

Toxicologic Pathology

<http://tpx.sagepub.com>

The Liver Toxicity Biomarker Study: Phase I Design and Preliminary Results

Robert N. McBurney, Wade M. Hines, Linda S. Von Tungeln, Laura K. Schnackenberg, Richard D. Beger, Carrie L. Moland, Tao Han, James C. Fuscoe, Ching-Wei Chang, James J. Chen, Zhenqiang Su, Xiao-Hui Fan, Weida Tong, Shelagh A. Booth, Raji Balasubramanian, Paul L. Courchesne, Jennifer M. Campbell, Armin Graber, Yu Guo, Peter J. Juhasz, Tricin Y. Li, Moira D. Lynch, Nicole M. Morel, Thomas N. Plasterer, Edward J. Takach, Chenhui Zeng and Frederick A. Beland

Toxicol Pathol 2009; 37; 52 originally published online Jan 26, 2009;
DOI: 10.1177/0192623308329287

The online version of this article can be found at:
<http://tpx.sagepub.com/cgi/content/abstract/37/1/52>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Society of Toxicologic Pathology](#)

Additional services and information for *Toxicologic Pathology* can be found at:

Email Alerts: <http://tpx.sagepub.com/cgi/alerts>

Subscriptions: <http://tpx.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

The Liver Toxicity Biomarker Study: Phase I Design and Preliminary Results

ROBERT N. MCBURNEY,¹ WADE M HINES,¹ LINDA S. VON TUNGLN,² LAURA K. SCHNACKENBERG,² RICHARD D. BEGER,² CARRIE L. MOLAND,² TAO HAN,² JAMES C. FUSCOE,² CHING-WEI CHANG,² JAMES J. CHEN,² ZHENQIANG SU,² XIAO-HUI FAN,² WEIDA TONG,² SHELAGH A. BOOTH,¹ RAJI BALASUBRAMANIAN,¹ PAUL L. COURCHESNE,¹ JENNIFER M. CAMPBELL,¹ ARMIN GRABER,¹ YU GUO,¹ PETER J. JUHASZ,¹ TRICIN Y. LI,¹ MOIRA D. LYNCH,¹ NICOLE M. MOREL,¹ THOMAS N. PLASTERER,¹ EDWARD J. TAKACH,¹ CHENHUI ZENG,¹ AND FREDERICK A. BELAND²

¹BG Medicine, Inc., Waltham, MA, USA

²National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR, USA

ABSTRACT

Drug-induced liver injury (DILI) is the primary adverse event that results in withdrawal of drugs from the market and a frequent reason for the failure of drug candidates in development. The Liver Toxicity Biomarker Study (LTBS) is an innovative approach to investigate DILI because it compares molecular events produced in vivo by compound pairs that (a) are similar in structure and mechanism of action, (b) are associated with few or no signs of liver toxicity in preclinical studies, and (c) show marked differences in hepatotoxic potential. The LTBS is a collaborative preclinical research effort in molecular systems toxicology between the National Center for Toxicological Research and BG Medicine, Inc., and is supported by seven pharmaceutical companies and three technology providers. In phase I of the LTBS, entacapone and tolcapone were studied in rats to provide results and information that will form the foundation for the design and implementation of phase II. Molecular analysis of the rat liver and plasma samples combined with statistical analyses of the resulting datasets yielded marker analytes, illustrating the value of the broad-spectrum, molecular systems analysis approach to studying pharmacological or toxicological effects.

Keywords: liver toxicity; biomarker; entacapone; tolcapone.

INTRODUCTION

Drug-Induced Liver Injury and the Liver Toxicity Biomarker Study

Drug-induced toxicity is the primary reason for the withdrawal of drugs from the market and a frequent reason for the

failure of drug candidates in the clinical phases of drug development (Fung et al. 2001; Lasser et al. 2002; Shuster et al. 2005; Temple and Himmel 2002; Zimmerman 1999). From the analyses of available information on drug withdrawals and drug development terminations, it is apparent that the liver is the most frequent site of drug-induced toxicity (Fung et al. 2001; Kaplowitz 2001; Schuster et al. 2005).

The consequences of drug-induced liver injury (DILI) can be catastrophic. In a prospective study, transplant-free survival for patients with DILI, excluding that caused by acetaminophen overdose, occurred in only 25% of cases (Ostapowicz et al. 2002). Beyond the direct cost in human suffering and health care expenditure, the cost of DILI to pharmaceutical companies is substantial, with one company alone estimating that DILI cost it \$2 billion over a ten-year period (Rotman 2004). Clearly, there is a need to prevent DILI, either by preventing drugs that have a liability for causing liver injury from entering the marketplace or by close monitoring of patients who take drugs that carry such a liability, particularly if the drug is unique in its potential for therapeutic benefit.

DILI that arises in clinical trials, but has not been seen in preclinical studies, can have an incidence in trial subjects or patients of one in ten to one in several thousand. DILI that is first recognized only when a drug is on the market generally has an incidence in the target patient population that is too low to be detected in common phase III clinical trials involving hundreds to single-digit thousands of patients. This type of

Address correspondence to: Robert Nicholas McBurney, 610 N. Lincoln St., Waltham, MA 02451; e-mail: rmcburney@bg-medicine.com.

Abbreviations: ALT, alanine aminotransferase; ANOVA, analysis of variance; AST, aspartate aminotransferase; CRADA, cooperative research and development agreement; CV, coefficient of variation; DTT, dithiothreitol; DFTMP, 1,1-difluoro-1-trimethylsilyl methyl phosphonic acid; DILI, drug-induced liver injury; DNA, deoxyribonucleic acid; EH, high-dose, entacapone-treated cohort; EL, low-dose, entacapone-treated cohort; EM, medium-dose, entacapone-treated cohort; FDA, U.S. Food and Drug Administration; FT-MS, fourier transform mass spectrometer or spectrometry; FWHM, full width at half maximum; HPLC, high-performance liquid chromatography; LC/MS, liquid chromatography coupled to mass spectrometry; LTBS, The Liver Toxicity Biomarker Study; MALDI, matrix-assisted laser desorption ionization; MIAME, Minimal Information About a Microarray Experiment; MS, mass spectrometer or mass spectrometry; MS/MS, tandem mass spectrometry; NMR, nuclear magnetic resonance; PBS, phosphate-buffered saline; PCA, principal components analysis; QC, quality control; RIN, RNA integrity number; RNA, ribonucleic acid; RNase, ribonuclease; RT, retention time; SCX, strong cation exchange; TEAB, tetraethylammonium bicarbonate; TFA, trifluoroacetic acid; TCEP, tris(2-carboxyethyl)phosphine; TH, high-dose, tolcapone-treated cohort; TL, low-dose, tolcapone-treated cohort; TM, medium-dose, tolcapone-treated cohort; ToF, time of flight; V, vehicle-treated cohort.

liver toxicity, which has an incidence of one in ten thousand to one in one hundred thousand, is referred to as “idiosyncratic,” indicating the key role of individual susceptibility in the generation of the liver injury. It is not clear whether these two different DILI situations are distinct in nature or simply represent a continuum of susceptibilities involving the same general biochemical mechanisms.

Irrespective of the frequency of DILI in patients taking a certain drug, the fact that it occurs at all at therapeutic doses, even though no indications of liver toxicity were detected in high-dose toxicity studies performed in at least two animal species, is sufficient reason to reconsider the current approach to preclinical drug safety testing in the area of liver toxicity. In rethinking the current practice of determining a drug candidate’s potential for causing liver toxicity, the following hypothesis arises.

Despite the absence of conventional indicators of liver toxicity in preclinical studies, there exist biochemical signals (molecular biomarkers) in liver or body fluids that can be used to distinguish between a drug candidate that has the potential to cause DILI in susceptible patients and drugs that do not have this potential.

The Liver Toxicity Biomarker Study (LTBS) is an innovative approach to addressing the dilemma of unanticipated clinical drug-induced liver injury and its negative impact on individuals, the health care system, and pharmaceutical companies. The overall goal of the LTBS is to test the above hypothesis and, if the hypothesis is not refuted, to discover candidate biomarkers that could, following validation, be incorporated into the drug development process.

The LTBS was conceived in late 2005 as a collaborative research effort in molecular systems toxicology between the US Food and Drug Administration’s (FDA’s) National Center for Toxicological Research (NCTR) and BG Medicine, Inc., and is being carried out under a Cooperative Research and Development Agreement (CRADA). The research program is supported financially and with scientific expertise by an international group of seven pharmaceutical companies (Mitsubishi Chemical Holdings Corporation, Eisai Co. Ltd., Daiichi Sankyo Co. Ltd., UCB Pharma, Orion Pharma, Johnson and Johnson, Inc. and Pfizer, Inc.) and is supported with access to certain technologies by Applied Biosystems, Inc., Affymetrix, Inc. and TIBCO Software, Inc. Oversight of the LTBS is provided by a scientific advisory committee, chaired by Dr. Paul Watkins of the University of North Carolina, which includes Dr. Neil Kaplowitz, University of Southern California; and Dr. John Senior, FDA, as well as representatives from the supporting companies, NCTR, and BG Medicine.

The LTBS does not address the nature of an individual patient’s susceptibility for DILI (Kaplowitz 2005) but focuses on the biochemical effects of drugs that can interact with such susceptibility to cause DILI. Individual patient susceptibility is being addressed in clinical studies, such as those currently being conducted by the Drug-Induced Liver Injury Network (DILIN, <http://diln.dcri.duke.edu/index.html>).

The LTBS is based on the fundamental assumption that a specific type of DILI, such as hepatocellular necrosis, is likely

to result from one of a small number of biochemical mechanisms. Therefore, drugs of different chemical structures and different primary mechanisms of therapeutic benefit, but which cause that specific type of liver injury in certain susceptible patients, are likely to share some common biochemical effects, most probably not related to their primary mechanism of action. By comparing, in preclinical studies, the “off-target” biochemical effects of a number of compounds that are known to cause clinical DILI, it should be possible to discover preclinical predictive biomarkers for DILI.

The key feature of the research strategy for the LTBS is the comprehensive molecular systems analysis of the *in vivo* effects in rats of related compound pairs, one a “Clean Compound” and the other a “Toxic Compound” (see below for definitions), to discover the “off-target” biochemical response differences between the two compounds in each pair. The “on-target” biochemical response should be similar between the two compounds and, therefore, should be revealed by the comparison of the drug-effect biomarker sets for each compound in a compound pair.

Clean Compound: An approved drug that exhibited no signs of liver toxicity in preclinical studies, in clinical trials, or on the market.

Toxic Compound: A drug candidate, withdrawn previously marketed drug, or marketed drug with a warning label, of similar chemical structure and identical primary target activity to its corresponding Clean Compound, which exhibited no signs of liver toxicity problems in preclinical studies but which, at a certain dose in any phase I, II, or III clinical trial or on the market caused the appearance of clinical chemistry indicators or other symptoms consistent with hepatocellular injury in a proportion of trial subjects or patients sufficient to trigger a decision to terminate the development of the drug candidate, to withdraw the drug from the market, or to include a warning on the drug label.

The LTBS is addressing its objectives through a molecular systems analysis of liver tissue, blood plasma, and urine samples derived from an experimental paradigm that is based on a standard, three-dosage-level, twenty-eight-day rat toxicity study with animal sacrifice on day 29. Additional features of the LTBS are: an early sacrifice group for each compound at one dosage level (day 4 sacrifice after three days of dosing) and twenty-four-hour urine collections (days 0-4, and day 27 in the twenty-eight-day dosing study).

Five compound pairs will eventually be studied in this experimental paradigm. A differential marker set, representing the comprehensive molecular differences between the tissue and/or body fluid biochemical response of rats to each member of a compound pair, will be determined. Five such differential marker sets will be generated in the LTBS. If no molecular components or predicted biochemical pathways are found to be common among at least some of the differential marker sets, the hypothesis will be considered refuted. If common molecular components

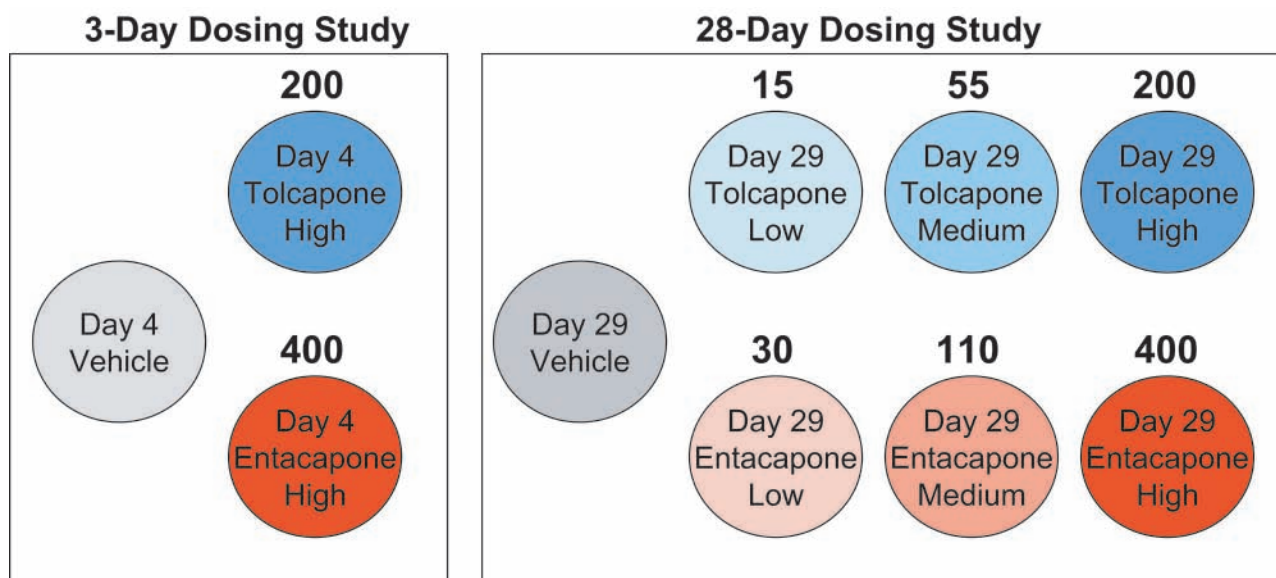


FIGURE 1.—A schematic of the experimental design of the three-day dosing and twenty-eight-day dosing parts of the LTBS. The colored circles represent the day of sacrifice samples that were generated for gene expression analysis, proteomic analyses, and metabolomic analyses. The mg/kg/day dosing for each drug for the study cohort represented by each colored circle is shown in bold above the colored circle. The colors of the circles for the drug-treated cohorts correspond to the color scheme used for the box-plots in Figure 3.

or common predicted biochemical pathways are found among the differential marker sets, such molecular components or biochemical pathways will be designated putative liver toxicity predictive biomarkers. Any putative liver toxicity predictive biomarker will represent a new hypothesis that will require testing and validation in subsequent experiments that are currently outside the scope of the LTBS. To be useful practically, the sensitivity and specificity of a liver toxicity predictive biomarker must be determined through additional studies involving sufficient numbers of compounds to generate robust statistical results.

Phases I and II of the LTBS

The LTBS is being conducted in two phases. In phase I, which commenced in early 2007 and is close to completion, a single pair of compounds has been studied in order to provide an opportunity for the scientific advisory committee and supporting pharmaceutical companies to review its results and operational aspects prior to committing to the design and funding for the remainder of the LTBS. Phase II will be initiated by the scientific advisory committee after review of the phase I results and will encompass the remaining four compound pairs.

Phase I Compound Pair

Twelve compound pairs were considered by the scientific advisory committee as candidates for the phase I compound pair. The catechol-O-methyl transferase inhibitors entacapone (Clean Compound) and tolcapone (Toxic Compound), indicated for the adjunctive treatment of Parkinson's disease (Bonifati and Meco 1999; Davis 1998), were selected as the

phase I compounds. Both compounds were reported to be free of liver toxicity in preclinical studies (Entacapone Product Monograph 1999; Tasmar Product Monograph 1997); however, clinical trials of tolcapone demonstrated dose-related increases in transaminases, and postmarketing studies revealed four cases of acute hepatotoxicity with three fatalities during the first 60,000 patient exposures (Assal et al. 1998; Olanow 2000; Olanow and Watkins 2007; Watkins 2000). Tolcapone was withdrawn from the market but has been reintroduced with a warning label. No cases of life-threatening hepatotoxicity have been reported for patients taking entacapone, although elevations of liver enzymes probably linked to entacapone exposure have been reported in two patients (Fisher et al. 2002).

Contents of this Article

This article is the first report arising from phase I of the LTBS and is intended to provide the rationale for the entire study, details on the study design, and some preliminary findings from phase I of the study.

MATERIALS AND METHODS

The overall design of phase I of the LTBS is shown in Figure 1. The overall workflow for the analysis of liver, plasma, and urine samples is shown in Figure 2.

Drugs

Entacapone was provided by Orion (Espoo, Finland) as a single lot. Tolcapone was purchased from US Pharmacopeia (Rockville, MD, USA) as a single lot.

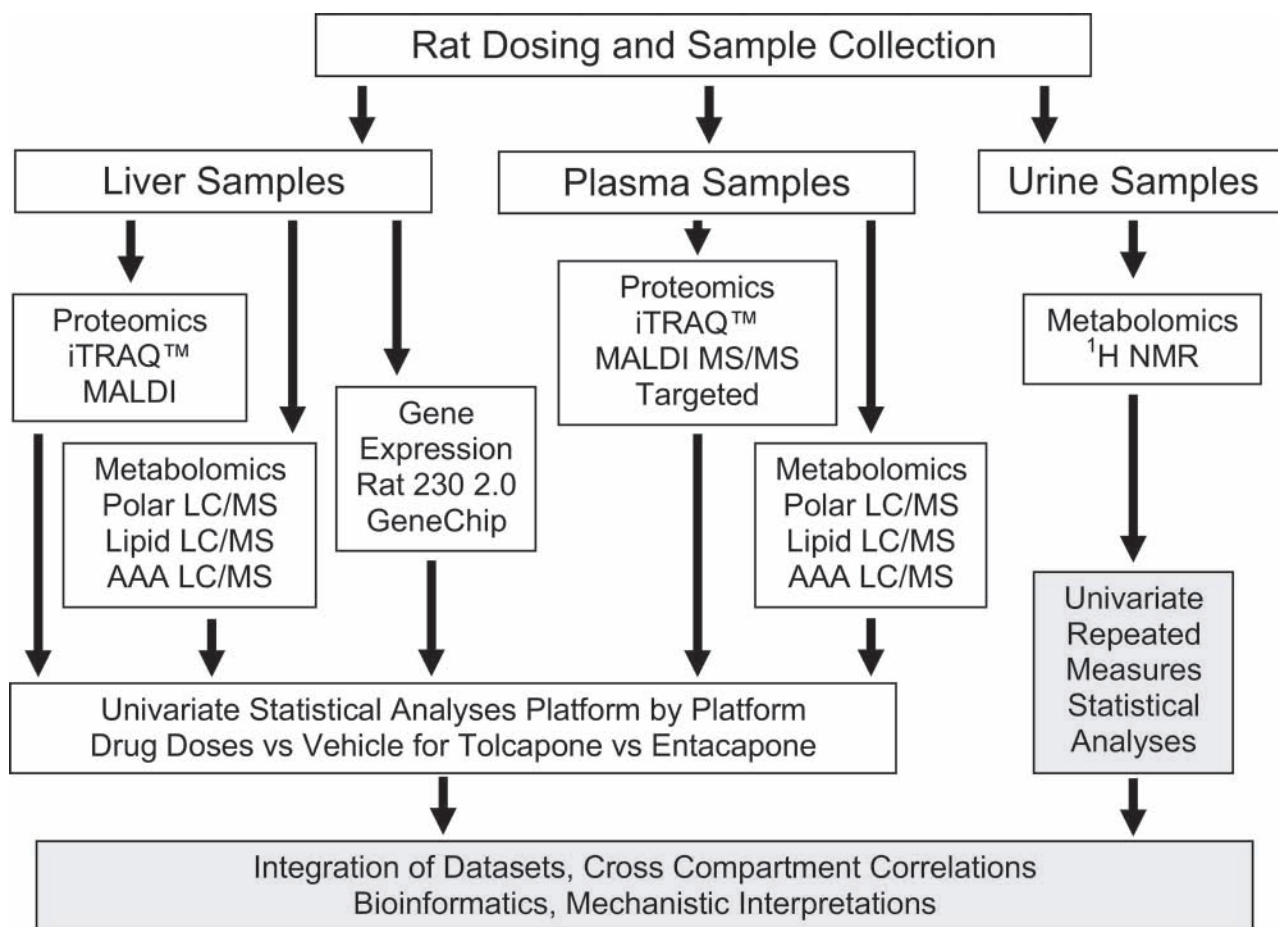


FIGURE 2.—A schematic that illustrates the overall workflow of the LTBS beginning with the rat dosing and sample collection stage and ending with the integration of datasets and mechanistic interpretations of the results of the statistical analysis. The methods employed in each of the bioanalytical platforms applied to liver, plasma, or urine samples are described in the Materials and Methods section.

Animal Study Design, Conduct, and Sample Collection

Each treatment cohort contained twelve male and twelve female Sprague-Dawley rats. Power calculations were performed to set the single-sex number of animals in each cohort. The goal of the power calculations was to have approximately 80% power to detect mean fold changes between pairs of cohorts of as little as 1.3.

Although both male and female rats were included in the dosing and sample collection component of the study, the preliminary results reported here are for the male rats only. The gross and microscopic evaluation of the tissue samples and the clinical chemistry measurement on the plasma samples were undertaken on both sexes. Only the samples from the male rats have been subjected to the bioanalytical profiling at this time.

Dose Levels

The dose levels for the study were based on available efficacy and safety information for entacapone and tolcapone (Haasio et al. 2001; Haasio et al. 2002; Hoffman-LaRoche, Inc.

1998; Napolitano et al. 2003; Olanow 2000; Orion Corporation 1999; Watkins 2000; Tornwall and Mannisto 1993; <http://chem.sis.nlm.nih.gov/chemidplus/direct.jsp/regno=130929-57-6|134308-13-7> see also Supplemental Information). The dosing range combined consideration of efficacious dosing in rat and man (lowest doses) and highest doses tolerated without conventional indications of toxicity, with the middle doses set at an approximate geometric mean.

Twenty-eight-day Dosing Study

Treatment began when the animals were six weeks of age. Entacapone was administered daily by gavage at doses of 30, 110, and 400 mg per kg body weight. Tolcapone was administered daily by gavage at doses of 15, 55, and 200 mg per kg body weight. Both drugs were administered in 0.5% methyl cellulose as the vehicle. Control rats were treated with the vehicle.

Two days before the initiation of dosing, the rats were transferred to metabolism cages (day -2) and allowed to acclimatize for forty-eight hours. A twenty-four-hour control urine sample was collected (beginning on day 0) on ice in 50-mL

polypropylene tubes containing 1 mL of 1% sodium azide. The urine samples were immediately stored at -60°C for subsequent metabolomic analyses. Dosing was initiated on day 1 and continued for twenty-eight consecutive days. Body weights were obtained daily to determine the appropriate dosing volume. Urine samples were collected for twenty-four hours after each of the first four doses. On dose day 25 the animals were placed once again in metabolism cages and allowed to acclimatize before collecting a twenty-four-hour urine sample on day 27. The rats were euthanized by exposure to carbon dioxide one day after the last dose.

Blood was collected by cardiac puncture, placed in tubes containing lithium heparin, and plasma was prepared, aliquoted into 500- μL portions, and frozen at -60°C . The livers were weighed as soon as possible after dissection. The median lobe of the livers from all animals was processed for histopathological examination. The remaining two lobes of the livers were immediately frozen in liquid nitrogen and stored at -60°C .

Three-day Dosing Study

This study was conducted in a manner identical to the twenty-eight-day study, with the following modifications. The treatment cohorts (twelve of each sex) consisted of entacapone at a dose of 400 mg per kg body weight, tolcapone at a dose of 200 mg per kg body weight, and a vehicle control. Dosing was conducted for three dose days, and the animals were euthanized one day after the last dose.

Pathology

The livers (median lobe) of all animals were examined grossly, removed, and preserved in 10% neutral buffered formalin. The livers were trimmed, processed, and embedded in infiltrating media (Formula R, Surgipath, Richmond, IL, USA), sectioned at approximately 5 microns, and stained with hematoxylin and eosin for subsequent microscopic evaluation. When applicable, the severity of lesions was graded as 1 (minimal), 2 (mild), 3 (moderate), or 4 (marked).

Plasma Clinical Chemistry Measurements

The following measurements were made on plasma collected at the scheduled terminal sacrifice using a Cobra Mira Plus analyzer (Roche Diagnostic Instruments, Indianapolis IN, USA): alanine aminotransferase (ALT), alkaline phosphatase, total bile acids, creatinine, blood urea nitrogen, aspartate aminotransferase (AST), sorbitol dehydrogenase, albumin, total protein, total bilirubin, lactate dehydrogenase, 5'-nucleotidase, glutamate dehydrogenase, and γ -glutamyl transferase.

Analysis of Liver Samples

Sample Preparation: Whole pieces of frozen liver lobes (weight approximately 5 g) were crushed in a CryoPrep bag (Covaris,

Woburn, MA, USA) on dry ice with a mallet to generate a fine powder. The pulverized tissue powder was aliquoted into one 100-mg portion for transcript analyses and two 40-mg portions for proteomic and metabolomic analyses, using RNase free cryovials. Small amounts from each pulverized primary sample were combined to create a reference sample pool (or QC sample pool) for the analytical platforms. The remainder was kept in the bags and stored at -80°C .

For both liver and plasma sample analyses on all analytical platform, sample randomization schemes were created to distribute the primary samples and the QC pool samples randomly throughout the analytical platform run order and among analytical batches (see, for example, van der Greef et al. 2007).

Gene Transcript Analysis: RNA was isolated from the frozen liver powder aliquots in batches of twelve samples at a time. The aliquots for RNA isolation were randomized with respect to treatment, dose, time, and necropsy date to reduce systematic bias. Prior to labeling of the RNA, the RNA samples were randomized with respect to treatment, dose, time, necropsy date, and RNA isolation date.

Qiagen RNeasy mini kits (Qiagen, Chatsworth, CA, USA) were used for RNA extraction following the standard Qiagen protocol. The concentration of the final RNA solution was determined by absorption at 260 nm. The RNA solution was stored at -80°C until use. The quality of extracted RNA was evaluated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). All RNA samples had an RNA Integrity Number (RIN) greater than 8.0.

Affymetrix GeneChip Rat Genome 230 v2.0 arrays (Affymetrix, Inc., Santa Clara, CA USA) were used for DNA microarray analysis. The built-in quality control was used to access the labeling and hybridization performance of each DNA microarray, and all were within recommended ranges. Protocols from Affymetrix were followed to process these microarray chips. Twelve GeneChips were processed in a batch. Each batch included eleven primary samples and one aliquot from the QC sample pool.

The Affymetrix one-cycle amplification protocol was followed for cRNA amplification and biotin labeling. The Affymetrix GeneChip Instrument System, which includes work station, fluidics station, hybridization oven, and scanner, was used for the microarray process (hybridization, washing, staining, and scanning). The resulting images were analyzed by GCOS software (Affymetrix) to generate .cel files, which contain the raw intensity values of each probe.

The data and experimental annotation were stored in the NCTR ArrayTrack database (Tong et al. 2003) that is Minimum Information About a Microarray Experiment (MIAME) compliant and a component of the NCTR Toxicoinformatics Integrated System (TIS) (Tong et al. 2004). PLIER, which is one of the multiple array-based normalization methods in the Affymetrix Expression Console Software, was then used to convert the probe-level intensity data into normalized probeset-level intensity data.

Proteomic Analysis

Protein Extraction and iTRAQ Labeling: Frozen tissue aliquots (40 mg) were immersed in 400 μ L homogenization buffer containing 6 M guanidinium hydrochloride, 1% Triton-X 100, 50 mM dithiothreitol (DTT), 50 mM tetraethylammonium bicarbonate (TEAB), and "Complete" Protease inhibitor cocktail (Roche Diagnostics, Mannheim, Germany—1/2 tablet dissolved in 50 mL buffer). Sample tubes were loaded into a Covaris E100 Ultrasonic Tissue Homogenizer (Covaris) and irradiated in four cycles, fifteen seconds each. The homogenates were briefly vortexed and centrifuged for fifteen minutes at 10,000rpm. Twenty μ L of the homogenate was aspirated for further processing. One μ L of 1M DTT was used to complete protein reduction (at 70°C, one hour), and cysteine alkylation was completed by adding 12 μ L 0.5 M iodoacetamide (room temperature, one hour). Unreacted iodoacetamide was quenched with 2.6 μ L of 1 M DTT. The protein sample was diluted with 250 μ L 15:10:75 formic acid–acetonitrile–water solvent and injected onto a Poros R1 reversed-phase column on a Vision Workstation LC System (Applied Biosystems, Framingham, MA, USA). After washing with 30% acetonitrile, the protein fraction was eluted with 30:70 formic acid–isopropanol solvent, and the column was then washed with 70:30 formic acid–isopropanol to remove any retained material. The protein fraction was dried and resuspended in 40 μ L buffer containing 2 M urea, 1 M TEAB, and 1% *n*-octylglucoside. Trypsin digestion was carried out by adding 5 μ g Trypsin Gold (Promega, Madison, WI, USA) twice (at the start and at the two-hour mark) to the sample. Digestion completed after four hours. Half of the digestion mix was labeled with the 8-plex iTRAQ reagent (Applied Biosystems) following the instructions of the manufacturer (for the labeling step only). Excess reagent was quenched by adding 2 μ L of 1 M ammonium bicarbonate (one hour). The eight samples constituting the 8-plex iTRAQ mix were combined and dried by speed vacuum. In each iTRAQ mix, two lanes—tagged with the *m/z*-113 and -117 labels—were reserved for QC pool samples for relative quantification.

Peptide Chromatography

The peptide mixes were initially reconstituted in 60 μ L of 8 M urea and then with a 500 μ L aliquot of the resuspension buffer (1% trifluoroacetic acid [TFA] and 5% acetonitrile). First the sample (500 μ L) was loaded onto a Poros R2 reversed-phase column (Applied Biosystems), washed with the loading buffer (5% acetonitrile and 0.1% TFA), and eluted onto a poly-sulfoethyl strong cation exchange (SCX) column (PolyLC, Columbia, MD, USA) with 35% acetonitrile and 0.1% TFA in an Agilent Technologies 1200 Chemstation. Buffer A (10 mM KH_2PO_4 in 25% acetonitrile buffered to pH 3.5 with 1 mM HCl) and Buffer B (10 mM KH_2PO_4 with 1 M KCl in 25% acetonitrile buffered to pH 3.5 with 1mM HCl) were used to elute the peptides from the SCX column using a gradient of 0%–10% B for 8.7 minutes, 10%–35% B for 16.0 minutes, 35%–100% B for 1.9 minutes, and finally 100% B for 26.6 minutes. Six fractions were collected from the SCX column.

These six SCX fractions were dried and resuspended in HPLC loading buffer (50–80 μ L 5% acetonitrile, 0.1% TFA). HPLC fractions were spotted onto matrix assisted laser desorption ionization (MALDI) plates.

Mass Spectrometry and Data Processing

The mass spectrometric analysis was completed on an Applied Biosystems/MDS Analytical Technologies 4800 MALDI ToF/ToF Mass Spectrometer. MS spectra were acquired in reflector mode (500–5000 *m/z* mass range, 1000 shots per spectrum, two-point internal calibration) on all six fractions (two MALDI plates) prior to precursor selection for MS/MS. Precursor selection was completed using a proprietary software program that enables selection of precursors in both LC dimensions.

The MS/MS spectra were searched against the International Protein Index rat protein database (version 3.36, European Bioinformatic Institute, www.ebi.ac.uk), using the Mascot Search Algorithm Version 2.0 (Matrix Biosciences, Boston, MA, USA). A mass tolerance of 50 ppm and 0.4 Da was used for precursor and fragment ions, respectively. Two missed cleavage sites were allowed. The following variable modifications were used: iTRAQ 8Plex (N-terminus and K), deamidation (N only), oxidation (M), Pyro Cmc N-term CamC), and Pyro-Glu (N-term Q).

Lipid LC/MS Metabolomic Analysis

Protein precipitation and lipid extraction were carried out using methanol:water (80:20) and further extracted using dichloromethane:isopropanol (10:90). The lipid extracts were analyzed by LC/MS using a Dionex Ultimate 3000 HPLC system (Dionex, Sunnyvale, CA, USA) coupled to a Q-ToF instrument (QSTAR, Applied Biosystems, Foster City, CA, USA). The lipids were separated using a Gemini C6-Phenyl 3 μ (2 mm x 150 mm) column (Phenomenex, Torrance, CA, USA) working at 55°C. The data processing for lipid profiling was carried out using peak picking and peak alignment programs developed at BG Medicine. A unique feature (*m/z* with RT) was selected for each analyte for the downstream statistical analysis. Each unique feature of a lipid was identified by combining the information of its precursor ion *m/z* and RT, accurate mass measurement obtained from FT-MS, and its MS/MS fragmentation pattern.

Amino Acid Analysis (AAA)

Protein precipitation and amino acid extraction were carried out using methanol:water (80:20). Amino acids in the extracts were labeled with the 115-iTRAQ reagent. The forty-two targeted amino acids were labeled with the 114-iTRAQ reagent and spiked into the primary and QC pool samples as internal standards. The iTRAQ reagent-labeled amino acids were separated using an AAA C18 column (4.6 mm x 150 mm, Applied

Biosystems) working at 50°C. The LC/MS multiple reaction monitoring data were generated using a Dionex Ultimate 3000 HPLC system (Dionex) coupled to a quadrupole linear ion-trap mass spectrometer (4000 QTRAP, Applied Biosystems). The absolute quantification of each amino acid was calculated from the ratio of intensities measured for the 115-labeled sample and the 114-labeled internal standard, which was added at a known concentration.

Polar LC/MS Metabolomics Analysis

For the liver samples, the polar metabolite extraction procedure was the same as for the AAA platform. The polar compounds in the extract were butylated prior to the LC/MS analysis. The butyl-ester derivatives were separated using a Varian/Chrompack Inertsil 5 µm ODS-3 (3 mm x 100 mm) column (Varian, Walnut Creek, CA, USA) followed by on-line MS detection using an LTQ Orbitrap (Thermo Fisher, San Jose, CA, USA).

Analysis of Plasma Samples

Sample Preparation

Individual rat plasma samples collected into Li-heparin-coated tubes were aliquoted for the different analytical platforms. Similarly to the tissue process, portions of each of the primary samples were combined to generate a QC sample pool that was re-aliquoted according to the needs of the individual analytical platforms.

Proteomic Analysis

Protein Extraction and iTRAQ Labeling: This method is based on the method of Lynch (2005). 20 µL of each plasma sample was diluted to 80 µL with 1x PBS. 40 µL of tetrachloroethylene was added, and the mixture was vigorously vortexed for 3x5 seconds to form an emulsion. The emulsion was centrifuged at 4°C for ten minutes at 14,000rpm. The top, aqueous layer was carefully removed. The sample was diluted to 600 µL with 1x PBS and injected onto an IgY R7 column (Beckman Coulter, Fullerton, CA, USA) containing IgY antibodies against seven abundant rodent proteins (albumin, IgG, IgM, transferrin, fibrinogen, haptoglobin and alpha-antitrypsin). The material not retained was captured on a Poros R1 column (Applied Biosystems), which was washed and eluted with 95% acetonitrile. The collected protein fraction was dried in a Speedvac. The RP column was then washed with 70:30 formic acid-isopropanol solvent and the IgY column was regenerated after eluting the abundant proteins with 100 mM glycine HCl buffer (pH 2.5). In between runs, the IgY column was re-equilibrated with 100 mM HEPES buffer (pH 8). The dried protein was reconstituted in 20 µL resuspension buffer (1 M TEAB, 2 M urea, 1% *n*-octyl-glucoside), and 2 µL of 50 mM tris(2-carboxyethyl)phosphine (TCEP) was added. The protein reduction was carried out for one hour at 70°C. One µL 84 mM solution of iodoacetamide was added, and the alkylation was completed at room temperature for thirty minutes. Five µL

Trypsin Gold was added at 1 mg/mL concentration also containing 4 mM N-acetyl cysteine to quench excess iodoacetamide and to digest the proteins overnight.

iTRAQ labeling with the 8-plex reagent was carried out as described above for the liver sample proteomic analysis. Peptide chromatography, mass spectrometry, and data processing were carried out as described above for the liver sample proteomic analysis.

Multiplexed Protein Immunoassays

Rat plasma samples were analyzed with the RodentMAP Version2 multiplex immunoassay platform (Rules Based Medicine, Austin, TX, USA). The antigen panel consisted of fifty-nine proteins, which included proteins involved in inflammation, cytokines, and proteases. Bead-based immunoassays were developed and optimized at Rules Based Medicine and read out on a Luminex 200 system (Luminex, Austin, TX, USA). Aliquots (100 µL) of primary samples and QC samples were used for the analysis.

Immunoassay results were analyzed in terms of coefficients of variations (CVs) as observed from twenty-four replicate QC pool samples and of number of valid measurements, that is, measurements above the limit of detection. Any antigen for which the CV exceeded 20% or for which the number of measurements below the limit of detection exceeded 20% of the samples (twenty-four or more samples) was excluded from the statistical analysis. This filtering procedure resulted in the use of nineteen analytes in the statistical analysis. Failing the QC or limit of detection criteria typically occurred together, indicating that poor CVs are associated with inadequate sensitivity of the assay for the given sample type. The main reason for the relatively few analytes satisfying the QC criteria is in the nature of "generic rodent" assays in RodentMAP Version2. These assays were not optimized specifically for rat antigens, but rather to deliver a balanced performance on both rat and mouse plasma samples.

Lipid LC/MS Metabolomic Analysis

Protein precipitation and lipid extraction were carried out using dichloromethane:isopropanol (10:90). All other aspects of the plasma sample analysis were identical to that described above for the lipid LC/MS analysis of the liver samples.

Amino Acid Analysis

The amino acid analysis of plasma samples was carried out by a procedure identical to that described above for the amino acid analysis of the plasma samples.

Polar LC/MS Metabolomic Analysis

The Polar LC/MS metabolomic analysis of the plasma samples was carried out by a procedure identical to that described above for the polar LC/MS analysis of the liver samples.

Analysis of Urine Samples

Urine samples were thawed, and three 500 μL aliquots of the urine samples removed from each 50 mL conical tube and deposited into separate 1.5 mL Eppendorf tubes. Additional 100 μL aliquots of each sample were pooled for the QC pool sample.

Urine and QC samples were prepared for analysis by the addition of 250 μL of sodium phosphate buffer (pH 7.4) and 75 μL of a mixture of 10 mM 1,1-difluoro-1-trimethylsilyl methyl phosphonic acid (DFTMP) (Bridge Organics, Vicksburg, MI, USA) and 1 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS-d6) (Sigma, St. Louis, MO, USA) in deuterium oxide (D_2O , Cambridge Isotope Laboratories, Andover, MA, USA) to 500 μL of urine. The samples were centrifuged at 12,000 rpm for twelve minutes at 10°C, and 550 μL were transferred to 5 mm o.d. NMR tubes.

For quality assurance purposes, each day, a sensitivity standard was run first and the S/N recorded to evaluate performance before starting sample analysis. The QC sample was run every twenty samples throughout the NMR analysis. 2D NMR JRES and HSQC experiments were run overnight on the last urine sample evaluated each day.

NMR spectra were acquired on a Bruker Avance 600 MHz spectrometer (Bruker, Billerica, MA, USA) operating at 600.133 MHz for proton. NMR spectra were acquired as previously described (Schnackenberg et al. 2007). The analysis time for each sample was approximately ten minutes using the parameters described above and including the autoshimming routine. Only ten urine samples were placed in the sample carousel at any one time to reduce the amount of time that the samples were exposed to room temperature prior to NMR analysis.

Spectra for urine, QA, and QC samples were processed using ACD/Labs 1D NMR Manager Version 10 (ACD/Labs, Toronto, Ontario, Canada) as previously described (Schnackenberg et al. 2007). All spectra were autoreferenced to the DSS-d6 peak at 0.0 ppm. The table of integrals was exported as a text file for statistical analyses. Spectra were exported as JCAMP files for quantitative analysis.

The full width at half maximum (FWHM) of the DSS-d6 peak at 0.00 ppm was assessed for each spectrum, and if the FWHM of DSS-d6 was greater than 3 Hz, the spectrum was rejected and the sample reanalyzed.

The raw NMR data were processed using ACD/Labs 1D NMR Manager (ACD/Labs). JCAMP files for each spectrum were used for analysis of specific metabolites using the Chenomx Eclipse software (Chenomx, Edmonton, Alberta, Canada). The Chenomx NMR spectral database was used to determine the concentrations of specific metabolites in the spectra. The concentrations of the metabolites in each spectrum were normalized by the absolute spectral intensity of that spectrum. For integration, the total binned intensity across each spectrum was normalized to a nominal value of 100 in the ACD/Labs integration module.

Metabolites were identified and quantified using the Chenomx NMR Suite (Chenomx), which has a database of more than 250

metabolites. Identification will be verified by 2D HSQC and JRES data.

Statistical Methods for Biomarker Discovery

Principal Components Analysis (PCA) was applied to explore the natural groupings of, and outliers among, the data derived from each platform in the study. Score plots were used to visualize separation between animals randomized to different treatment cohorts and time points at which samples were obtained. The contributions of specific analytes on respective outliers were subsequently inspected by means of loading plots.

Univariate statistical analysis was performed to address the primary objective of the study, namely to discover absence versus presence, fold-change, and change-in-CV markers that are indicative of the toxic potential of the drug tolcapone. For each univariate analysis, derived vectors of p values corresponding to the list of all analytes measured in each platform individually and combined were adjusted for multiple comparisons using the method by Benjamini and Hochberg (1995). For clarity, we use the term “marker” to refer to the initial results of a statistical analysis and the term “biomarker” to describe a marker, or marker set, that has received additional evaluation to assess its candidacy as an indicator of a biological process (Biomarkers Definition Working Group 2001).

Absence versus Presence Markers

The univariate analysis consists of logistic regression models that evaluate the statistical significance of the association of each analyte’s percentage of missingness with treatments cohorts and the vehicle. Analyte values were transformed to binary variables, 0 indicating missing values and 1 otherwise, and the respective missingness was regressed on the different treatment cohorts and the vehicle and time points.

Fold-Change Markers

The primary univariate analysis is based on ANOVA models taking into account the different treatment cohorts and the vehicle, and time points. Each analyte (transformed to the natural logarithmic scale) measured by respective assays and platforms is considered an outcome in the ANOVA models. For the analysis of serum clinical chemistry and liver and plasma “-omics” analytes, the statistical significance of each of the pairwise comparisons was assessed via likelihood ratio tests. A p value corresponding to each of the comparisons was obtained for every analyte, with at least 50% of data present in all treatment cohorts.

Change-in-CV Markers

The univariate analysis to detect analytes showing significant differences in coefficient of variation (CV) between treatment cohorts and the vehicle is based on permutation tests (Pitman 1937). A p value corresponding to each pairwise group

comparison was obtained for every analyte, with at least 50% of data present in all treatment cohorts. The p value was based on 1000 treatment group label permutations.

Platform Performance Metrics

Except for iTRAQ proteomics, platform analytical variability per profiled analyte was determined by calculation of the coefficient of variation (CV) per profiled analyte based on the repeated measurements obtained on the QC pool samples. For iTRAQ proteomics, analytical variability was determined from an internal consistency metric calculated for proteins with multiple reporter peptides. For this estimate, a protein iTRAQ ratio is calculated as the median ratio from the set of peptides mapped to the same protein. Each of the individual peptide iTRAQ ratios in this set is then compared to the protein iTRAQ ratio; the reported CV is the coefficient of variation of the peptide iTRAQ ratios / protein iTRAQ ratio. This calculation is repeated for each sample, and the per analyte CV is the average across the study.

RESULTS AND DISCUSSION

The results presented here are intended to provide an overview of some preliminary findings from phase I of the LTBS. They include the evaluation of liver toxicity indicators, histopathology and clinical chemistry, the performance of the bioanalytical platforms (except the NMR urine analysis), and an initial overview discussion of marker findings. For a summary comparison of entacapone and tolcapone, please see the supplemental appendix to this article published electronically only at <http://tpx.sagepub.com/supplemental>.

Gross and Microscopic Pathology

Overall, the gross and microscopic evaluation of the livers obtained at sacrifice from both the twenty-eight-day dosing study and the three-day dosing study did not find any treatment-related lesions. This finding represents a key result for the study, namely the lack of conventional indicators of liver toxicity caused by tolcapone.

Consistent with the presence of glycogen within hepatocytes, minor hepatocellular cytoplasmic vacuolization was noted in the rat livers and was considered a result of the animals not being fasted prior to necropsy. The incidence and severity of these cytoplasmic vacuoles were similar in all treatment cohorts, indicating that they were unrelated to treatment of the rats with entacapone or tolcapone.

Plasma Clinical Chemistries

No statistically significant findings judged as consistent with liver toxicity being caused by any of the treatments were found. This finding also represents a key result for the study.

Inspection of the individual box-plots for the ALT and AST measurements across all treatment cohorts in the twenty-eight-day

dosing study revealed that the three highest measured values for ALT occurred in the high-dose, tolcapone-treated cohort (see Figure 3A, animal numbers 173, 83, and 75). Furthermore, in the high-dose, tolcapone-treated cohort, the same animals had the three highest AST values within that cohort (Figure 3B). Based on this intriguing observation, results for rats 173, 83, and 75 will receive attention in the analysis of the “-omics” datasets to identify further trends involving these rats that might signal low-grade hepatotoxicity associated with tolcapone exposure.

Overall Performance of Bioanalytical Profiling Platforms

An understanding of the reproducibility of an analytical platform is an important basis upon which to interpret the results obtained with the platform. The primary metric used here for evaluating platform performance is the distribution of the coefficient of variation (CV) for each measured analyte derived from multiple analyses of the QC pool samples.

Tables 1 and 2 present the distributions of CVs for all analytes measured in the QC samples created from the rat liver samples and rat plasma samples, respectively. A total of 32,895 analytes was measured in the liver QC samples by the five bioanalytical platforms applied to those samples (Table 1). The CVs were less than 20% for the majority of analytes measured in the liver QC samples (Table 1). A total of 678 analytes was measured in the plasma QC samples by the five bioanalytical platforms applied to those samples (Table 2). The CVs were less than 20% for the majority of analytes measured in the plasma QC samples (Table 2). For both liver and plasma analyses, the Polar LC/MS platform had the highest proportion of analytes, with CVs greater than 20%. The higher CVs obtained for analytes measured by this platform most probably reflect additional variability introduced by the chemical modification (butylation) step employed in the sample preparation prior to profiling on the LC/MS.

Preliminary Statistical Results for Comparisons of Tolcapone Effect and Entacapone Effect in Liver and Plasma from the Twenty-eight-day Dosing Study

The comprehensive molecular systems analysis of liver and plasma samples combined with statistical analyses of the resulting datasets has revealed many similarities and differences between the *in vivo* biochemical effects of the two drugs. All bioanalytical platforms applied to either liver or plasma samples contributed marker analytes, illustrating the value of the broad-spectrum, molecular systems analysis approach to studying pharmacological or toxicological effects.

For both the three-day dosing study and the twenty-eight-day dosing study, on a platform-by-platform basis, each dataset was subjected to ANOVA to determine the median fold changes and statistical significance of each pairwise comparison between the treatment cohorts. Analytes meeting a certain criterion of statistical significance in certain cohort comparisons, such as a p value cutoff, were declared to be “markers.”

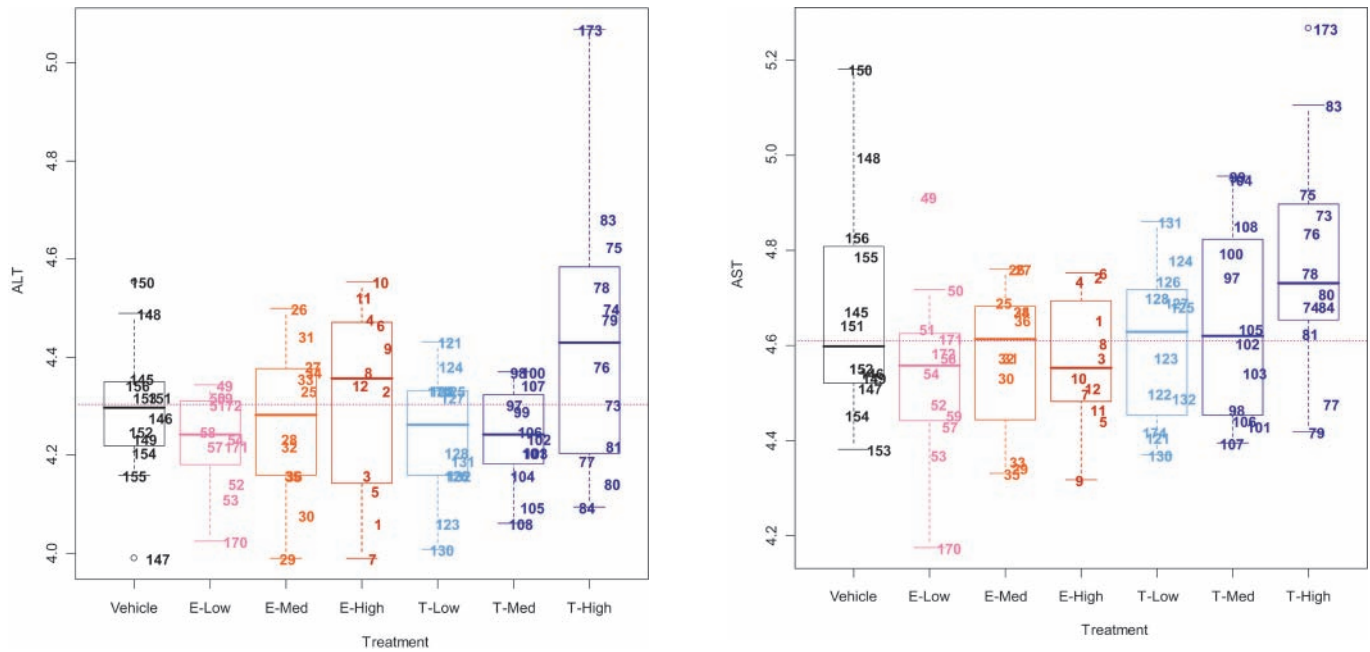


FIGURE 3.—A, Box-plots for the levels of alanine aminotransferase (ALT) measured in the plasma samples obtained at sacrifice (day 29) for the animals in all treatment cohorts of the twenty-eight-day dosing study. The ordinate is the natural logarithm of the ALT value. Note that animals 173, 83, and 75 in the high-dose, tolcapone-treated cohort have the highest measured values of ALT across all cohorts. B, Box-plots for the levels of aspartate aminotransferase (AST) measured in the plasma samples obtained at sacrifice (day 29) for the animals in all treatment cohorts of the twenty-eight-day dosing study. The ordinate is the natural logarithm of the AST value. Note that animals 173, 83, and 75 have the highest measured values of AST in the high-dose, tolcapone-treated cohort, but other high values of AST are found in the vehicle-treated cohort.

TABLE 1.—Coefficients of variation (CVs) for analytical platforms applied to the liver samples.

CV Range	mRNA analysis	Discovered proteins	Lipids	Amino acids	Polar analytes
0%–10%	11,957	630	108	34	103
>10%–20%	10,935	142	15	3	138
>20%–30%	4219	1	9	0	51
>30%–50%	2933	0	4	0	21
>50%	1054	0	1	0	6
NA	1	529	—	1	—
Total	31,099	1302	137	38	319

TABLE 2.—Coefficients of variation (CVs) for the analytical platforms applied to plasma samples.

CV Range	Discovered proteins	Targeted proteins	Lipids	Amino acids	Polar analytes
0%–10%	41	11	96	27	74
>10%–20%	100	7	6	3	68
>20%–30%	13	1	3	1	81
>30%–50%	5	0	0	1	7
>50%	0	0	2	0	2
NA	124	0	5	0	0
Total	283	19	112	32	232

For the twenty-eight-day dosing study with seven treatment cohorts, such an effort had the potential to yield a large number of different cohort comparisons, and markers of many different comparisons. Some of these markers would be of little or no interest in light of the overall objective of the study. We therefore focused on markers derived from just nine cohort comparisons of interest (TL-V, TM-V, TH-V, EL-V, EM-V, EH-V, TL-EL, TM-EM and TH-EH; see Abbreviations for definitions of cohort abbreviations).

To manage the results of the ANOVA in a fashion consistent with the overall study objective, four major classes of marker behavior were defined as follows: tolcapone-specific; entacapone-specific; common to both drugs; and divergent between both

drugs. The purpose of this marker classification approach was to consolidate separate lists of markers to elucidate similarities and differences between the biochemical pathways affected by the drug treatments. We found the process of selecting and classifying markers of the actions of tolcapone and entacapone to be considerably more challenging than anticipated during the study design stage. The following is an overview of the criteria used for marker classification and of the number of markers in each class on a platform-by-platform basis.

In a strict sense, tolcapone-specific markers should be statistically significant in the tolcapone-treated to vehicle-treated (T-V) and tolcapone-treated to entacapone-treated (T-E) cohort comparisons, but not statistically significant in entacapone-treated

TABLE 3.—Number of markers by bioanalytical platform and classification in the twenty-eight-day dosing study.

Platform	<i>p</i> -value cutoff	total	Marker class							
			E	e	T	t	C	c	D	U
Plasma clinical chemistry	.05	6	1	2	0	1	0	1	0	3
Liver gene transcripts	.002	1786	383	886	44	95	257	572	2	546
Liver proteomics	.02	192	19	69	9	34	13	26	0	78
Liver lipid LC/MS	.01	26	2	2	6	10	0	0	0	14
Liver AAA	.02	13	7	9	2	2	0	1	0	2
Liver polar LC/MS	.02	57	8	18	2	3	4	7	0	32
Plasma proteomics	.02	51	2	8	3	7	5	8	0	31
Plasma targeted proteins	.05	9	0	1	2	4	0	0	0	4
Plasma lipid LC/MS	.01	12	0	0	2	2	0	0	0	10
Plasma AAA	.02	11	6	7	1	1	0	0	0	3
Plasma polar LC/MS	.02	141	15	25	4	5	1	2	0	110

U signifies a marker of unusual behavior that is not readily interpretable.

to vehicle-treated (E-V) cohort comparisons. The number of such strict-sense, tolcapone-specific markers discovered by each bioanalytical platform for the twenty-eight-day dosing study is presented in Table 3 in the column labeled "T." In a liberal sense, a marker qualifying in the T-V but not in the E-V nor in the T-E cohort comparisons could be considered tolcapone specific, though this definition encompasses a wide range of marker behavior, from one extreme where the entacapone-treated cohorts are nearly indistinguishable from vehicle-treated cohorts (most desirable), to another extreme, where the entacapone-treated cohorts are nearly indistinguishable from tolcapone-treated cohorts (less desirable). Markers significant in T-V, E-V, and T-E cohort comparisons, where the magnitude of change in abundance of an analyte from its abundance level in the vehicle-treated cohort is greater in the tolcapone-treated cohort than in the entacapone-treated cohort, can be considered as liberal sense, tolcapone specific. Markers for the case where T-V and E-V are not significant but T-E was significant are also included in the liberal-sense, Tolcapone-specific class. The number of such liberal-sense, Tolcapone-specific markers discovered by each bioanalytical platform for the twenty-eight-day dosing study is presented in Table 3 in the column labeled "t."

Markers were classified as strict sense or liberal sense entacapone specific following logical principles identical to those described above for classifying markers as tolcapone-specific. The number of strict-sense, entacapone-specific markers discovered by each bioanalytical platform for the twenty-eight-day dosing study is presented in Table 3 in the column labeled "E," and the number of liberal-sense, entacapone-specific markers discovered by each bioanalytical platform for the twenty-eight-day dosing study is presented in Table 3 in the column labeled "e." Strict-sense Common markers encompass cases where the T-V and E-V cohort comparisons are statistically significant but the T-E comparison is not statistically significant (see Table 3 column labeled "C," for the numbers of such markers from the twenty-eight-day dosing study), and liberal-sense Common markers are those where T-V and E-V cohort comparisons are statistically significant and the T-E

cohort comparison is statistically significant but the direction of change of both T and E cohorts from the V cohort is identical (see Table 3 column labeled "c" for the number of such markers from the twenty-eight-day dosing study). Divergent markers are the case where T-V and E-V cohort comparisons are statistically significant but the T-V marker changes its abundance in a different direction relative to vehicle from the E-V marker. The number of Divergent markers discovered by each bioanalytical platform for the twenty-eight-day dosing study is presented in Table 3 in the column labeled "D."-

For rats dosed for twenty-eight days with high (H), medium (M), and low (L) doses, emphasis in this preliminary analysis was placed on the high and medium doses. Table 3 presents a compendium of markers revealed by the ANOVA applied to the datasets generated from the application of the bioanalytical platforms to the terminal liver and plasma samples from the twenty-eight-day dosing study. The table arranges these markers into the classifications described above. Markers were revealed in all platform datasets. There were only two Divergent markers found in the twenty-eight-day dosing study, and both of these were found with the Liver Transcript platform.

The number of markers in each classification depends on the *p* value cutoff criterion used to declare an analyte a marker for a particular comparison of cohorts. For the results shown in Table 3, different *p* values have been used for different platforms based on the number of analytes measured in that platform and a consideration of false discovery rates (Benjamini and Hochberg 1995), particularly for a dataset such as the gene transcript dataset that contains over 30,000 analytes.

In Table 3, it is apparent that seven out of eleven platforms generated higher numbers of entacapone-specific markers than tolcapone-specific markers. One possible explanation for this result is that for each dose cohort, the dosing for entacapone on a molar basis was approximately twice the dosing for tolcapone. This molar difference between entacapone and tolcapone at the different dosing levels is a consequence of the design aspect of the study, which attempted to set the doses at equi-efficacious levels so that the pharmacological effects of each drug on its primary target, catechol-O-methyltransferase, would be equivalent.

In light of this possible explanation, the greater number of tolcapone-specific markers than entacapone-specific markers in four out of eleven platforms (highlighted with a yellow background) might be considered particularly worthy of attention during the ongoing data analysis and interpretation aspects of phase I and when the results of phase I are compared to the results of phase II. The liver lipid LC/MS results are especially intriguing, given the substantial role of certain lipid species in inflammation and membrane degeneration/regeneration.

The preliminary results for phase I of the LTBS contain differential marker sets composed of both liver and plasma molecules that were derived from comparisons of the effects on rats of tolcapone and entacapone in a three-day dosing study or a twenty-eight-day dosing study. The existence of such differential marker sets, under conditions where no overt toxicity is observed in tolcapone-treated rats, is a prerequisite condition for the future discovery of a liver toxicity predictive biomarker from the LTBS. Specific statistical results and candidate biomarker identities will be reported in subsequent publications.

Implementation of Phase I of the LTBS

Phase I of the LTBS demonstrated that despite the complexity of the study in both scientific and management aspects, the overall approach to address the fundamental hypothesis is feasible. A public-private partnership of this type involving a public regulatory agency and industrial partners and supporters is able to undertake a project that would likely be too resource intensive for any one of the contributors alone to attempt. Furthermore, the collaborative aspects of the study design and critique of results enhance the study's prospects of success.

Completion of Phase I of the LTBS and Initiation of Phase II

Two activities are currently proceeding in parallel: the completion of the statistical analysis and biological interpretation of the results of phase I of the LTBS, and planning for phase II of the LTBS.

With regard to the overall biological interpretation of the results of phase I, a variety of data integration and data mining approaches are being used at the NCTR and at BG Medicine, including ArrayTrack (Tong et al. 2003; Tong et al. 2004), Ingenuity Pathways Analysis (Ingenuity Systems, Redwood City, CA, USA), and Correlation Network Analysis (Adourian et al. 2008). One focus of this analysis is to determine whether there is evidence of mitochondrial uncoupling in the integrated molecular profiling dataset for the effects of tolcapone on rat liver, as has been suggested by the previous comparisons of the actions of entacapone and tolcapone (Haasio et al. 2001; Haasio et al. 2002).

Phase I of the LTBS is just the first step on the path to testing the overall hypothesis of the study and, if possible, to discovering liver toxicity predictive biomarkers that can be employed in pre-clinical drug development studies in rats to evaluate the a drug candidate's potential to cause liver toxicity in patients despite the absence of conventional signs of liver toxicity in those preclinical studies. Four additional compound pairs will be studied in phase

II of the LTBS to provide a total of 5 differential marker sets. When comparisons are made between the analytes of these 5 differential marker sets, it will be possible to determine whether there are analytes in common that would represent hypothetical liver toxicity predictive biomarkers.

ACKNOWLEDGMENTS

We thank Ralph Patton, Toxicologic Pathology Associates, NCTR, for conducting the clinical chemistry measurements and William M. Witt, Toxicologic Pathology Associates, NCTR, for performing the histopathological analyses. We also thank Dr. Paul Watkins for chairing the scientific advisory committee and for his comments on the manuscript and Drs. Neil Kaplowitz and John Senior for their advice and counsel. We are grateful to the following companies for financial support: Mitsubishi Chemical Holdings Corporation, Eisai Co. Ltd., Daiichi Sankyo Co. Ltd., UCB Pharma, Orion Pharma, Johnson and Johnson, Inc., and Pfizer, Inc. In addition, we thank the following companies for technology support: Applied Biosystems, Inc., Affymetrix, Inc., and TIBCO Software, Inc. The views expressed in this paper do not necessarily represent those of the U.S. Food and Drug Administration.

REFERENCES

- Adourian, A., Jennings, E., Balsubramanian, R., Hines, W.M., Damian, D., Plasterer, T.N., Clish, C.B., Stroobant, P., McBurney, R., Verheij, E.R., Bobeldijk, I., van der Greef, J., Lindberg, J., Kenne, K., Andersson, U., Hellmond, H., Nilsson, K., Salter, H. and Schuppe-Koistinen, I. (2008) Correlation network analysis for data integration and biomarker selection. *Mol BioSyst* **4**, 249–59.
- Assal, F., Spahr, L., Hadengue, A., Rubbia-Brandt, L., Burkhard, P. R. (1998) Tolcapone and fulminant hepatitis. *Lancet* **352**, 958.
- Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* **57**, 289–300.
- Biomarkers Definitions Working Group (2001). Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* **69**, 89–95.
- Bonifati, V. and Meco, G. (1999) New, selective catechol-O-methyltransferase inhibitors as therapeutic agents in Parkinson's disease. *Pharmacol Ther* **81**, 1–36.
- Comtan (Entacapone) Tablets NDA 20-796 (December 30, 1999). Available at: www.fda.gov/cder/foi/nda/99/20796_Comtan.htm.
- Davis, T.L. (1998) Catechol-O-methyltransferase inhibitors in Parkinson's disease: guidelines for effective use. *CNS Drugs* **10**, 239–46.
- Fisher, A., Croft-Baker, J., Davis, D., Purcell, P. and McLean, A.J. (2002) Entacapone-induced hepatotoxicity and hepatic dysfunction. *Mov Disord* **17**, 1362–65.
- Fung, M., Thornton, A., Mybeck, K., Wu, J.H.-H., Hornbuckle, K. and Muniz, E. (2001) Evaluation of the characteristics of safety withdrawal of prescription drugs from worldwide pharmaceutical markets – 1960 to 1999. *Drug Inf J* **35**, 293–317.
- Haasio, K., Sopanen, L., Vaalavirta, L., Linden, I.-B. and Heinonen, E.H. (2001) Comparative toxicological study on the hepatic safety of entacapone and tolcapone in the rat. *J Neural Transm* **108**, 79–91.
- Haasio, K., Koponen, A., Penttila, K.E. and Nissinen, E. (2002) Effects of entacapone and tolcapone on mitochondrial membrane potential. *Eur J Pharmacol* **453**, 21–26.
- Lasser, K.E., Allen, P.D., Woolhandler, S.J., Himmelstein, D.U., Wolfe, S.M. and Bor, D.H. (2002) Timing of new black box warnings and withdrawals for prescription medications. *JAMA* **287**, 2215–20.

- Kaplowitz, N. (2001) Drug-induced liver disorders: implications for drug development and regulation. *Drug Saf* **24**, 483–90.
- Kaplowitz, N. (2005) Idiosyncratic drug hepatotoxicity. *Natl Rev Drug Disc* **4**, 489–99.
- Napolitano, A., Bellini, G., Borroni, E., Zurcher, G. and Bonucelli, U. (2003). Effects of peripheral and central catechol-*O*-methyltransferase inhibition on striatal extracellular levels of dopamine: a microdialysis study in freely moving rats. *Parkinsonism Relat Disord* **9**, 145–50.
- Liang, K.-Y. and Zeger, S.T. (1986) Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.
- Lynch, M. (2005) An automated method for highly abundant protein partitioning of human plasma samples prior to proteomic analysis. *Am Biotechnol Lab* **23**, 24–26.
- Olanow, C.W. (2000) Tolcapone and hepatotoxic effects. *Arch Neurol* **57**, 263–67.
- Olanow, C.W. and Watkins, P.B. (2007) Tolcapone: an efficacy and safety review (2007). *Clin Neuropharmacol* **30**, 287–94.
- Ostapowicz, G., Fontana, R.J., Schiodt, F.V., Larson, A., Davern, T.J., Han, S.H.B., McCashland, T.M., Shakil, A.O., Hay, J.E., Hynan, L., Crippin, J.S., Blei, A.T., Samuel, G., Reisch, J., Lee, W.M. and the U.S. Acute Liver Failure Study Group. (2002) Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. *Ann Intern Med* **137**, 947–54.
- Pitman, E.J.G. (1937). The closest estimates of statistical parameters. *Proc Camb Philol Soc* **33**, 212–22.
- Rotman, D. (2004) Can Pfizer deliver? *Technol Rev* February 2004.
- Schnackenberg, L.K., Dragan, Y.P., Reily, M. D., Robertson, D.G., and Beger, R.D. (2007) Evaluation of NMR spectral data of urine in conjunction with measured clinical chemistry and histopathology parameters to assess the effects of liver and kidney toxicants. *Metabolomics* **3**, 87–100.
- Shuster, D., Laggner, C., and Langer, T. (2005) Why drugs fail—a study on side effects in new chemical entities. *Curr Pharma Des* **11**, 3545–59.
- Tasmar (Tolcapone) Tablets NDA 20-697 (January 29, 1998). Available at http://www.fda.gov/cder/foi/nda/98/20697_Tasmar.htm.
- Temple, R.J. and Himmel, M.H. (2002) Safety of newly approved drugs. *JAMA* **287**, 2273–75.
- Tong, W., Cao, X., Harris, S., Sun, H., Fuscoe, J., Harris, A., Hong, H., Xie, Q., Perkins, R., Shi, L. and Casciano, D. (2003) ArrayTrack: Supporting toxicogenomic research at the U.S. Food and Drug Administration National Center for Toxicological Research. *Environ Health Perspect* **111**, 1819–26.
- Tong, W., Harris, S., Cao, X., Fang, H., Shi, L., Sun, H., Fuscoe, J., Hong, H., Xie, Q., Perkins, R. and Casciano, D. (2004) Development of a Public Toxicogenomics Software for Microarray Data Management and Analysis. *Mutat Res* **549**, 241–53.
- Tornwall, M. and Mannisto, P.T. (1993). Effects of three types of catechol *O*-methylation inhibitors on L-3,4,-dihydroxyphenylalanine-induced circling behaviour in rats. *Eur J Pharmacol* **250**, 77–84
- van der Greef, J., Martin, S., Juhasz, P., Adourian, A., Plasterer, T., Verheij, E.R. and McBurney, R.N. (2007) The Art and Practice of Systems Biology in Medicine: Mapping Patterns of Relationships. *J Proteome Res* **4**, 1540–59.
- Watkins, P. (2000) COMT inhibitors and liver toxicity. *Neurology* **55(11Suppl 4)**, S51–52.
- Zimmerman, H. (1999) *Hepatotoxicity: The Adverse Effects of Drugs and Other Chemicals on the Liver*. Lippincott Williams and Wilkins, Philadelphia, PA.