ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Toxicology and Applied Pharmacology 187 (2003) 137–146

Toxicology and Applied Pharmacology

www.elsevier.com/locate/taap

Contemporary issues in toxicology
The role of metabonomics in toxicology and its evaluation by the COMET project

John C. Lindon,[a,*] Jeremy K. Nicholson,[a] Elaine Holmes,[a] Henrik Antti,[a,1] Mary E. Bollard,[a] Hector Keun,[a] Olaf Beckonert,[a] Timothy M. Ebbels,[a] Michael D. Reily,[b] Donald Robertson,[b] Gregory J. Stevens,[c] Peter Luke,[d] Alan P. Breau,[e] Glenn H. Cantor,[f] Roy H. Bible,[e] Urs Niederhauser,[g] Hans Senn,[g] Goetz Schlotterbeck,[g] Ulla G. Sidelmann,[h] Steen M. Laursen,[h] Adrienne Tymiak,[i] Bruce D. Car,[i] Lois Lehman-McKeeman,[i] Jean-Marie Colet,[j] Ali Loukaci,[j] and Craig Thomas[j,2]

[a] *Biological Chemistry, Biomedical Sciences Division, Faculty of Medicine, Imperial College of Science, Technology and Medicine, Sir Alexander Fleming Building, South Kensington, London SW7 2AZ, UK*
[b] *Pfizer Global R&D, 2800 Plymouth Road, Ann Arbor, MI 48105, USA*
[c] *Pfizer Global R&D, La Jolla, CA, 92121, USA*
[d] *Pfizer Global R&D, Sandwich, Kent CT13 9NJ, UK*
[e] *The Pharmacia Corporation, 4901 Searle Parkway, Skokie, IL 60077, USA*
[f] *The Pharmacia Corporation, 301 Henrietta St., Kalamazoo, MI 49007, USA*
[g] *Hoffmann-La Roche AG, Grenzacherstrasse, CH-4070 Basel, Switzerland*
[h] *Novo Nordisk, Novo Nordisk Park, 2760 Maaloev, Denmark*
[i] *Bristol-Myers-Squibb Company, PO Box 4000, Princeton, NJ 08543, USA*
[j] *Lilly Research Laboratories, Eli Lilly and Co., Lilly Development Centre S.A., 1348 Mont-Saint-Guibert, Belgium*

Received 19 September 2002; accepted 9 December 2002

Abstract

The role that metabonomics has in the evaluation of xenobiotic toxicity studies is presented here together with a brief summary of published studies. To provide a comprehensive assessment of this approach, the Consortium for Metabonomic Toxicology (COMET) has been formed between six pharmaceutical companies and Imperial College of Science, Technology and Medicine (IC), London, UK. The objective of this group is to define methodologies and to apply metabonomic data generated using ^{1}H NMR spectroscopy of urine and blood serum for preclinical toxicological screening of candidate drugs. This is being achieved by generating databases of results for a wide range of model toxins which serve as the raw material for computer-based expert systems for toxicity prediction. The project progress on the generation of comprehensive metabonomic databases and multivariate statistical models for prediction of toxicity, initially for liver and kidney toxicity in the rat and mouse, is reported. Additionally, both the analytical and biological variation which might arise through the use of metabonomics has been evaluated. An evaluation of intersite NMR analytical reproducibility has revealed a high degree of robustness. Second, a detailed comparison has been made of the ability of the six companies to provide consistent urine and serum samples using a study of the toxicity of hydrazine at two doses in the male rat, this study showing a high degree of consistency between samples from the various companies in terms of spectral patterns and biochemical composition. Differences between samples from the various companies were small compared to the biochemical effects of the toxin. A metabonomic model has been constructed for urine from control rats, enabling identification of outlier samples and the metabolic reasons for the deviation. Building on this success, and with the completion of studies on approximately 80 model toxins, first expert systems for prediction of liver and kidney toxicity have been generated.

Keywords: Metabonomics; Toxin; Pharmaceutical; NMR; Chemometrics; Expert system

* Corresponding author. Fax: +44-0-20-7594-3066.
E-mail address: j.lindon@ic.ac.uk (J.C. Lindon).
[1] Current address: Metabometrix Ltd., RSM, Prince Consort Road, London SW7 2BP, UK.
[2] Current address: Lilly Research Laboratories, Greenfield IN 46163, USA.

0041-008X/03/$ – see front matter
doi:10.1016/S0041-008X(02)00079-0

Introduction

The importance of postgenomic technologies for improving the understanding of drug adverse effects has been highlighted recently (Aardema and MacGregor, 2002; Cockerell et al., 2002) and these approaches have been recognized to include metabonomics. While there is a comprehensive literature on the use of metabonomics to investigate xenobiotic toxicity and this has been reviewed recently (Nicholson et al., 2002), a rigorous and comprehensive evaluation would be of considerable value. To this end, a consortium has been formed to investigate the utility of NMR-based metabonomic approaches to the toxicological assessment of drug candidates. The main aim of the consortium is to use ^{1}H NMR spectroscopy of biofluids (and, in selected cases, tissues), with the application of computer-based pattern recognition and expert system methods to classify the biofluids in terms of known pathological effects caused by administration of substances causing toxic effects. The project is hosted at Imperial College of Science, Technology and Medicine (IC), University of London, UK, and involves funding by six pharmaceutical companies, namely, Bristol-Myers-Squibb, Eli Lilly and Co., Hoffman–La Roche, NovoNordisk, Pfizer Incorporated, and The Pharmacia Corporation.

The main objectives of the project are (1) provision of a detailed multivariate description of normal physiological and biochemical variation of metabolites in urine, blood serum, and selected tissues, for primarily selected male rat and mouse strains, based on ^{1}H NMR spectra; (2) development of a database of ^{1}H NMR spectra from animals dosed with model toxins, initially concentrating on liver and kidney effects; (3) development of expert systems for the detection of the toxic effects of xenobiotics based on a chemometric analysis of their NMR-detected changes in biofluid metabolite profiles; (4) identification of combination biomarkers of the various defined classes; (5) testing of the methods to assess the ability of metabonomics to distinguish between toxic and nontoxic analogues and to assess the specificity of the predictive expert systems. The classes of chemicals used and the types of toxicity investigated are as diverse as possible to assist the validation of NMR methods for use in early "broad" screening of candidates for toxicity.

In this concise review, the background to metabonomics is presented together with a survey of literature results where metabonomics has been used to probe xenobiotic toxicity based on target organ, regions within target organs, and biochemical mechanisms of action. The methods used in COMET are briefly summarized and results are used to exemplify the approach. Finally, a perspective on the future uses of metabonomics in drug safety assessment is given.

Metabonomics background

NMR spectroscopy of biofluids

When toxins interact with cells and tissues they disturb the ratios, concentrations, and fluxes of endogenous biochemicals in key intermediary cellular metabolic pathways. Under mild toxic stress, cells attempt to maintain homeostasis and metabolic control by varying the composition of the body fluids that either perfuse them or are secreted by them. In more severe toxicity states, cell death leads to loss of organ function and more marked biochemical changes occur in biofluids due to loss of whole body homeostasis and metabolite leakage from damaged cells. Consequently, following either scenario there are characteristic organ-specific and mechanism-specific alterations in biofluid composition. Clearly, the detection of toxic lesions via biochemical effects is most difficult close to the toxic threshold, yet these are often the most important effects to define. Previously, the detection of novel biomarkers of toxic effect has mainly been serendipitous. However, it is now possible to use a combined NMR–expert system approach to systematically explore the relationships between biofluid composition and toxicity and to generate novel combination biomarkers of toxicity. The approach of characterizing the metabolic profile of a specific tissue or biofluid has been termed "metabonomics" by analogy with genomics and proteomics. Metabonomics has been defined as the study of the time-related quantitative multivariate metabolic response to pathophysiological processes or genetic modification in cells, tissues, and whole organisms (Nicholson et al., 1999).

The successful application of ^{1}H NMR spectroscopy of biofluids to study a variety of metabolic diseases and toxic processes has now been well established and many novel metabolic markers of organ-specific toxicity have been discovered (Nicholson and Wilson, 1989). ^{1}H NMR spectroscopy is well suited to the study of toxic events, as multicomponent analyses on biological materials can be made simultaneously, without bias imposed by expectations of the type of toxin-induced metabolic changes. This is particularly true for NMR spectra of urine in situations where damage has occurred to the kidney or liver. Quantitative changes in NMR spectroscopic metabolite patterns have also been shown to give information on the location and severity of toxic lesions, as well as insights into the underlying molecular mechanisms of toxicity. (Nicholson et al., 1985; Nicholson and Wilson, 1989).

The first studies of using PR to classify biofluid samples used a simple scoring system used to describe the levels of 18 endogenous metabolites in urine from rats which either were in a control group or had received a specific organ toxin which affected the liver, the testes, the renal cortex, or the renal medulla (Gartland et al., 1990, 1991). This study showed that samples corresponding to different organ toxins mapped into distinctly different regions. Various refinements in the data analysis were investigated, including tak-

ing scored data at three time points after the toxin exposure for the nephrotoxins only (this used only 16 metabolites, as taurine and creatine were not altered in this data subset) as well as using a simple dual scoring system (the time and magnitude of the greatest change from control). The maps derived from the full time course information provided the best discrimination between toxin classes. This study was further extended (Anthony et al., 1994b) to incorporate actual metabolite NMR resonance intensities rather than simple scores. This was carried out for the nephrotoxins in the earlier group plus additional nephrotoxic compounds. A good separation of renal medullary from renal cortical toxins was achieved. In addition, it was possible to differentiate cortical toxins according to the region of the proximal tubule which was affected and also by the biochemical mechanism of the toxic effect.

The time course of metabolic urinary changes induced by two renal toxins was first investigated in detail by metabonomics using Fisher 344 rats administered a single acute dose of the renal cortical toxin mercuric chloride and the medullary toxin 2-bromoethanamine (Holmes et al., 1992). The rat urine was collected for up to 9 days after dosing and was analyzed using ^{1}H NMR spectroscopy. The onset, progression, and recovery of the lesions were also followed using histopathology to provide a definitive classification of the toxic state relating to each urine sample. The concentrations of 20 endogenous urinary metabolites were measured at eight time points after dosing and mapping methods were used to reduce the data dimensionality. These showed that the points on the plot can be related to the development of, and recovery from, the lesions.

A wide range of toxins has now been investigated using this metabonomics approach, including the kidney cortical toxins mercury chloride (Nicholson et al., 1985), *p*-aminophenol (Gartland et al., 1989; Sanins et al., 1990b), uranyl nitrate (Anthony et al., 1994a), ifosfamide (Foxall et al., 1996), cephaloridine (Anthony et al., 1992), the kidney medullary toxins propylene imine and 2-bromoethanamine hydrochloride (Holmes et al., 1992; Robertson et al., 2000), and the liver toxins hydrazine, allyl alcohol, thioacetamide, α-naphthylisothiocyanate, and carbon tetrachloride (Sanins et al., 1990a, 1990b; Nicholls et al., 2001; Robertson et al., 2000). The testicular toxin cadmium chloride has also been investigated in detail (Nicholson et al, 1989), including the effects of chronic exposure at environmentally realistic levels (Griffin et al., 2001). The incidence of phospholipidosis caused by amiodarone, chloroquine, DMP-777 (a neutrophil elastase inhibitor), and a number of Glaxo Wellcome compounds has been evaluated using metabonomics (Espina et al., 2001; Nicholls et al., 2000), as has the effect of dexamethasone on vascular lesions (Slim et al., 2002). Other studies include the toxicity of the aldose reductase inhibitor HOE-843 (Hoyle et al., 1992) and lanthanum nitrate (Feng et al., 2002). Toxic stress in earthworms has also been investigated using metabonomics (Bundy et al., 2001).

Extensions to the earlier chemometric approaches include a toxicological assessment approach based on neural network software to ascertain whether the methods provide a robust approach which could lead to automatic toxin classification (Anthony et al., 1995). The neural network approach to sample classification, based on ^{1}H NMR spectra of urine, was in general predictive of the sample class. It appears to be reasonably robust and once the network is trained, the prediction of new samples is rapid and automatic. However, the principal disadvantage is common to all neural network studies in that it is difficult to ascertain from the network which of the original sample descriptors are responsible for the classification. Nevertheless as shown recently, probabilistic neural networks appear to be a useful and effective method for sample classification (Holmes et al., 2001). The use of statistical batch processing has been explored as a means of characterizing the biochemical changes associated with hydrazine-induced toxicity (Antti et al., 2002). Recently, comprehensive studies have been published using pattern recognition to predict and classify drug toxicity effects, including lesions in the liver (Beckwith-Hall et al., 1998) and kidney (Holmes et al., 1998a), and using supervised methods as an approach to an expert system (Holmes et al., 1998a, 1998b; Holmes et al., 2000).

It appears that significant analytical advantages may be conferred by using ^{1}H NMR spectroscopy to follow the biochemical responses of animals or cells to foreign compounds. Currently, in order to evaluate the toxicity of a drug candidate by conventional toxicological procedures the subjective selection of a range of biochemical methods is required. This is necessarily a complex and time-consuming process and if an inappropriate range of biochemical methods or metabolic parameters are used important metabolic disturbances may be overlooked. The role of NMR spectroscopy in analytical toxicology is thus essentially one of biochemical exploration, i.e., determining the range of biochemical perturbations caused by exposure to a toxin and whether these are biologically significant.

MAS NMR of tissues

As described above, NMR spectroscopy of biofluids when coupled with pattern recognition analysis can be an efficient new method of investigating toxicity profiles of xenobiotics. In addition, while NMR spectroscopy in vivo might be used to investigate abnormalities in whole animals, such studies are hampered by the heterogeneity of the sample, low magnetic fields of whole body scanners (low sensitivity and poor spectral dispersion), and short NMR relaxation times, all leading to broad lines and loss of resolution. Within the last few years, with the development of high-resolution magic angle spinning (MAS) technology, it has become possible to obtain very high quality ^{1}H NMR spectra on small (~10-mg) samples of whole tissue with no sample pretreatment, other than addition of a very small amount of D_2O to the sample.

MAS involves spinning the sample about an axis at 54.7°

to the magnetic field direction. This process removes the line broadening caused by dipolar couplings, chemical shift anisotropies, and anisotropic magnetic susceptibility artifacts. Tissue metabolites already enjoy a degree of molecular mobility such that the line broadenings are greatly reduced from their true solid values and this means that only modest rotation rates (<10 kHz) are required. This approach has now been applied to cells and tissues (Cheng et al., 1996, 1997, 1998; Millis et al., 1997; Moka et al., 1997, 1998; Tomlins et al., 1998; Weybright et al., 1998).

The technique opens up many diagnostic possibilities since information on a variety of metabolites in different cellular environments can be rapidly obtained and specialized NMR experiments, such as those to measure molecular diffusion coefficients, can be used to probe compartmentation. Confirmation of biochemical composition can be obtained using standard high-resolution NMR of both aqueous (protein-free) and methanolic extracts. This produces a comprehensive set of metabolic information that can be used in integrated metabonomics studies.

Animal tissues can be examined after exposure to model organ-specific toxins and drugs using MAS and standard NMR methods, allowing investigation of the time-related changes in biochemical profiles. MAS NMR data, like biofluid NMR spectra, can also be subjected to computer pattern recognition methods in order to classify toxicity type (target organ and biochemical mechanism) and to map time-related biochemical trajectories associated with drug-induced biochemical changes (Nicholson et al., 2002). The ability to compare biofluid and tissue NMR spectra may provide further insight into mechanisms of toxicity or target organ identification.

Chemometric analysis of metabolic NMR data

One general procedure which has found wide application is to first simplify ^{1}H NMR spectra of biofluids by means of a data compression, by producing a segmentation of each NMR spectrum (usually with about 250 intensity values per spectrum), integrating peak intensity in each segment. Each of these acts as a metabolic descriptor with which to classify the NMR spectra according to biochemical features. These data are then constructed in the form of a spreadsheet which is used as the input into a pattern recognition/multivariate statistics software suite. Appropriate data reduction routines, such as principal components analysis (PCA) or partial least squares discriminant analysis (PLS-DA), can be used to classify the NMR-generated toxicity data in terms of toxin type and dose. Multidimensional metabolic trajectories can be constructed in order to visualize the biochemical time course of the toxic episodes (Holmes et al., 1992). More complex expert systems based on chemometric models in the multidimensional metabolic space can be constructed and used for class prediction (Holmes et al., 1998b).

Identification of novel biomarkers of toxicity

Pattern recognition maps can be examined for evidence of clustering of data according to site and type of toxic lesion. For example, PCA-derived classification can be interrogated in relation to the input data by examination of the PC loadings in order to identify key portions of the NMR spectra giving rise to classification. These parts of the spectra contain the biomarker signals. The spectra are then analyzed in detail in the uncompressed mode to assign the NMR signals and identify the metabolites. If necessary, new biomarker metabolites which cannot be identified using multipulse and multidimensional NMR spectroscopy can be separated and identified using solid-phase extraction–NMR–mass spectrometry and HPLC.

Expert system development

A variety of different supervised pattern recognition methods have been evaluated for detecting abnormalities in metabolite profiles caused by toxins (Holmes et al., 1998a, 1998b, 2000, 2001; Antti et al., 2002). These are of two main types, those which relate to overall variance in the data sets, such as those based on latent variables, e.g., SIMCA and PLS-DA, and those which examine relationships in the data in different ways, such as neural networks and genetic algorithms. The basic procedure is to train types of expert systems to produce classifications of biofluid samples based on known toxicity type according to standard methods of evaluation, such as histopathology. These systems are then tested against standard toxicological assessment procedures using toxins unknown to the model in order to evaluate the robustness of each expert system approach for toxicity screening.

The COMET project

The pharmaceutical industry is increasingly interested in the possibility of metabonomics being useful for toxicity screening and so an initial meeting of interested parties was held in January 1999 and a Steering Committee to oversee the project was formed from selected and agreed personnel from each company and IC. The consortium operation commenced in January 2001 and is funded for 3 years.

The consortium is generating databases of ^{1}H NMR spectra of rat and mouse urine and blood serum initially in male animals. In addition, selected samples will be subjected to two-dimensional NMR spectroscopy. Selective sampling of tissues is also being performed and these will form the basis of subsequent databases. Over the lifetime of the project, it is intended that approximately 150 compounds (i.e., 25 per company) will be studied. The choice of compounds is determined by the Steering Committee, with close attention paid to the interests of each company but balanced against the need to provide a useful resource for all participants. However, the primary selection criteria are

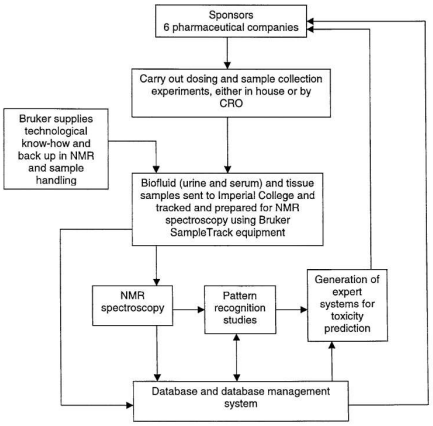

Fig. 1. Operational flow chart for the COMET project.

based on the need to provide robust predictive models that cover an extensive range of metabolic space. The NMR databases are being used to derive statistical models of various pathological situations so that biofluid spectra can be classified in these terms on the basis of the observed biochemical changes. The endogenous metabolites which are changed significantly between classes will thus be determined ("biomarkers"). The NMR spectra are measured at IC and the pattern recognition analyses are also being undertaken there. The methodology and results are transferred to the member pharmaceutical companies on a stepwise basis electronically. Training in applicable technologies will be given to each member pharmaceutical company such that at the end of the nominal 3-year period, each will be capable of conducting studies in-house and be able to internally expand the database established through the consortium. The operations of the project and the ways in which they are interconnected are shown in Fig. 1.

It must be emphasised that COMET is an academic project and it is intended and agreed that all results will be published in the scientific literature. Indeed, several studies have already been the subject of open presentations, including a number of conference posters at the annual Society of Toxicology meetings and the first manuscripts arising from COMET have now been accepted for publication.

Experimental methods used in COMET

Animal experiments and sample collection

The organization of the animal experiments is subject to a standard protocol. All animal studies are carried out in accordance with relevant national legislation and are subject to appropriate local review. Histopathology and clinical chemistry results are also included in the database. Tissues are being saved and toxicity assessment will be selected as appropriate according to the animal model used, compound pharmacology, pharmacokinetics, and toxicity.

Ten Sprague–Dawley rats or 8 B6C3F mice per group are randomly assigned to control, low-dose, or high-dose treatment groups. Animals are fed and watered ad libitum. Half of the animals from each group are euthanized at 48 h postdose and the remaining 5 animals at 168 h postdose. Where possible compounds are administered as solutions in saline or corn oil. The high-dose group, defined as a dose which exerts a clear toxic effect following a single administration, is chosen based on literature or in-house data where possible. An acceptable low dose is one which generates a threshold response. Where possible, the dose administration is a single po administration, but multiple dosing is used where necessary.

All data are recorded in Excel templates and submitted to IC together with the urine samples. Blood is sampled into microcentrifuge tubes at 24, 48, and 168 h postdosing, by puncture of the tail vein. Serum samples are obtained from blood and standard clinical chemistry parameters are measured. Any serum sample remaining is retained frozen at $<-70°C$ for NMR analysis. Urine is collected into labeled containers that are maintained at $0 \pm 2°C$ and contain sodium azide and then stored frozen at $<-70°C$ until analysis. Remaining samples are used for classical urinalysis measurements. Samples are collected over a period of 8 days, which includes a 1-day baseline collection. The usual range of conventional parameters is measured.

All main study animals are euthanized on day 8 by overdose of anesthesia. Satellite animals are euthanized as above after collection of the last blood sample (approximately 48 h after administration of the compound). Detailed external and internal examination is conducted on all animals. Representative samples of kidneys, liver, and other organs of interest are preserved in 10% buffered formol saline. For histology, tissues are processed through to paraffin wax, sectioned, stained with hematoxylin and eosin, and examined microscopically. Spare tissue is retained for MAS NMR spectroscopy. Tissue samples are snap frozen in liquid nitrogen and then transferred to solid CO_2 and stored at $-70°C$ for examination by ^1H MAS NMR spectroscopy from all animals.

Sample preparation and NMR spectroscopy

All NMR spectra of biofluids (urine and serum) are measured at a ^1H NMR frequency of 600 MHz at Imperial College using a flow-injection process, avoiding the need for glass NMR tubes. Selected tissue samples will also be measured at 400 or 600 MHz using the MAS NMR technique.

Rat and mouse urine samples are prepared for NMR

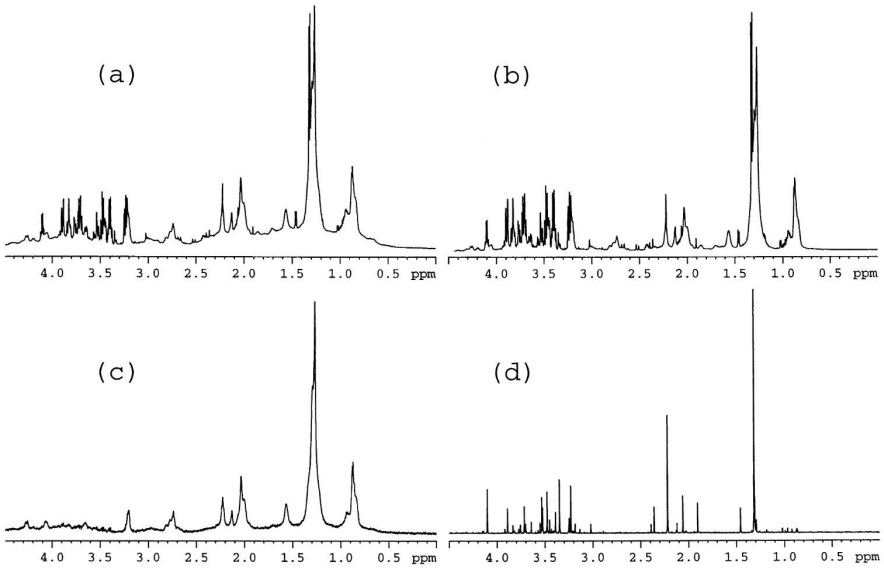

Fig. 2. Typical 600-MHz ^{1}H NMR spectra of rat serum. The assignments have been given previously (Nicholson et al., 1995). (a) The normal water-suppressed spectrum showing a broad background from both macromolecules, such as albumin, broad peaks from lipoproteins, and sharp peaks from small molecule metabolites; (b) the CPMG spin–echo spectrum in which macromolecule peaks are suppressed and lipoprotein peaks are somewhat attenuated, enhancing the importance of small molecule peaks; (c) a diffusion-edited spectrum in which the small molecule peaks are essentially suppressed and (d) the sum projection onto the chemical shift axis of a two-dimensional J-resolved spectrum in which all spin–spin coupling multiplets are converted to single peaks and severe editing of the macromolecule and lipoprotein peaks has been achieved.

spectroscopy in covered 96-well plates using the Bruker SampleTrack system and a Gilson 215 Preparation robot. In order to provide some stabilization of the urinary pH, rat or mouse urine is mixed with an aliquot of buffer in 96-well plates using a Bruker preparation robot. The volumes can be adjusted according to the NMR probe used. Approximately 500 μl of sample is injected into the probe with no air gaps in between the sample and the push solvent.

The ^{1}H NMR spectra of the urine are measured using a standard water suppression pulse sequence at 303K, resulting in a total acquisition time of about 4 min per sample.

For serum, a suite of four types of NMR spectra are measured, namely, a water-suppressed 1D ^{1}H NMR spectrum, a spin–echo spectrum using the CPMG experiment, a two-dimensional J-resolved spectrum, and a diffusion-edited spectrum. Although the two-dimensional J-resolved spectrum is obtained, pattern recognition is carried out on the sum projection onto the chemical shift axis. To illustrate the different types of metabonomic response which can easily be observed for serum, Fig. 2 shows typical ^{1}H NMR spectra of rat serum based on the four approaches above. The major peaks have been assigned previously (Nicholson et al., 1995).

Overview of results to date

Critical issues requiring resolution

It was clear from the outset that there would be a number of complex and critical issues to be resolved if this project were to be successful. Thus, considerable time was spent in ensuring that the group derived robust protocols and operating procedures which could be used at all of the laboratories involved.

First, it was important to define a set of compounds which gave reproducible and well-defined toxic effects and that the observed effects could be described in a common terminology across the many toxicologists and pathologists involved in the companies, the university, and in contract research organizations (CROs) used by some companies. The standard operating protocols used by the companies and

CROs were refined such that unwanted biological variation was kept to a minimum. To this end, detailed consideration was given to animal strain, food, adaptation to metabolic cages, diurnal variation, and animal stress. Thought was also given to the possibility of the presence of the dosed xenobiotic and its metabolites in the biofluids. By these means it was hoped that the biochemical consistency of biofluid samples from the different companies would be high.

For such a multisite set of studies it was also important to consider the analytical reproducibility of the techniques used for obtaining the metabonomic data. NMR spectroscopy was the analytical method of choice as it is known to provide highly reproducible data for replicate samples and it allows high precision and accuracy. The sample preparation and NMR measurement protocols were devised taking these requirements into account. Thus, for example, the samples are processed using dilution with strong buffer to stabilize the pH range of urine samples, thereby minimizing any pH-dependent chemical shift differences. Additionally sodium azide is used during urine collection to prevent bacterial contamination. During the first period of the project, a detailed examination of the biological quality control was undertaken, with each company dosing the same toxin to rats and with IC performing all NMR measurements on all samples. The NMR spectroscopic quality control was further tested by a procedure whereby a comparison was conducted of IC-measured and company-measured NMR spectra.

As mentioned above, the NMR spectra can show peaks that arise from the dosed toxin or its metabolites. To avoid confounding the chemometric models, the NMR peaks associated with toxins or their metabolites are removed before the pattern recognition analysis proceeds. It was therefore necessary to have a method which could achieve this in an automated fashion on large batches of spectra. In the situation where there are only a few toxins in a database and the xenobiotic-related peaks are few, then it is possible to simply cut out all the regions which are "contaminated." However, this has to be done for all regions for all spectra so that there are no missing NMR descriptor values. For large databases where there are many diverse toxins being studied, with a wide range of molecular structures and possibilities of metabolism, it is easy to see that essentially all of the spectral regions might have to be eliminated. This has been overcome by development of software which replaces the different "contaminated" regions of each spectrum, allowing most of the different regions of the various spectra to be retained.

A statistical model for control rat and mouse urine

Given the availability of thousands of control urine samples for both rat and mouse, the project has provided an excellent opportunity to develop an understanding of the biochemical variation in these fluids by building statically based models of control urine. These models are continuously evolving as more samples become available. The control model successfully detects outlier samples known to be deviating from normal control behavior based on clinical chemistry measurements and also successfully detects early time point deviations in animals for which a full time course was not available. Early fault detection such as this could be used as an integrated part of a fully automated expert system in order to discard abnormal samples at the earliest possible stage.

Intersite consistency of biofluid sample provision

In order for a database of spectra generated on samples from a number of different laboratories to be useful, it is necessary that interlaboratory differences be small compared to the biochemical effects caused by the treatments. To this end, a study has been conducted in all of the sponsors' laboratories using hydrazine as a model toxin (Amenta and Johnston, 1962; Waterfield et al., 1993; Sanins et al., 1990a, 1990b, 1992; Nicholls et al., 2002) and working from an agreed, standard protocol. In fact, one member carried out two studies, one in-house and the other at a contract research organization in the UK. Urine and plasma samples were collected according to the standard protocol and were shipped to IC, where the NMR and pattern recognition analyses were carried out.

The ^{1}H NMR spectra of the urine and plasma samples showed a high degree of consistency across companies, with clear differences observable between the samples from control and dosed animals. The spectra were subjected to pattern recognition analysis and chemometric models were built. Pattern recognition derived metabolic trajectories for the low-dose group animals showed that the largest changes were seen in the 0- to 8-h period, followed by slower changes over the period 8–48 h, with apparent recovery by 72 h postdose. In the high-dose dose case, toxin-induced changes occurred throughout the study period, up to 168 h postdose, at which point the animals had, in general, not fully recovered. In the model, differences between companies were small compared to differences between the control animals and dosed animals. Clear separation of control, low-dose, and high-dose animals could be seen.

A number of outlying rats were detected and many of the endogenous metabolites causing the separation of these samples from the main cluster of the data, such as citrate, taurine, and hippurate, are known to have a high interanimal variation as a result of hormonal influences, stress, or diet. The small but observable variability of animals to hydrazine toxicity observed between the different companies may be due to external factors such as food intake, temperature, and light intensity, which can vary slightly between laboratories. Despite this variability, the average trajectories for each company were similar to one another.

In summary this study showed that the differences between samples derived from different laboratories were gen-

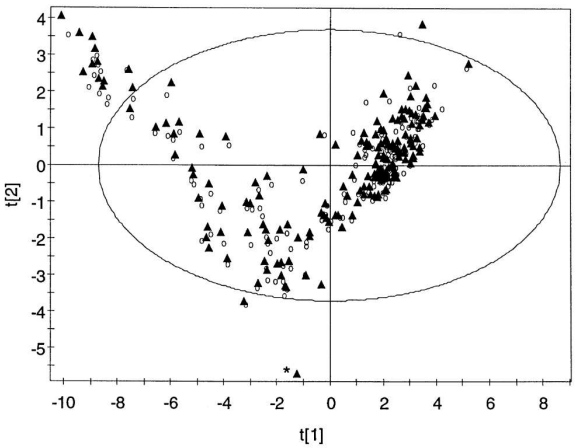

Fig. 3. An example of the high degree of biochemical consistency between NMR spectra measured at IC and at Roche on samples from the hydrazine toxicity study. This is a plot of the first and second PC scores for all NMR spectra. Triangles and circles represent NMR spectra measured at IC and Roche, respectively. The ellipse denotes the 95% significance limit. Samples appear as pairs showing similarity of biochemical profiles. The asterisk denotes a sample which gave an anomalous spectrum when measured automatically at IC, due to an instrumental artifact.

erally much smaller than the differences in biochemistry elicited by the toxin, providing a strong rationale for continuing the database generation.

Consistency of NMR measurement study

In addition, the NMR spectra measured at IC were compared to NMR spectra measured on aliquots of the same samples at two of the participating companies, Pfizer and Roche. Again, a high degree of consistency was observed between samples measured in different laboratories in the UK, Europe, and the USA, even though the NMR measurements were made using different magnetic field strengths on instruments provided by different vendors (Keun et al., 2002). This is illustrated in Fig. 3, which shows a PC scores plot comparing ^{1}H NMR spectra measured at both IC and Roche with close pairing in the plot between the same sample measured at both sites. The spread of points arises because of the onset of, and recovery from, the toxic episode caused by the single dose of hydrazine. The triangles and circles represent spectra measured at IC and Roche, respectively.

Overview of toxicity results

The project has been fully implemented for approximately 20 months and, to date, studies on about 80 compounds have been completed. Preliminary chemometric models have been derived for each study on the basis of three classes, namely, control, low-dose, and high-dose animals. At present, a detailed evaluation of prototype expert systems is being carried out, in order to be able to predict whether a sample is from a control animal or is typical of liver toxicity, typical of kidney toxicity, allows a differential diagnosis of liver versus kidney toxicity, or is of a toxicity class not yet defined.

Information is reported electronically to and from the sponsoring companies, including protocols, for both dosing and sample collection, histopathology reports and clinical chemistry data, NMR spectral data, lists of compounds studied and to be studied, and reports issued. To date, the use of electronic data capture and information exchange has allowed a relatively seamless interaction among COMET members.

The future: a perspective on metabonomics

Single markers of toxicity are unlikely to be of value in the detection of subtle lesions that are the most problematic in primary toxicology studies on novel drugs. However, such lesions would cause multiple low-level disruptions of intermediary metabolites in biofluids that would closely reflect the site and mechanisms of damage. Pattern recognition and multivariate statistical analysis of NMR spectra offers a realistic prospect of identifying these novel combination biomarkers of toxic effect for use in toxicological screening at the discovery/development stage. In COMET,

the scope of the NMR/pattern recognition approach is being broadened and NMR-based expert systems suitable for "high-throughput" in vivo toxicology screening are being devised. This effort is expected to generate new markers of drug toxicity in vivo and further the understanding of the systemic toxic mechanisms and biochemical effects of novel drugs. The output of the consortium will facilitate the development of valuable new approaches to in vivo drug toxicity screening at the discovery stage and in the investigation of more subtle lesions emerging for specific compounds during the development stage.

It is perhaps noteworthy that the U.S. Food and Drug Administration has publicized on its web pages that through its national Center for Toxicological Research, it is evaluating metabonomics as a tool for leveraging the development of new medicines and the FDA's Vascular Injury Expert Working Group is exploring the use of biomarkers via genomics, proteomics, and metabonomics. Given both the increasing numbers of publications from different laboratories in which metabonomics has been used to probe xenobiotic toxicity and the interest that the pharmaceutical industry has shown through the COMET project and by other means, it is clear that metabonomics will play an increasing role in drug safety evaluation. It can thus be expected that metabonomics will enter the mainstream activities of pharmaceutical companies and environmental agencies for identifying toxic compounds and their mechanism of action. At least for the pharmaceutical industry, this should lead to a reduction in the attrition rate of compounds failing at later stages in the R&D process.

References

Aardema, M.J., MacGregor, J.T., 2002. Toxicology and genetic toxicology in the new era of "toxicogenomics:" impact of "-omics" technologies. Mutat. Res. Fund. Mol. Mech. Mutagen. 499, 13–25.

Amenta, J.S., Johnston, E.H., 1962. Hydrazine-induced alterations in rat liver. A correlation of the chemical and histologic changes in acute hydrazine intoxication. Lab. Invest. 11, 956–62.

Anthony, M., Gartland, K.P.R., Beddell, C.R., Lindon, J.C., Nicholson, J.K., 1992. Cephaloridine-induced nephrotoxicity in the Fischer 344 rat: proton NMR spectroscopic studies of urine and plasma in relation to clinical chemical and histopathological assessments of nephronal damage. Arch. Toxicol. 66, 525–37.

Anthony, M., Gartland, K.P.R., Lindon, J.C., Beddell, C.R., Nicholson, J.K., 1994a. Biochemical characterisation of uranyl nitrate-induced nephrotoxicity in the F344 rat. Arch. Toxicol. 68, 43–53.

Anthony, M.L., Rose, V.S., Nicholson, J.K., Lindon, J.C., 1995. Classification of toxin-induced changes in ^{1}H NMR spectra of urine using an artificial neural network. J. Pharm. Biomed. Anal. 13, 205–11.

Anthony, M.L., Sweatman, B.C., Beddell, C.R., Lindon, J.C., Nicholson, J.K., 1994b. Pattern recognition classification of the site of nephrotoxicity based on metabolic data derived from proton nuclear magnetic resonance spectra of urine. Mol. Pharmacol. 46, 199–211.

Antti, H., Bollard, M.E., Ebbels, T., Keun, H., Lindon, J.C., Nicholson, J.K., Holmes, E., 2002. Batch statistical processing of ^{1}H NMR-derived urinary spectral data. J. Chemometrics 16, 461–68.

Beckwith-Hall, B.M., Nicholson, J.K., Nicholls, A.W., Foxall, P.J.D., Lindon, J.C., Connor, S.C., Abdi, M., Connelly, J., Holmes, E., 1998. Nuclear magnetic spectroscopic and principal components analysis investigations into biochemical effects of three model hepatotoxins. Chem. Res. Tox. 11, 260–72.

Bundy, J.G., Osborn, D., Weeks, J.M., Lindon, J.C., Nicholson, J.K., 2001. An NMR-based metabonomic approach to the investigation of coelomic fluid biochemistry in earthworms under toxic stress. FEBS Lett. 500, 31–5.

Cheng, L.L., Chang, L-W., Louis, D.N., Gonzalez, R.G., 1998. Correlation of high resolution magic angle spinning proton magnetic resonance spectroscopy with histopathology of intact human brain tumor specimens. Cancer Res. 58, 1825–32.

Cheng, L.L., Lean, C.L., Bogdanova, A., Wright, S.C., Ackerman, J.L., Brady, T.J., Garrido, L., 1996. Enhanced resolution of proton NMR spectra of malignant lymph nodes using magic angle spinning. Magn. Reson. Med. 36, 653–8.

Cheng, L.L., Ma, M.J., Becerra, L., Ptak, T., Tracey, I., Lackner, A., Gonzalez, R.G., 1997. Quantitative neuropathology by high resolution magic angle spinning proton magnetic resonance spectroscopy. Proc. Nat. Acad. Sci. USA 94, 6408–13.

Cockerell, G.L., McKim, J.M., Vonderfecht, S.L., 2002. Strategic importance of research support through pathology. Toxicol. Pathol. 30, 4–7.

Espina, J.R., Shockcor, J.P., Herron, W.J., Car, B.D., Contel, N.R., Ciaccio, P.J., Lindon, J.C., Holmes, E., Nicholson, J.K., 2001. Detection of *in vivo* biomarkers of phospholipidosis using NMR-based metabonomic approaches. Magn. Reson. Chem. 39, 559–65.

Feng, J.H., Li, X.J., Pei, F.K., Chen, X., Li, S.L., Nie, Y.X., 2002. ^{1}H NMR analysis for metabolites in serum and urine from rats administrated chronically with $La(NO_3)_3$. Anal. Biochem. 301, 1–7.

Foxall, P.J.D., Lenz, E.M., Lindon, J.C., Neild, G.H., Wilson, I.D., Nicholson, J.K., 1996. NMR and HPLC-NMR studies on the toxicity and metabolism of ifosfamide. Therapeut. Drug Monit. 18, 498–505.

Gartland, K.P.R., Beddell, C.R., Lindon, J.C., Nicholson, J.K., 1990. A pattern recognition approach to the comparison of ^{1}H NMR and clinical chemical data for classification of nephrotoxicity. J. Pharm. Biomed. Anal. 8, 963–8.

Gartland, K.P.R., Bonner, F., Nicholson, J.K., 1989. Investigations into the biochemical effects of region-specific nephrotoxins. Mol. Pharmacol. 35, 242–51.

Gartland, K.P.R., Lindon, J.C., Beddell, C.R., Nicholson, J.K., 1991. Application of pattern recognition methods to the analysis of toxicological data generated by NMR spectroscopy. Mol. Pharmacol. 39, 629–42.

Griffin, J.L., Walker, L.A., Shore, R.F., Nicholson, J.K., 2001. Metabolic profiling of chronic cadmium exposure in the rat. Chem. Res. Tox. 14, 1428–34.

Holmes, E., Bonner, F., Sweatman, B.C., Lindon, J.C., Beddell, C.R., Rahr, E., Nicholson, J.K., 1992. NMR spectroscopy and pattern recognition analysis of the biochemical processes associated with the progression and recovery from nephrotoxic lesions in the rat induced by mercury(II)chloride and 2-bromoethanamine. Mol. Pharmacol. 42, 922–30.

Holmes, E., Nicholls, A.W., Lindon, J.C., Connor, S.C., Connelly, J.C., Haselden, J.N., Damment, S.J.P., Spraul, M., Neidig, P., Nicholson, J.K., 2000. Chemometric models for toxicity classification based on NMR spectra of biofluids. Chem. Res. Tox. 13, 471–8.

Holmes, E., Nicholls, A.W., Lindon, J.C., Ramos, S., Spraul, M., Neidig, P., Connor, S.C., Connelly, J., Damment, S.J.P., Haselden, J.N., Nicholson, J.K., 1998a. Development of a model for classification of toxin-induced lesion using ^{1}H NMR spectroscopy of urine combined with pattern-recognition. NMR Biomed. 11, 235–44.

Holmes, E., Nicholson, J.K., Nicholls, A.W., Lindon, J.C., Connor, S.C., Polley, S., Connelly, J., 1998b. The identification of novel biomarkers of renal toxicity using automatic data reduction techniques and PCA of proton NMR spectra of urine. Chemomet. Intel. Lab. Systems 44, 245–55.

Holmes, E., Nicholson, J.K., Tranter, G., 2001. Metabonomic characterization of genetic variations in toxicological and metabolic responses using probabilistic neural networks. Chem. Res. Tox. 14, 182.

Hoyle, V.R., Gilbert, P.J., Troke, J.A., Vose, C.W., Nicholson, J.K., 1992. Studies on the biochemical effects of the aldose reductase inhibitor 2,7-difluorospirofluorene-9,5′-imidazolidine-2′,4′-dione (AL-1576, HOE-843)—detection of deuterium-glucaric and deuterium-glucuronic acid excretion by high-resolution ^{1}H NMR and ^{13}C NMR spectroscopy. Biochem. Pharmacol. 44, 231–41.

Keun, H.C., Ebbels, T.M.D., Antti, H., Bollard, M., Beckonert, O., Schlotterbeck, G., Senn, H., Niederhauser, U., Holmes, E., Lindon, J.C., Nicholson, J.K., 2002. Physiological variation and analytical reproducibility in metabonomic urinalysis. Chem. Res. Tox. 15, 1380–1386.

Millis, K.K., Maas, W.E., Cory, D.G., Singer, S., 1997. Gradient high resolution magic angle spinning NMR spectroscopy of adipose tissue. Magn. Reson. Med. 38, 399–403.

Moka, D., Vorreuther, R., Schicha, H., Spraul, M., Humpfer, E., Lipinski, M., Foxall, P.J.D., Nicholson, J.K., Lindon, J.C., 1997. Magic angle spinning ^{1}H nuclear magnetic resonance spectroscopy of intact kidney tissue samples. Anal. Commun. 34, 107–9.

Moka, D., Vorreuther, R., Schicha, H., Spraul, M., Humpfer, E., Lipinski, M., Foxall, P.J.D., Nicholson, J.K., Lindon, J.C., 1998. Biochemical classification of kidney carcinoma biopsy samples using magic-angle-spinning ^{1}H nuclear magnetic resonance spectroscopy. J. Pharm. Biomed. Anal. 17, 125–32.

Nicholls, A.W., Holmes, E., Lindon, J.C., Shockor, J.P., Farrant, R.D., Haselden, J.N., Damment, S.J.P., Waterfield, C.J., Nicholson, J.K., 2002. Metabonomic investigations into hydrazine toxicity in the rat. Chem. Res. Tox. 14, 975–87.

Nicholls, A.W., Nicholson, J.K., Haselden, J.N., Waterfield, C.J., 2000. A metabonomic approach to the investigation of drug-induced phopspholipidosis: an NMR spectroscopy and pattern recognition study. Biomarkers 5, 410–23.

Nicholson, J.K., Connelly, J., Lindon, J.C., Holmes, E., 2002. Metabonomics: a platform for studying drug toxicity and gene function. Nature Rev. Drug Disc. 1, 153–61.

Nicholson, J.K., Foxall, P.J.D., Spraul, M., Farrant, R.D., Lindon, J.C., 1995. 750 MHz ^{1}H and ^{1}H-^{13}C NMR spectroscopy of human blood plasma. Anal. Chem. 67, 793–811.

Nicholson, J.K., Higham, D., Timbrell, J.A., Sadler, P.J., 1989. Quantitative ^{1}H NMR urinalysis studies on the biochemical effects of acute cadmium exposure in the rat. Mol. Pharmacol. 36, 398–404.

Nicholson, J.K., Lindon, J.C., Holmes, E., 1999. "Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica 29, 1181–9.

Nicholson, J.K., Timbrell, J.A., Sadler, P.J., 1985. Proton NMR spectra of urine as indicators of renal damage: Mercury nephrotoxicity in rats. Mol. Pharmacol. 27, 644–51.

Nicholson, J.K., Wilson, I.D., 1989. High resolution proton magnetic resonance spectroscopy of biological fluids. Prog. NMR Spectrosc. 21, 449–501.

Potts, B.C.M., Deese, A.J., Stevens, G.J., Reily, M.D., Robertson, D.G., Theiss, J., 2001. NMR of biofluids and pattern recognition: assessing the impact of NMR parameters on the principal component analysis of urine from rat and mouse. J. Pharm. Pharmacol. 26, 463–76.

Robertson, D.G., Reily, M.D., Sigler, R.E., Wells, D.F., Paterson, D.A., Braden, T.K., 2000. Metabonomics: evaluation of nuclear magnetic resonance (NMR) and pattern recognition technology for rapid in vivo screening of liver and kidney toxicants. Toxicol. Sci. 57, 326–37.

Sanins, S.M., Nicholson, J.K., Elcombe, C., Timbrell, J.A., 1990a. Hepatotoxin-induced hypertaurinuria: a proton NMR study. Arch. Toxicol. 64, 407–11.

Sanins, S.M., Timbrell, J.A., Elcombe, C.R., Nicholson, J.K., 1990b. Hypertaurinuria induced by exposure to hepatotoxins. Arch. Toxicol. 64, 407–11.

Sanins, S.M., Timbrell, J.A., Elcombe, C., Nicholson, J.K., 1992. Proton NMR spectroscopic studies on the metabolism and biochemical effects of hydrazine in vivo. Arch. Toxicol. 66, 489–95.

Slim, R.M., Robertson, D.G., Albassam, M., Reily, M.D., Robosky, L., Dethloff, L.A., 2002. Effect of dexamethasone on the metabonomics profile associated with phosphodiesterase inhibitor-induced vascular lesions in rats. Toxicol. Appl. Pharmacol. 183, 109–16.

Tomlins, A.M., Foxall, P.J.D., Lindon, J.C., Lynch, M.J., Spraul, M., Everett, J.R., Nicholson, J.K., 1998. High resolution magic angle spinning ^{1}H NMR analysis of intact prostatic hyperplastic and tumor tissues. Anal. Commun. 35, 113–15.

Waterfield, C.J., Turton, J.A., Scales, M.D.C., Timbrell, J.A., 1993. Investigations into the effects of various hepatotoxic compounds on urinary and liver taurine levels in rats. Arch. Toxicol. 67, 244–54.

Weybright, P., Millis, K., Campbell, N., Cory, D.G., Singer, S., 1998. Gradient high resolution magic angle spinning ^{1}H NMR spectroscopy of intact cells. Magn. Reson. Med. 39, 337–44.