America with over 130 million prescriptions in 2010, has recently been reformulated to decrease the acetaminophen dose from 750 mg to 300 mg while retaining hydrocodone dosage.

While select individuals may have greater or lower sensitivity to acetaminophen associated toxicities, if dosed high enough, 100% of individuals would be expected to experience hepatic damage leading to liver failure. NAPQI is electron deficient and capable of reacting with multiple biomolecules which gives rise to an array of cellular damage. Papers have been published suggesting that the reactive metabolite formed by acetaminophen overdose causes toxicity via changing the cellular concentration of glutathione [83, 84], changing the redox potential of the cell, covalently modifying key cellular proteins [69, 85-87], causing mitochondrial damage [88-91] and endoplasmic reticulum stress [92]. All of these are likely correct and combine to amplify the toxicity that would be observed from any one pathway. While acetaminophen is certainly acutely toxic, it also shares some similarities to an idiosyncratic toxin. While observed in very few patients, a subset of the population is hypersensitive to acetaminophen and exhibit significant toxicity even when taking the drug at recommended doses. This is likely due to an immunological response to NAPQI-conjugated proteins. Acetaminophen specific IgE antibodies and acetaminophen positive skin tests have been detected in hypersensitive patients [93, 94].

Because of the large number of people that safely take acetaminophen every year, it is frequently brought up in discussions about reactive metabolites as the reason why compounds that covalently modify proteins can be safe drugs. This argument is tenuous. The FDA has recently expanded the warning labels and restricted acetaminophen doses due to the danger of liver toxicity [95-104]. It is highly likely that acetaminophen would not gain approval if it was submitted for regulatory approval today. Benchmarking against such a compound is exceedingly risky.

What Assays are Available to Evaluate Reactive Metabolite Formation?

A range of methods to identify compounds that generate reactive metabolites and try to predict future risk of serious drug-induced toxicity are discussed in the remainder of this review. They are listed in Table 1 and where they might be incorporated in the drug discovery and development timeline is depicted in Fig. (2).

IN VITRO METHODS

Radiolabel Incorporation/Covalent Adduction

The easiest assay to understand, and paradoxically the most controversial method for detection of reactive
Table 1. Methods to evaluate reactive metabolites.

<table>
<thead>
<tr>
<th>Method</th>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong> In vitro – microsomal radiolabel incorporation/covalent adduction</td>
<td>• simple • potential to identify reactive secondary or tertiary metabolites • quantitative</td>
<td>• requires radiolabeled compound • P450-centric • lack of detoxification pathways • many compounds that bind extensively are safe drugs</td>
</tr>
<tr>
<td><strong>2</strong> In vitro – reactive metabolite trapping</td>
<td>• established methodology • amenable to moderate throughput</td>
<td>• unknown significance • usually not quantitative</td>
</tr>
<tr>
<td><strong>3</strong> In vitro – P450 time-dependent inhibition</td>
<td>• simple • data is already collected to predict pharmacokinetic drug-drug interactions</td>
<td>• correlation with toxicity is modest • TDI mechanism may be independent of reactive metabolite formation</td>
</tr>
<tr>
<td><strong>4</strong> In vitro – genotoxicity (Ames test)</td>
<td>• data is already required for IND package</td>
<td>• mechanism of genotoxicity may not involve a reactive metabolite</td>
</tr>
<tr>
<td><strong>5</strong> In vitro/in vivo – proteomics</td>
<td>• data may eventually help identify a fingerprint or profile for compounds that have higher probabilities of clinical toxicity</td>
<td>• methodology needs improved • generally not quantitative • minimal historical data</td>
</tr>
<tr>
<td><strong>6</strong> In vivo – micronucleous assay</td>
<td>• part of the preclinical safety package</td>
<td>• mechanism of genotoxicity may not involve a reactive metabolite</td>
</tr>
<tr>
<td><strong>7</strong> In vivo – whole body radiolabel incorporation or PET</td>
<td>• identifies accumulation in tissues • pharmacokinetics, transporters, and cellular detoxification pathways are accounted for • quantitative</td>
<td>• requires radiolabeled (or positron-emitting) compound • does not distinguish between the parent drug and metabolites • requires knowledge of the metabolic pathway to assure that the label is not lost</td>
</tr>
<tr>
<td><strong>8</strong> In vivo – glutathione and mercapturate adducts</td>
<td>• possible with urine, feces, and/or bile • structural information can be obtained via mass spectrometry experiments to identify the likely reactive metabolite</td>
<td>• uncertainty if the detected adduct is from reaction with cellular glutathione pools, protein cysteines, or glutathione S-transferase</td>
</tr>
<tr>
<td><strong>9</strong> In vivo – markers of hepatocellular injury (ALT, ALP, bilirubin)</td>
<td>• identifies hepatic injury • biliary vs hepatocellular injury can be identified • plasma samples are sufficient</td>
<td>• elevation of transaminases have been observed without progression to significant hepatic injury</td>
</tr>
<tr>
<td><strong>10</strong> In vivo – HLA association</td>
<td>• may identify at-risk populations</td>
<td>• usually performed retrospectively • association may be complicated if multiple HLA are involved • unclear if HLA association would be useful for mild hepatic injury</td>
</tr>
<tr>
<td><strong>11</strong> In vivo – lymphocyte proliferation assay</td>
<td>• only a blood draw is required</td>
<td>• labor intensive • only approximately half of patients with serious immune-mediated hepatic or cutaneous injury are positive in LTT test • uncertain if evaluation of patients with mild ALT elevation vs control patients would predict increased risk of significant toxicity</td>
</tr>
<tr>
<td><strong>12</strong> In vivo – patch/skin hypersensitivity test</td>
<td>• simple to administer</td>
<td>• may harm patient • compounds that are irritants may be difficult to test</td>
</tr>
<tr>
<td><strong>13</strong> In vivo – eosinophilia/DRESS</td>
<td>• demonstrates immune-mediated toxicity</td>
<td>• observed in a small percentage of patients • eosonophilis may be elevated by non-related allergens</td>
</tr>
<tr>
<td><strong>14</strong> In vivo – re-activation of virus</td>
<td>• non-invasive patient sampling • high percentage of positive samples among patients that experience severe immune mediated toxicity</td>
<td>• mechanism is not certain • unclear if the method would have predictive value for later toxicity if used with asymptomatic patients during clinical trials</td>
</tr>
<tr>
<td><strong>15</strong> In vivo – autoantibodies or drug-protein antibodies</td>
<td>• high throughput ELISA based assays for autoantibody detection could be developed</td>
<td>• utility in prediction of severe toxicity is unknown</td>
</tr>
</tbody>
</table>
metabolite formation, is the radiolabel incorporation assay. This was extensively published by researchers from Merck and they used the methodology for many years [9, 105-108]. Radiolabeled drug is incubated with a high concentration of liver microsomes and NADPH for an extended time period. The protein is precipitated, extensively washed, and the amount of radioactivity covalently incorporated into the microsomal protein fraction is determined [109, 110]. The assay casts a wide net and is typically able to identify compounds oxidized to reactive metabolites by microsomal enzymes even if the ultimate reactive metabolite is a tertiary or quaternary metabolite due to the long incubation time.

Modifications to this general study design have been used for compounds that are not primarily metabolized in the liver. For instance, the generation of reactive clozapine metabolites by neutrophils has been demonstrated through the isolation of neutrophils and incubation with radiolabeled clozapine [111].

Two major objections are raised against this method. The first centers on experimental set-up due to the lack of detoxification pathways in the assay and the bias towards P450 catalyzed reactions through the use of microsomes. The second frequently raised objection to the radiolabel incorporation/covalent addition assay is that the correlation between the amount of radiolabel incorporated (pMoles/mg protein) and observed toxicity in human patients or rodents is poor [112].

While it is doubtful many people would argue that covalent modification of proteins by reactive metabolites is beneficial, there are people that argue that the lack of correlation between covalent binding and observed toxicity justifies not using reactive metabolite formation within the criteria for candidate selection. An internal review by Merck in 2009 found that a quarter of lead optimization projects had issues with covalent binding of reactive metabolites and that it took on average 6 months to redesign the molecules to dial out such metabolites. The assay casts a wide net and is typically able to identify compounds oxidized to reactive metabolites by microsomal enzymes even if the ultimate reactive metabolite is a tertiary or quaternary metabolite due to the long incubation time.

A perspective piece written by scientists from GlaxoSmithKline examined covalent binding data along with chemical trapping and P450 inactivation data for over 200 marketed drugs [8]. Their data demonstrated trends between observed hepatotoxicity in human patients and in vitro measurements of reactive metabolite formation when the assay results were combined with information about daily dose. Similar results were reported from a team at Bristol-Myers Squibb where 50 approved drugs were evaluated. Observed hepatotoxicity marginally correlated with detection of reactive metabolites, assessed by glutathione trapping studies [1]. This correlation improved when the levels of thiol adduct formation and daily dose were incorporated into the analysis.

It would be of interest if data from individual patients enrolled in extended dosing studies were made available where hepatic function is tracked via weekly or bi-weekly blood chemistry analysis. Correlation of ALT elevation during the first few days of treatment (acute toxicity) and 3-10 weeks after starting the drug (common for immune-mediated toxicity) could then be evaluated in context with binding data and administered dose.

In Vitro Chemical Trapping Experiments

While there are several compounds that have been utilized as trapping agents for detection of reactive metabolites, some of the most common are glutathione [1, 113-116], N-acetylcysteine [117-120], cyanide [121-124], semicarbazide [125], methoxylamine [126, 127], N-acetylysine [128, 129], and N-acetyltyrosine [130-132]. Glutathione and N-acetylcysteine are used to trap soft electrophiles such as quinones. N-acetylysine and cyanide are used to trap harder electrophiles. Semicarbazide and methoxylamine [133] are primarily used to trap reactive aldehydes. N-acetyltyrosine can dimerize with some free radicals. In a typical experiment, human liver microsomes or human hepatocytes are incubated in the presence and absence of trapping agent. While a few detection methods have been utilized to detect trapped metabolites, mass spectrometry is the most common.

Glutathione covalently binds to electrophiles and can be detected using multiple scan types including positive ion mode with a neutral loss scan following loss of 129 Da, corresponding to the collision-induced fragmentation/loss of glutamate. An alternative method using negative ion mode looks for a fragment ion of 272 Da corresponding to breaking the thioether bond [132, 134-136]. The negative ion approach is generally more sensitive and gives fewer false positives, but does not provide information regarding the position of glutathione binding to the reactive metabolite because the negative charge resides on glutathione. Modern mass spectrometers can rapidly switch between negative and positive ion mode and it is now possible to run the experiment in both modes using a single injection [135].

Glutathione and N-acetylcysteine are not suitable for trapping reactive aldehydes. N-acetylysine is sometimes used, but is not ideal because the generated imine is reversible. Semicarbazide and methoxylamine are soluble, minimally inhibit P450 enzymes, and react with aldehydes to form stable products. Reaction with semicarbazide and methoxylamine typically increase the mass by 57 and 29 Da, respectively [125-127]. Other hydrazine containing molecules such as phenylhydrazine or 2,4-dinitrophenylhydrazine can also be used depending on the specialized needs of the researcher.

Cyanide adducts can often be detected as a neutral loss of 27 Da corresponding to the collision-induced loss of the cyanide group. The use of a 1:1 ratio of stable labeled cyanide (13C14N:13C13N) gives a characteristic doublet that can help differentiate true cyanide adducts from other components within the reaction that can generate a 27 Da neutral loss
Enzyme Modification

[123, 132, 137]. Because of the high sensitivity of modern mass spectrometers, care needs to be taken to avoid false positives. In our experience we are generally distrustful of results from cyanide trapping experiments without the inclusion of significant controls. We have seen many compounds where non-enzyme catalyzed cyanide adducts can clearly be identified by LC-MS/MS. Control samples +/- NADPH are not sufficient and parallel samples +/- NADPH, samples with cyanide added after stopping the reaction with organic solvent, heat inactivated enzyme controls, and compound + cyanide in buffer or media should be included and carried through all sample processing steps to assure the detected cyanide adduct is enzymatically generated.

**Inactivation of P450**

Because many reactive metabolites are generated within the active site of a cytochrome P450, it follows that cytochrome P450s are a likely target. Tienilic acid was removed from the market three years after gaining regulatory approval. Tienilic acid is a mechanism-based inhibitor of CYP2C9 and in some patients that experienced severe hepatic toxicity anti-CYP2C9 antibodies were detected [138]. A GlaxoSmithKline perspective piece presented an analysis of P450 time-dependent inhibition for 200 marketed drugs and showed a moderate correlation with hepatic toxicity [8]. Compounds that demonstrate time-dependent P450 inhibition should be considered as potentially generating a reactive metabolite capable of oxidizing or covalently modifying cellular biomolecules until experiments are conducted to determine otherwise.

For most of the known P450 inactivators, there is not a 1:1 ratio of inactivated enzyme per metabolic event [139-146]. Instead, individual P450 proteins may be able to oxidize the inhibitor tens or hundreds of times before becoming inactivated. This phenomena results in the inhibition having a time-dependence where the observed inhibition increases with extended incubation time. While P450s catalyze reactions at varying rates depending on the individual enzyme and substrate, all are relatively slow with typical rates ranging between 0.1 to 50 substrates oxidized per second. In a theoretical reaction where a P450 metabolizes an inhibitory compound once per second with one in every 200 metabolic cycles leading to enzyme inactivation, the enzymatic half-life would be 2.3 minutes. Because of the time-dependence, remaining enzymatic activity can be evaluated after various incubation times, e.g. 0, 1, 3, 5, 10, and 15 minutes, and the rate of enzyme inactivation can be calculated. Compounds where the inactivation efficiency approaches 1:1 are more difficult to detect by measuring time-dependent inhibition because enzyme inactivation is completed so quickly.

Time-dependent P450 inhibition is typically tested using in vitro experiments with liver microsomes or recombinant P450. Compounds that show time-dependent inhibition fall into at least four categories, two of which are related to reactive metabolites, but all have clinical significance in terms of drug-drug interactions. These are discussed below:

**Enzyme Modification**

The generated reactive metabolite can oxidize key amino acids or covalently bind within the P450 active site or substrate access channel and thus restricts access of other substrates. Several compounds have been identified that covalently modify the reactive site of P450s. Some of these are complemented by mutational studies where the added amino acid was mutated and the enzyme was shown to no longer be inactivated [147-152]. Covalent modification is usually associated with, but does not necessitate enzyme inactivation. Work from the labs of Hollenberg and Halpert demonstrated covalent modification of CYP2B4 which resulted in reduction, but not elimination of enzymatic function [153, 154]. The observed time-dependent decrease in enzymatic activity was caused by a metabolite of t-butylphenylacetylene covalently binding to threonine 302 of helix I in the active site. As the percentage of covalently modified threonine 302 increased, the rate of enzymatic activity decreased but did not go to zero. This is an interesting observation which may extend to other P450s and time-dependent inhibitors.

Whereas covalent modification of P450 has a high probability of causing pharmacokinetic drug-drug interactions, it is unclear how significant this is in terms of hepatic toxicity. There appear to be at least a subset of compounds where reactive metabolites can induce an immune response to P450, e.g. tienilic acid where anti-CYP2C9 antibodies were detected in patient serum [138] or with halothane [155] or disulfiram [156] where anti-CYP2E1 and anti-CYP1A2 antibodies have been reported. Reactive metabolites that inactivate P450s can often exit the active site where they can react with non-P450 proteins leading to the potential for toxicity.

**Heme Modification**

Alkylation or cleaving the porphyrin ring of the heme can result in P450 inactivation. Several compounds including mibebradil and methadone have been shown to directly modify the P450 heme [157-159]. The pyrrole nitrogens, vinyl, and propionate groups of the heme have been reported to be oxidized [160-164]. Significant work in this area has been done by the Correia lab where they have shown that P450 heme destruction can sometimes be reversed in cells by reconstitution with fresh heme. In cases where the heme becomes covalently cross linked to the protein or when there is insufficient heme to reconstitute the P450 the protein is rapidly degraded via ubiquitin-dependent proteasomal degradation [165-169]. Using in vitro assays, test compounds can be incubated with recombinantly expressed enzyme and destruction of the heme can be monitored using the heme absorption spectrum in a process sometimes referred to as “heme bleaching”, or by quantitating the heme by LC-UV or LC-MS [152, 157, 162, 170, 171]. It is best to use preparations that have minimal cytochrome b5 to avoid confusion with the b5 heme which will not be lost during P450 inactivation. Hepatic toxicity specifically due to heme destruction is not expected, but molecules that cause heme destruction may also react with amino acids within the P450 active site, or escape the active site and react with other cellular biomolecules.

**MI Complex**

Metabolite-Intermediate (MI) complexes exhibit time-dependent decreases in enzyme activity, but do not irreversibly modify the P450. An MI complex can be considered a metabolite “trapped” within the active site and often involve
reactive nitroso or carbone intermediates [172-178]. The compound is coordinated to the heme iron and experiences a stable geometry which greatly decreases the rate of release. The coordination of the substrate to the heme iron results in characteristic spectral changes causing an increase in absorbance at approximately 455 nm [179]. For many compounds the MI complex can be reversed through the addition of potassium ferricyanide [178, 180, 181]. While the enzyme is not truly modified and activity is restored upon removal of the trapped intermediate, this does not imply MI complexes are benign. Numerous compounds that form MI complexes with CYP3A4 have been shown to cause clinically significant drug-drug interactions, e.g. erythromycin [182-187], nicardipine [188-190], and troleandomycin [177, 191, 192]. Erythromycin is a weak inhibitor of CYP3A4, but during the N-demethylation of erythromycin, a nitroso intermediate is formed which ligates the heme iron and forms a stable complex which prevents further catalysis by the enzyme. Whereas formation of an MI-complex can cause significant drug-drug interactions, it is unlikely that this would directly cause increased toxicity.

**Inhibitory Metabolite**

Time dependent inhibition can be observed in an in vitro setting due to the generation of a metabolite which is a stronger inhibitor than the parent compound. For example, both fluoxetine and its demethylated metabolite norfluoxetine potently inhibit CYP2D6 with similar observed IC50 values. However, norfluoxetine inhibits CYP3A4 approximately 4- to 10-fold more potently than fluoxetine [193-195]. This gives the appearance of fluoxetine having weak CYP3A4 time-dependent inhibition as fluoxetine is converted to norfluoxetine during the incubation. Such a mechanism would not be expected to directly influence hepatotoxicity.

**Assessing Covalent Binding of Reactive Drug Metabolites by Proteomics**

Electrophilic drug metabolites are capable of forming stable covalent protein adducts [196, 197]. This remains a poorly understood cause of drug toxicity. Even though in vitro trapping experiments help in the identification of reactive metabolites, they do not help in determining the proteins that are modified. Sometimes short-lived reactive metabolites can react with nucleophilic residues present in the enzyme’s active site leading to enzyme inactivation, but they may be too reactive to diffuse out of the enzyme active site, thus precluding detection using trapping techniques [198]. The identification of proteins modified by reactive metabolites allows information to be obtained at the structural level and can aid in understanding the role between covalent binding and drug-induced toxicity. Similar patterns in protein modifications among reactive metabolite species may help in predicting the outcome of toxicity with greater specificity.

The analysis of proteins modified by reactive metabolites is an analytical challenge due to the complexity of protein samples. Since protein targets are diverse and adducted proteins represent only a small fraction of all the proteins present in the cell, this scenario was rightly described as a “needle in a haystack problem” by Liebler et al. [199]. Detection of target proteins is aided if they can be enriched from other proteins for better analysis and there are promising new affinity chemistries that can be utilized. Traditional in gel digestion methods have successfully identified several protein targets of electrophilic drugs. For instance, with the advent of proteomics in late 1990’s, 23 liver proteins modified by acetaminophen were determined using Two-dimensional (2D) gel-based proteomic approaches [200]. By using autoradiography the location of adducted proteins were revealed and the adducted protein spots were excised from the gel and digested with trypsin. Peptides were then analyzed by MALDI-MS to identify proteins. The 2-D gel methods based on autoradiography is dependent on the specific activity of the radiolabeled drug and is biased towards proteins that are abundantly expressed [199]. A variation that does not use radioactivity is the use of antibodies directed against adducted protein (2-D Western Blotting). Labeled secondary antibodies can then be used to identify bands that contain adducted protein.

As an alternative to using electrophoresis to separate proteins, complex samples can be enriched through the incorporation of affinity tags such as biotin into the molecule. The molecule proposed to form a reactive metabolite is synthesized with a biotin tag and when the mixture of adducted and unadducted proteins are passed through a streptavidin or avidin column, the biotin labeled adducted proteins are captured [201-203]. Elution of these proteins followed by digestion enriches the sample for adducted peptides. Alternatively the proteins can be digested prior to loading them onto the affinity column [199]. The incorporation of fluorescent or biotin tags into the test molecule can prevent metabolism or alter the metabolite profile and prevent the formation of the reactive metabolite even if the tag is spatially removed from the portion of the molecule that is responsible for the formation of the reactive metabolite.

If the target protein can be expressed recombinantly, analysis can be simplified. For example, upon incubation of recombinant CYP3A4 with raloxifene, a single raloxifene was found to bind by examining the mass shift of intact 3A4 using MADLI-TOF whole protein mass spectrometry. Despite observing a mass shift indicative of the binding of a single raloxifene, upon enzymatic digestion of the adducted 3A4, followed by tandem MS, two sites of addition were elucidated for raloxifene, Cys 239 and Tyr 75 [204, 205]. Using similar techniques, 17-R-ethynylestradiol was shown to modify Ser360 in both P450 2B1 and 2B6 [206].

Pähler et al. have described a targeted proteomic approach to enrich and identify peptides modified by reactive drug metabolites of 2-amino-pyrimidine obtained from rat liver microsomal proteins [198]. Their procedure involved the use of an equimolar mixture of non- and radioisotope-labeled (14C) substrate. The protein mixture was resolved by 1D SDS-PAGE and protein bands were visualized by autoradiography, and in-gel digested with trypsin. Peptides were then separated by strong cation exchange (SCX) chromatography and radioactivity containing fractions analyzed by nanoLC-MS/MS.

Non-enriched shotgun proteomics can be used to identify proteins in a complex mixture with minimal specific enrichment. The protein mixture is digested to peptides, followed by their analysis by LC-MS-MS to generate tandem MS-MS
Genotoxicity

The role of metabolites in genotoxicity has long been appreciated. Genotoxicity is typically evaluated using at least one and usually a battery of in vitro assays as a preliminary evaluation and followed up with in vivo evaluation during the preclinical toxicology package in IND enabling studies. One of the most common in vitro assays is the Ames test which evaluates the mutagenic potential of compounds in engineered Salmonella typhimurium bacterial cultures [215-218]. Bacteria are modified to be deficient in histidine biosynthesis and must be cultured in histidine augmented media. Strains can be modified to identify compounds that initiate point mutations or frame shifts that re-instate the ability to synthesize histidine. In practice, the bacteria are plated at a low density in media containing a pH indicator with sufficient histidine to allow for minimal growth. Metabolic processes acidify the media and wells containing revertant cells continue to grow causing further acidification of the media. Due to the presence of the pH indicator, revertant wells can be identified by simple visual inspection. Compounds that cause higher than background rates of inversion are considered potential mutagens.

Hepatic S9 (the soluble fraction from liver homogenate after centrifugation at 9,000 x g) can be added to the Ames test to improve predictions by incorporation of a metabolism component [219]. Known carcinogens such as benzo[a]pyrene are positive only upon the addition of S9 [220, 221]. Benzo[a]pyrene does not directly modify DNA. CYP1A1 catalyzes the formation of an epoxide which is hydrolyzed to form a diol [222-224]. The diol metabolite is further metabolized to benzo[a]pyrene-7,8-dihydrodiol-9,10-epoxide which covalently binds DNA. Human S9 or induced rat S9 microsomes are used. Rats can be treated with Aroclor 1254 for two to three days prior to harvesting the liver and preparation of S9. The specific content of many of the drug metabolism enzymes is greatly increased in Aroclor 1254 induced S9 [225, 226].

IN VIVO METHODS

There are several assays that can be utilized to examine in vivo generation of reactive metabolites and try to gauge their toxicity risk. Based on the results from in vitro experiments, additional in vivo experiments may be warranted. It is important to have realistic expectations when evaluating in vivo data. With the current state of knowledge, experiments to verify observed in vitro reactive metabolites are possible through several mechanisms. Utilization of this data to predict the toxicological implication is more difficult. The focus will be on rodent and human clinical studies. Other mammals including primates can be similarly used.

Among FDA approved drugs that are associated with idiosyncratic hepatic toxicity, rates of drug-induced hepatic failure are frequently one in ten-thousand to one in a few hundred thousand [227-231]. Even among drugs that were withdrawn due to their risk of hepatotoxicity the typical incidence rate of hepatic failure was one case per ten-thousand patients. This represents a significant challenge for discerning those compounds with higher probabilities of causing serious toxicity from those that may cause mild hepatic issues that do not progress to more serious toxicity. Currently studies can be designed to evaluate multiple factors that correlate with serious adverse events, but none have 100% predictive value. Fig. (1) presents a map of physiological steps associated with chemical-induced immune-mediated toxicity. Most of the common experimental observations used to evaluate reactive metabolites focus on the first few steps, but interruptions at multiple stages along the path would be expected to mitigate toxicity. Experimental observations are likely to provide parameters to incorporate in risk assessment, but based upon current understanding, the predictive value is not sufficient to provide a definitive answer on the likelihood of an individual compound to cause rare idiosyncratic toxicity prior to exposure of high numbers of patients.
Reactive Metabolites – Experimental Evaluation

**In vivo Genotoxicity**

As part of the preclinical toxicology evaluation, *in vivo* genotoxicity is evaluated using a micronucleus assay [232-240]. During cell division, the parent cell should result in two daughter cells each containing identical copies of DNA. In the presence of compounds or reactive metabolites that can cause chromosomal damage, it is possible to have one of the daughter cells lacking a part or all of a chromosome. The missing chromosomal fragment can develop a nuclear membrane where some cells appear to have one large and one small nucleus (the micronuclei). Comparison of the percentage of micronucleus positive cells in the treated vs. control groups provide a readout of genotoxicity. The assay is typically done with erythrocytes in the peripheral blood or bone marrow [241-245]. Erythrocytes are anucleated which simplifies detection because the micronuclei are not extruded as efficiently.

**In vivo Binding Studies**

Experimental options focusing on binding are quite similar to those used *in vitro*, but incorporate factors such as pharmacokinetics and tissue levels which are not easily replicated *in vitro*. Similarly, the *in vivo* systems have a full complement of detoxification pathways such as antioxidants, glutathione S-transferase, superoxide dismutase/catalase, glutathione peroxidase, epoxide hydrolase, etc. Alternative methodologies to evaluate covalent modification are presented. Similar to the *in vitro* systems, radiolabel incorporation is the most frequently utilized methodology [87, 246, 247].

**Detection of Glutathione or Mercapturate Adducts**

Positive results from glutathione adduction studies *in vitro* can be easily followed up with *in vivo* studies using preclinical species or humans. This is typically done through obtaining samples of urine and either bile or feces from dosed rats or humans and comparing to vehicle dosed controls [248-250]. In rodent studies it is also possible to get a liver sample to evaluate adduct levels in tissue [251]. *In vivo*, glutathione conjugated drugs are typically further metabolized and detected as mercapturic acid adducts [252-254]. The epsilon-peptide linkage between glutamate and cysteine is resistant to normal proteases which make glutathione adducts relatively stable in most tissue and allows for the direct detection of intact drug adducts. However, GSH-adducts can be enzymatically degraded by the enzyme γ-glutamyl transferase which removes the glutamate followed by further processing by cellular dipeptidases which remove glycine. The amine group of the cysteine is then acetylated to form a mercapturate. The tissue with the highest expression level of these enzymes is the kidney and in many cases the glutathione conjugate is exported from the liver then processed in the kidney where it can be excreted in urine or be reabsorbed by the liver and excreted in the bile [253]. Detection of mercapturate adducts in bile or feces does not necessarily indicate inter organ transport, especially in rodent where expression levels of γ-glutamyl transferase and cellular dipeptidases in the bile duct are sufficient to facilitate the degradation of GSH-adducts [253, 254]. Detection of mercapturates using mass spectrometry indicates that the compound was metabolized to a nucleophilic metabolite, but may not indicate a high level of protein modification [255]. The mercapturate could have been derived from the direct reaction with glutathione from cellular glutathione pools, enzymatically catalyzed by glutathione S-transferase, or even derived from adducted proteins where cysteine is covalently modified. If the reactive electrophile is a substrate for glutathione it is possible to have one of the daughter cells each containing identical copies of DNA. The amine group of the cysteine is then acetylated to form a mercapturate. The tissue with the highest expression level of these enzymes is the kidney and in many cases the glutathione conjugate is exported from the liver then processed in the kidney where it can be excreted in urine or be reabsorbed by the liver and excreted in the bile [253]. Detection of mercapturate adducts in bile or feces does not necessarily indicate inter organ transport, especially in rodent where expression levels of γ-glutamyl transferase and cellular dipeptidases in the bile duct are sufficient to facilitate the degradation of GSH-adducts [253, 254]. Detection of mercapturates using mass spectrometry indicates that the compound was metabolized to a nucleophilic metabolite, but may not indicate a high level of protein modification [255]. The mercapturate could have been derived from the direct reaction with glutathione from cellular glutathione pools, enzymatically catalyzed by glutathione S-transferase, or even derived from adducted proteins where cysteine is covalently modified. If the reactive electrophile is a substrate for glutathione S-transferase it is possible to have one of the daughter cells each containing identical copies of DNA. The amine group of the cysteine is then acetylated to form a mercapturate. The tissue with the highest expression level of these enzymes is the kidney and in many cases the glutathione conjugate is exported from the liver then processed in the kidney where it can be excreted in urine or be reabsorbed by the liver and excreted in the bile [253].

**Whole-Body Radiography**

Quantitative Whole Body Autoradiography (qWBA) is an imaging technique used to provide information on the localization of radiolabeled compounds in lab animals [257-
It is usually performed on only advanced safety-assessment candidates during the preclinical drug development stage, and involves imaging of the whole-body, and/or specific tissues or organs. The data can provide important information about tissue pharmacokinetics, tissue distribution, metabolism, drug-drug interactions, elimination, and clearance and is especially helpful in pinpointing the site of compound retention [261]. In these experiments, animals are given a single dose of radioactive compound ($^{14}$C, $^{125}$I, $^{3}$H) and may be visualized live or euthanized at different time points. Autoradiograph results show the localization of the radiolabeled drug at different times and the amount of radioactivity from tissues is quantitated by comparative dosimetry.

qWBA studies can also provide tissue samples for analysis of reactive metabolites that may bind covalently to proteins in various tissues or organs over time [257]. The use of whole-body sections and/or residual frozen tissue can supply samples for bioanalysis using radio-HPLC and mass spectrometry to identify the adduct and modified proteins. For example, drugs like amodiaquin (withdrawn) and clozapine (black box warning), which cause severe and life-threatening agranulocytosis, have been evaluated by qWBA and showed retention of radioactivity in the bone marrow and spleen up to 168 hours after administration, Fig. (3) [262]. If the retention of these radiolabeled compounds is correlated with its covalent binding data, qWBA can be used for toxological risk assessment to estimate body burden in larger species or humans.

![Fig. (3)](image)

**Fig. (3).** Whole-body radiographs demonstrating compound retention in the liver and bone marrow 72 and 128 hours after a single 3mg/kg oral dose of radiolabeled amodiaquine (A) and clozapine (B) [31], reprinted with permission.

One of the main advantages of this technique is that it provides a comprehensive visual of radiolabel distribution in organs and tissues. Unlike other techniques such as tissue dissection/liquid scintillation counting where all tissue samples are not routinely tested, qWBA provides the distribution of radioactivity in all organs and tissues of an intact animal carcass. Phosphor-imaging technique has replaced photographic techniques due to its broad dynamic range and sensitivity, and its ability to provide qWBA data in days versus weeks to months [263-266]. Some of the shortcomings of this method are that these experiments require the synthesis of radiolabeled drug, which can be practically challenging. Thus this approach is not amenable for early discovery projects where the best-in-class compound is likely to change quickly. Another disadvantage is that this technique provides data on the concentration of radioactivity only, and it is not possible to determine if the radioactivity represents parent drug, metabolites and/or degradation products without further evaluation.

### TOXICITY BASED IN VIVO STUDIES

The literature regarding severe drug-induced hepatic toxicity can be confusing. Part of this is due to differences in the intended audiences. Publications targeting clinicians have often advised caution about over-interpreting measures of hepatic toxicity such as increases in serum transaminases [267-271]. This is primarily an underlying argument of probability. Most patients that demonstrate mild or moderate increased serum transaminases in the first few weeks after initiation of treatment revert to normal levels through tolerance or unknown adaptation mechanisms with continued exposure. For example, the drug tacrine is commonly associated with mild hepatic injury that can cause significant increases in plasma ALT levels three to six weeks after initiation of treatment, often 5- to 10-fold the upper limit of normal. However, severe hepatic injury caused by tacrine is exceedingly rare and for most patients ALT levels return to normal by twelve weeks [268]. If clinicians withdrew treatment with the first observation of elevated ALT, they may be denying their patient beneficial treatment. On the opposite spectrum are publications targeting drug discovery scientists where compounds such as troglitazone are highlighted. Elevated ALT levels in troglitazone patients are both lower and less frequent than what is seen with tacrine, but in the three years between the approval of troglitazone and its removal from the market, at least 94 cases of troglitazone-induced liver failure were reported to the FDA [272]. The relative probability of liver failure for an individual patient will be quite low for a compound exhibiting idiosyncratic hepatic toxicity, but the overall incidence may be too high to remain on the market [227, 273-276].

### Measures of In Vivo Hepatic Damage

During or upon completion of multi-day dosing studies, it is common to collect blood which can be evaluated for markers of hepatic damage. Numerous markers have been historically utilized to evaluate liver damage, with the four most common readouts being bilirubin, aspartate transaminase (AST), alanine transaminase (ALT), and alkaline phosphatase (ALP). The liver is not comprised of a single cell type nor is it uniform in its architecture. A hepatologist can detail numerous classifications to describe liver injury, but for most individuals there are four major categories to consider: hepatocellular, cholestatic, infiltrative, and autoimmune [269, 270, 275, 277-281].

Hepatocellular injury is a primary injury to the hepatocytes. In cases where hepatocyte membrane integrity is lost
conjugated bilirubin is exported to the bile. Bilirubin increases may be small or undetectable, but are elevated as the extent of hepatic injury progresses. Bilirubin is cleared from the blood through conjugation and excreted from the hepatocyte into the bile. There is excess hepatic capacity to carry out this function so increases in bilirubin are not seen with mild hepatoellular injury. Diseases that primarily affect hepatocytes, such as viral hepatitis, will cause disproportionate elevations of ALT and AST compared with alkaline phosphatase or bilirubin [229, 279, 282].

Cholestatic hepatic injury, sometimes referred to as obstructive liver injury, is primarily a bile duct injury [279, 282, 283]. This may be due to gall stones, tumors obstructing bile flow, or due to significant inflammation or necrosis of the cells forming the bile canaliculi which causes leakage of alkaline phosphatase. In addition to increases in alkaline phosphatase in the plasma, cholestatic injury is also associated with elevated bilirubin. This is not surprising that blocking biliary excretion leads to a buildup of bilirubin because conjugated bilirubin is exported to the bile.

Infiltrative hepatic injury is characterized by the liver being invaded or replaced by non-hepatic substances. This is commonly neoplasm (an abnormal growth or division of cells which may be benign or malignant) or amyloid [284-287]. Infiltrative hepatic injury is not usually associated with reactive metabolites or drug-induced toxicity but has a similar increase in alkaline phosphatase as seen with cholestatic injury. However, bilirubin changes tend to be reduced. Imaging studies or liver biopsy may be required to unambiguously diagnose infiltrative hepatic injury along with its cause.

Autoimmune mechanisms can have a hepatocellular pattern if hepatocyte damage is involved (autoimmune hepatitis) or cholestatic if the immune system targets the biliary duct (primary biliary cirrhosis) [288-291]. Immune-mediated hepatic toxicity is discussed in greater detail in a separate section.

Measureable Hepatic Markers

Elevated transaminases are indicative of hepatocellular injury. Both ALT and AST are at high concentrations in the liver cytosol. In cases where hepatocyte membrane integrity is lost, they can be detected at elevated levels in the blood. AST is not favored as a primary readout of liver function because it is also present at high levels in muscle leading to fluctuations in AST for benign reasons such as muscle mass and exercise [292-297]. Both transaminases require the prosthetic group pyridoxal-phosphate which is the active form of vitamin B6 [298-300]. Vitamin B6 deficiency is rare, but would be expected to decrease activity and mask the elevation of plasma transaminase levels [298]. Other interactions have been noted with chemical inhibitors. For example, isoniazid associated hepatic damage is largely masked because isoniazid inhibits ALT [301].

Alkaline phosphatase is a hydrolase enzyme which catalyzes the removal of phosphate from numerous compounds. It is primarily elevated in the bile duct, kidney, placenta, and bone [302, 303]. Liver levels have also been reported to be elevated but this may be due to bile canaliculi which infiltrate the liver and collect bile excreted from hepatocytes. Alkaline phosphatase derived from bone and bile duct can be differentiated based upon the reduced thermal stability of bone derived alkaline phosphatase [303]. Alkaline phosphatase levels naturally increase in pregnant woman and do not necessarily imply liver dysfunction.

Elevated bilirubin levels (jaundice) are often associated with damage to hepatic biliary excretion [275, 277, 282, 302]. Bilirubin is not released from the liver, but is from the breakdown of heme containing proteins, particularly hemoglobin and myoglobin. In the liver, bilirubin is conjugated by glucuronidation and sulfation and then exported into the bile [304-309]. This creates an important caveat in the evaluation of liver damage: increases in total bilirubin levels are not necessarily indicative of liver dysfunction and can be due to increased hemolysis. Elevated plasma concentration of conjugated serum bilirubin (not total) implies liver disease/dysfunction [310, 311]. However, most commercial blood chemistry analyzers provide tBIL for “total bilirubin”. When the two are differentiated, the term “direct bilirubin” is used to refer to conjugated bilirubin and “indirect bilirubin” for unconjugated. Bilirubin should not be present in the urine of a healthy patient, but conjugated bilirubin may be present in urine when biliary export is impaired leading to elevation of conjugated bilirubin in the blood.

Clinical Implications

In most cases where increases in ALT are detected, this indicates damage to hepatocytes. However, the liver is an amazingly resilient organ and mild damage is rarely associated with a significant toxicological event. In clinical trials, ALT is commonly measured and compared to the control group [271, 312-315]. A typical threshold of significance is 3-times the upper limit of normal (ULN). Higher levels of 5- and 10-times ULN are often used to indicate more significant hepatic damage.

Almost all compounds that cause hepatic failure also had an associated increase in the percentage of patients with ALT elevation in clinical trials. However, the opposite (increased ALT indicating imminent severe toxicity) is not always true. Notable compounds that cause ALT elevations but are rarely associated with severe adverse events are tacrine, aspirin, heparin, and statins [7, 271, 316-318]. Hy Zimmerman provided evidence that increases in transaminases were most predictive of serious hepatic toxicity when they were associated with a concomitant measure of altered liver function (measured as a greater than 2-fold increase in serum bilirubin). This observation has been refined and is sometimes referred to as Hy’s Law [319-326]. Unfortunately, by the time hepatocellular injury is severe enough to lead to greater than 2-fold increase in serum bilirubin there is approximately a ten percent chance the patient will progress to liver failure resulting in death or liver transplant. This represents a serious deficiency in our current knowledge base. Earlier biomarkers that describe the loss of liver function/capacity prior to reaching 2-times increase in bilirubin concentration may provide the opportunity for earlier intervention without the associated risk of patient mortality.

Summary of Hy’s Law: A drug that causes serious hepatic injury will generally:
Show an increased incidence in elevated plasma transaminases (>3xULN) than the control

Lead to total bilirubin >2xULN in a small percentage of patients without initial elevation of serum ALP indicating the increase in bilirubin is due to a decrease in hepatocellular function and not due to obstruction of bile flow. Complicating factors such as hepatitis, underlying liver disease and co-medications should be factored in, as should the presence of increased red blood cell lysis as this would increase bilirubin levels independent of liver function.

In the FDA issued guidance for evaluation of drug-induced liver injury, Hy’s law is prominently featured and used to estimate the expected incidence of serious hepatic events [7, 327].

In the clinical trials of a hypothetical compound where 6,000 total patients are dosed for extended periods and five Hy’s Law cases are observed (individuals with >3xALT and 2xbilirubin) the estimated rate of hepatic failure would be 1/12,000. This is based on 5/6,000 (1/1,200) patients exceeding Hy’s law and an estimate that 10% of Hy’s law cases proceed to hepatic failure. When there is an increase in the overall rate of >3xALT but no patients with associated bilirubin increases a different calculation is used. Evaluation of the same hypothetical 6,000 patients with an observed increase in the percentage of treated patients with >3xALT, but assuming no Hy’s Law cases are observed, the estimated rate of hepatic failure would be less than 1/20,000. This is because in the absence of findings the calculation is based upon a 95% confidence level. Using the “rule of 3”, three times more samples are required to say with 95% confidence that the lack of an observed case was not due to statistical chance. If the rate of hepatic failure were to equal 1/20,000 then 1/2,000 patients would be expected to be Hy’s Law cases, and three times that number (6000 patients) would be needed to say with 95% confidence that the expected rate of hepatic failure is less than 1/20,000 patients [7].

Immune-Mediated Toxicity

Immune-mediated toxicity has been implicated with numerous compounds and almost all of these compounds have been shown to generate reactive metabolites [328-335]. The evidence for immune involvement is strong and includes a delayed onset of toxicity with rapid recurrence upon rechallenge, detection of drug specific antibodies and drug responsive T-cells, macrophage recruitment to the site of toxicity, positive skin hypersensitivity tests, and positive correlations between patient susceptibility to drug-induced toxicity and human leukocyte antigen (HLA, the protein that binds proteolytic peptides and presents them to T-cells) [333-337]. The ability to predict what compounds will and will not elicit immune-mediated toxicity or to predict the frequency of severe events prior to entering the clinic is not currently validated. Severe immune-mediated toxicities where the patient fails to gain tolerance are not common, but can be particularly severe and have high rates of fatality.

The rate of drug-induced cholestatic liver disease with fluvoxacinilin is approximately 1/10,000 [338]. A genome-wide association study was conducted with patients that had experienced fluvoxacinilin-induced liver failure and a comparison was made to a control group of similar ethnicity. Sixty-three of seventy-four liver injury patients had a HLA-B*5701 genotype compared to four of sixty-four controls. The researchers concluded that the odds of a patient with HLA-B*5701 genotype experiencing severe liver injury from taking fluvoxacinilin were 80.6-times higher than for patients with alternative genetic HLA profiles [339]. While the 80-fold increase in risk is exceptionally large, there is still under a one percent chance for an individual to experience severe fluvoxacinilin-induced liver injury even if they do have the HLA-B*5701 genotype. Obviously there are additional factors involved in distinguishing what patients will and will not experience toxicity.

Other studies have demonstrated an immune component in serious drug-induced skin toxicity from compounds that are known to generate reactive metabolites or otherwise covalently modify proteins [340]. The risk of moderate carbamazepine hypersensitivity increased 12-fold and severe Stevens–Johnson syndrome or toxic epidermal necrolysis increased 26-fold for patients with HLA-A*3101 [341]. Additional studies associating antigen presentation and drug toxicity have been published. A partial list includes: lapatinib [334], abacavir [342], nevirapine [343], and lumiracoxib [344]. Results from numerous association studies were evaluated to see if there were common features in peptide binding between those HLA that have been associated with drug-induced toxicities and those that have not. Researchers found that within HLA class I alleles those that were associated with toxicity were actually less similar than the control set of HLA analyzed [345]. Similarity between HLA class II alleles associated with drug-induced toxicity were only slightly more similar than control [345]. This implies that antigen presentation is either not due to a common subset of presented antigens or that there are additional factors that we do not yet understand.

Despite not understanding all of the specifics leading to immune-mediated toxicity, patient benefit is starting to be realized. Two of the strongest HLA associations are with abacavir hypersensitivity (HLA-B*5701) and severe cutaneous injury due to carbamazepine in the Han Chinese and Thai populations (HLA-B*1502) [346, 347]. Through genetic pre-testing, reduced rates of toxicity have been achieved for both drugs [335]. One of the major obstacles to implementing such tests on a wider scale is the role of ethnicity. Whereas HLA-B*1502 increases the chance of serious cutaneous injury by over 100-fold in the Han Chinese, the increased risk of carbamazepine-induced toxicity due to HLA-B*1502 is minimal among Caucasian populations [348, 349]. Understanding if there are both primary and secondary drivers for toxicity and having the ability to appropriately weight their significance would be a major step forward.

Genome-wide association studies have been valuable in identifying increased-risk populations for several drugs, but this is a retrospective analysis [350]. Patient injury is required in order to have patient pools that are positive and negative for toxicity. The ideal situation would be to identify populations at increased risk prior to entering the clinic or as soon as possible after human patients are exposed. There are
additional tests that have demonstrated drug-induced immuno-mediated toxicity and statistically significant differences have been demonstrated between patients that experienced toxicity and untreated controls, but even with the best tests currently available, not all patients in the toxicity group give positive results. In some cases, patients that have been treated but show no observable toxicity have come back positive for an immune response. It is unclear if detection of an immunological response in non-symptomatic groups foretells an increased likelihood for future toxicity or what percentage of patients would experience serious adverse events.

**Lymphocyte Transformation Test (LTT)**

Evaluation of lymphocyte proliferation in response to drug challenge is much more common in Europe and Japan than in the United States. Blood collected from patients that experience potential immune-mediated drug-induced toxicity is processed to isolate lymphocytes which are evaluated immediately or frozen. Lymphocytes are then cultured in the presence or absence of the suspected drug and lymphocyte proliferation is evaluated [351, 352]. A common readout of proliferation is to measure radioactive thymidine incorporation. Generally, stimulation greater than two-fold is considered a positive response. Phytohemagglutinin can be used as a positive control to demonstrate that induction of proliferation is possible.

The LTT test is built upon the natural propensity for activated lymphocytes to increase their proliferation in the presence of the specific antigen that they are primed to recognize. In the LTT assay, lymphocytes are incubated with the parent drug and not with drug-protein or drug-peptide conjugates. It is generally accepted that a small molecule cannot directly illicit an immune response leading to the generation of specific antibodies. Instead, the molecule is presented as a larger complex, typically through a covalently modified peptide or protein (hapten hypothesis). The reverse is not true. An antibody (at least polyclonal antibodies), originated in response to haptenized proteins, do have the potential to directly bind free small molecules and their metabolites.

Anti-clozapine antibodies were raised by trapping the reactive clozapine metabolite with N-acetylcysteine (NAC). The NAC was then used to link clozapine to Keyhole limpet hemocyanin (KLH) and used for antibody generation [353]. Serum was tested and shown to recognize clozapine conjugated proteins. Of importance to the LTT assay is that antibody binding could be inhibited by clozapine, its close analog olanzapine, or clozapine-NAC which was the actual antigen used for antibody generation. Binding was not inhibited when structurally unrelated compounds were tested [353]. Similarly, antigen binding of antibodies generated against hydroxyfluperlapine, the metabolite of fluperlapine, could be inhibited by both fluperlapine and hydroxyfluperlapine [111]. The importance of these findings is that activated lymphocytes, which recognize antigens through surface bound antibodies, would be expected to behave similarly and may react with the parent drug even if the actual antigen was a metabolite bound to protein.

The previously described interactions with fluperlapine and clozapine all utilized 20-100 μM drug to inhibit antibody binding and it could be argued that these are not clinically relevant concentrations. However, small molecule-antibody interactions are possible at much lower concentrations and have even been utilized to establish analytical diagnostics. Commercial ELISA kits are available from abcam (ab20768), Pierce (PC10-242.6), Beckman Coulter (Emit 2000 procainamide assay), and other providers for quantitaiton of free procainamide in patient plasma. These are based on monoclonal antibodies that were generated using a procainamide-KLH immunogen through covalently attaching the amine group of procainamide to acidic amino acids of KLH. Through the fusion with myeloma cells, monoclonal antibodies were obtained and expanded in the presence of free procainamide.

The percentage of lymphocytes that sufficiently recognize free drug is likely different for individual drugs and for individual patients. This is likely the reason for the wide range in responders in LTT evaluation.

When evaluating increasing numbers of patients, LTT clearly demonstrates increases in the rate of positive responses for both cutaneous and hepatic immune-mediated drug-induced toxicity [354-362]. However, the rate of positive responders can fluctuate greatly and the severity of toxicity is not necessarily tied to a positive LTT. The drug lamotrigine is associated with severe cases of cutaneous injury including Stevens-Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN) [360, 363]. Using LTT, lamotrigine-naive control patients were compared to lamotrigine treated patients without observable adverse reaction, with mild epidermal eruption, active SJS/TEN, and patients recently recovered from SJS/TEN. Positive LTT did not trend with the severity of toxicity; however, positive tests were detected for seven of twenty nine patients that had a reaction and in zero of twelve lamotrigine-naive controls. Interestingly, one of six patients that had been exposed to lamotrigine but showed no reaction had a positive test [364].

In a test of 95 suspected cases of drug-induced liver injury associated with 33 different drugs, 25/95 of the liver injury patients had a positive LTT, whereas positive tests were obtained for 0/106 healthy controls. When the prostaglandin inhibitor indomethacin was added to the LTT cell media to inhibit decreased lymphocyte proliferation due to excreted prostaglandins, the percentage of positives in the LTT increased to 53/95 [365].

**Patch/Skin Hypersensitivity Test**

Positive skin hypersensitivity tests have been observed for several drugs including acetaminophen. The patch test applies drug to the skin and tests if the dermal dendritic cells (Langerhans cells), or other lymphocytes that migrate to the area are capable of responding to the presented antigen. Patch tests have been utilized to demonstrate drug hypersensitivity [366, 367]. In contrast to the LTT assay which offers absolute safety due to being an *in vitro* evaluation, the patch test has the potential for complications. While the FDA does not have firm regulations regarding the use of the patch test to evaluate hypersensitivity, they do specify guidelines for the area of testing, the use with abacavir, citing concerns of patient safety. Instead of using the patch test to evaluate sensitive vs. non-sensitive individuals, they recommend testing for the HLA-B*5701 allele and caution against abacavir use in such patients.
Elevation of Eosinophils

Through the collection of patient blood samples, blood counts can be performed to evaluate eosinophil count. Eosinophils are a portion of the circulating leukocyte population and elevations are often associated with allergic reaction. Eosinophil count is not causal with respect to the suspected treatment drug and can be elevated by certain parasites or non-related allergens [368, 369].

Eosinophilia is often referred to as Drug Reaction with Eosinophilia and Systemic Symptoms (DRESS). DRESS onset is typically three to eight weeks after drug initiation and is frequently associated with fever and skin eruption [370]. Some of the drugs most frequently observed in DRESS such as allopurinol, phenytoin, carbamazepine and dapsone are associated with other immune-mediated drug-induced toxicities and have been shown to generate reactive metabolites [369]. DRESS has been reported to positively correlate with viral reactivation.

Re-Activation of Viruses

Viral reactivation has been utilized in Japan more so than in other countries. An increase in herpes virus (HHV-6) was shown to be strongly associated with drug-hypersensitivity reactions, but other viruses have also been implicated [371-385]. The largest volume of data is available for reactions with anti-convulsants. The mechanism of the association is largely still hypothetical and this may explain why it has not been adopted in a wider setting. The simplicity of a rapid PCR based analysis along no additional patient risk is quite attractive. HHV-6 reactivation has a short delay from the initial signs of hypersensitivity [386] and differing results are reported for minor hypersensitivity vs. more severe reactions. While HHV-6 has been most widely studied, reactivation of three other herpes viruses: HHV-7, Epstein-Barr virus (EBV or HHV-4) and cytomegalovirus (HHV-5) have also been reported.

In one report HHV-6, HHV-7 or Epstein-Barr virus were reactivated in 76% of patients with DRESS associated with taking carbamazepine, allopurinol, or sulfamethoxazole. In all patients, circulating CD8+ T lymphocytes were activated with elevated secretion of cytokines [368].

Both the mechanism of increase and the predictive value of viral reactivation are unknown. While reactivation may track well for DRESS, it is unclear if the association is as strong with other drug-induced toxicities or if there is a threshold effect, e.g. is viral reactivation only seen with advanced toxicity. An ideal marker would be capable of identifying potentially toxic compounds early in clinical development and not be limited to classifying the mechanism of observed toxicity. The value of determining levels of HHV-6 or Epstein-Barr virus after 4-8 weeks dosing in treated vs. control patients during clinical trials is uncertain.

Anti-Self Antibodies

Several antibodies have been associated with autoimmune diseases including drug-induced autoimmune diseases. These include antinuclear antibody (ANA, also called anthistone antibody), anti-Smooth muscle antibody (ASMA), liver/kidney microsomal antibody (LKM-1, LKM-2, LKM-3), anti-soluble liver antigen (SLA), anti-actin antibody (AAA), and anti-mitochondrial antibody (AMA) [47, 138, 155, 156, 328, 387-390]. ASMA and AAA have been reported to associate with disease state in autoimmune hepatitis [391, 392]. Increased levels of autoantibodies have been reported in cases of chronic drug-induced liver injury relative to those seen for acute injury [388]. Autoantibodies are likely detectable as a marker of chronic drug-induced injury and may have little to do with the drug being presented as an antigen or specifically altering antigen processing. Patients with suspected drug-induced liver injury taking three biological anti-TNF-α agents, infliximab (n = 26), etanercept (n = 4), and adalimumab (n = 4) were evaluated for autoantibodies and 67 percent of the patients tested positive for antinuclear and/or smooth muscle antibodies. 88 percent of patients testing positive for ANA or ASMA had clear features of autoimmunity [393].

In contrast to the before mentioned autoantibodies, anti-P450 antibodies appear to be in direct response to hapten formation [155, 156, 390]. The drug tienilic acid was found to illicit the formation of anti-CYP2C9 antibodies. This correlates with tienilic acid metabolism as it is oxidized by CYP2C9 and is a potent mechanism based inactivator of 2C9 activity [138]. Similar to tienilic acid, anti-CYP2E1 antibodies were detected in 25 of 56 patients diagnosed with halothane-induced hepatitis [155] and anti-CYP1A2 antibodies have been found associated with disulfiram [156]. A recent paper reports the detection of both anti-isoniazid and anti-P450 antibodies in a patient with isoniazid-induced liver failure [328]. Where these accounts appear to have mechanistic utility, all deal with the appearance of anti-P450 antibodies in patients diagnosed with severe drug-induced injuries by comparing them to control patients. It would be of great benefit to know if less severely injured patients would have had detectable anti-P450 antibodies in there sera. Whereas viral infection or other causes of liver injury may increase the potential for detectable levels of anti-self antibodies, anti-P450 antibodies would be expected to correlate to metabolic activation. It is unknown if evaluation of anti-P450 antibodies during clinical development of compounds demonstrating time-dependent inhibition would be beneficial. For example, would evaluation of anti-CYP2C9 antibodies during the development of tienilic acid (approved by the FDA in 1979 and withdrawn in 1982) have shown differences between the tienilic acid vs control groups?

CONCLUDING REMARKS

Because of the low incidence of severe drug-induced toxicity, it remains a difficult challenge to identify idiosyncratic hepatic toxins during clinical development. The ability to identify reactive metabolites has outpaced the ability to ascertain the toxicological implication. There is currently no way to unequivocally provide answers to questions about the ultimate toxicity of an individual preclinical compound that is metabolized to a reactive metabolite. Prevention of reactive metabolite formation or decreasing the body-burden through lowering the dose have been demonstrated to reduce the risk of idiosyncratic toxicity based on statistical analysis,
but there are numerous outlier drugs that have been beneficial to patients despite being associated with high body burdens of reactive metabolites. There is strong evidence that the immune system plays a prominent role in many chronic drug-induced toxicities. To date, most studies have been retrospective to evaluate the mechanism of toxicity but much of the methodology could be adopted at an earlier stage to evaluate the potential for immune-mediated drug-induced toxicity during clinical development. Evaluation of the potential for compounds to be recognized as an antigen, activate an immune response and influence tolerance will increase the predictive value of early studies where reduced numbers of patients are tested and decrease the number of compounds that are eventually removed from market or fail late in clinical development due to hepatic toxicity.

LIST OF ABBREVIATIONS

ALT = Alanine transaminase
DRESS = Drug reaction with eosinophilia and systemic symptoms
HLA = Human leukocyte antigen
IND = Investigational new drug
Keap1 = Kelch-like ECH-associated protein 1
KLH = Keyhole limpet hemocyanin
LDH = Lactate dehydrogenase
LTT = Lymphocyte transformation test
NAC = N-acetylcysteine
Nrf2 = Nuclear factor erythroid 2-related factor
P450 = Cytochrome P450
qWBA = Quantitative whole body autoradiography
SJS = Stevens-Johnson syndrome
TEN = Toxic epidermal necrolysis

CONFLICT OF INTEREST
The Authors report no financial conflicts of interest. Salaries were partially paid for by NIH grant 1R21DK091630 and P01DA033622 to MDC.

ACKNOWLEDGEMENTS
Declared none.

REFERENCES


[23] Lewis, D.F.; Ioannides, C.; Parke, D.V., A combined COMPACT and HazardExpert study of 40 chemicals for which information on mutagenicity and carcinogenicity is known, including the results of


Hong, F.; Sekhar, K.R.; Freeman, M.L.; Liebler, D.C., Specific patterns of electrophile addition trigger Keap1 ubiquitination and Nrf2 activation. *J. Biol. Chem.*. 2005; 280, (36), 31768-31775.


[285] Koay, E.B.; Lin, C.Y.; Tsai, S.L.; Lee, C.; Lin, C.N.; Sheu, M.J.; Kuo, H.T.; Sun, C.S., Type 1 autoimmune hepatitis in Taiwan:


[293] Crouch, J.Y.; Solomon, L.R., Inhibition of rat liver transaminases by low levels of acetaldehyde and the pharmacologic effects of B6 deficiency. *Clinical nephrology*. 1995, 43, (6), 405-408.


Spraggs, C.F.; Budde, L.R.; Briley, L.P.; Bing, N.; Cox, C.J.; King, K.S.; Whittaker, J.C.; Moorss, V.E.; Preston, A.J.; Stein, S.H.; Cardon, L.R., HLA-DQA1*02:01 is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer. J. Clin. Oncol., 2011, 29, (6), 667-673.


Reactive Metabolites – Experimental Evaluation


