Use and value of metabolism databases

David R. Hawkins

A key objective during drug discovery is to ensure selection of lead compounds for development that have the optimum pharmacodynamic profile without undesirable toxicity. Metabolism plays a key role in determining the biological activity of compounds *in vivo*, and is therefore an important factor to consider during the discovery phase. The ability to predict the biotransformation pathways for specific chemical structures provides

the opportunity to implement structural modifications that might advantageously modify these processes. Potentially valuable tools include knowledge databases that contain searchable information on the known metabolites of existing compounds and databases that are designed to predict the metabolites.

Metabolism has been an established part of drug development for several decades because of the recognition of its important contribution to pharmacological and toxicological profiles. The compounds produced by metabolism are dependent on the chemical structure of the drug, and it is therefore logical that the effects of metabolism are examined during the discovery phase. Factors that affect which metabolites are formed include lipid solubility, which functional groups are present and their positional relationships, size, shape and steric factors, all of which will influence the fit and electrostatic interactions of a molecule to the active sites of drug metabolizing enzymes. The advent of combinatorial chemistry has both provided the means and accentuated the need to establish criteria and procedures for optimum selection of lead compounds for development. One approach has been to use high-throughput screening (HTS) techniques, mainly *in vitro*, although this inevitably has some limitations.

The objective is to ensure that, as far as possible, lead compounds are synthesized that possess the optimum pharmacodynamic profile for the intended clinical use without undesirable toxicity. One approach is to use existing knowledge to predict the effects of metabolism and then to propose modifications that might enhance or limit metabolism or direct it to alternative, more desirable pathways. Knowledge of likely species differences for specific structures might also be important to provide feedback on the relevance of some HTS techniques that involve non-human test systems. Two databases have therefore been developed for metabolism data, namely knowledge-based and predictive databases. The knowledge-based databases consist of compilations of published information on metabolites and can be searched in several different ways, including by chemical structure and sub-structure. This enables retrieval of relevant information on the known metabolites of compounds with similar structures or on compounds containing specific common moieties. By contrast, predictive databases attempt to portray the metabolites of a structure based on a set of knowledge rules.

Knowledge-based systems

There are two existing knowledge-based systems, namely Metabolite, a software product produced by MDL Information Systems Inc.[1] (San Leandro, CA, USA) and the book series, Biotransformations [*Biotransformations: A Survey of the Biotransformations of Drugs and Chemicals in Animals* (1988–1996, Vols 1–7) (Hawkins, D.R., ed.), The Royal Society of Chemistry], which has subsequently been produced as a software product called Metabolism (Synopsys Scientific Systems, Leeds, UK)[2]. These products contain information on the metabolism of compounds abstracted from the scientific

David R. Hawkins, Huntingdon Life Sciences, Woolley Road, Alconbury, Huntingdon, Cambridgeshire, UK PE17 5HS. tel: +44 1480 892121, fax: +44 1480 891685, e-mail: HawkinsD@UKOrg.Huntingdon.com

 1359-6446/99/$ – see front matter PII: S1359-6446(99)01401-4

literature. Provided the quality of the abstraction and transposition of data is high, the limitations of the information obtained are dependent upon the quality of the published data. The quality is affected by the degree of confidence given to a metabolite structure assignment and, in some instances, proposed structures might be speculative based on limited physicochemical data, sometimes leading to the perpetuation of myths. However, in the past ten years, the general increase in quality has been proportional to the greater importance being placed on metabolism data. Furthermore, information on the metabolites formed is often incomplete, particularly for *in vivo* data. In certain instances, only a few of the metabolites have been identified and this might only represent a few percent of the dose. This type of information does add to the knowledge base but does not provide a profile of the relative importance of certain pathways. Hence, some additional quality judgement might be necessary to complement the retrieved information. Knowledge databases subsequently provide the means to formulate rules for the second type of product, namely prediction databases.

Uses of knowledge databases include:

- Searching for information on metabolites of compounds with similar structures to those of interest
- Searching for information on metabolites involving common sub-structures
- Finding information on specific isozymes (e.g. P450 enzymes) involved in the oxidative metabolism of known structurally related compounds
- Providing alerts to differences between *in vitro* and *in vivo* metabolism and between different species
- Providing alerts to possible toxicity associated with specific pathways and sub-structures
- Supplying information to assist in the identification of unknown metabolites for new compound(s).

Metabolism of oxazole or oxadiazole ring systems

These databases can be best illustrated by specific examples. One such example might be finding information on the likely metabolism of oxazole or oxadiazole ring systems through the use of the Biotransformations database[2] (Fig. 1). A ring-cleaved metabolite (Fig. 2c) of zonisamide (Fig. 2a) has been reported to be formed both *in vitro* by rat liver microsomes[3] and also as a major urine metabolite[4]. It is presumed that this metabolite is formed by a reductive process to give an intermediate (Fig. 2b), which is then hydrolyzed to a ketone. Reduction of the imine also occurs, giving the amine (Fig. 2d) as a minor metabolite. An acetyl derivative of this amine has also been reported as a human metabolite[5].

Ring cleaved products of risperidone (Fig. 3a, structure 1) with the phenolic ketone partial structure (Fig. 3b,

Drug Discovery Today

Figure 1. *Heterocyclic sub-structures*

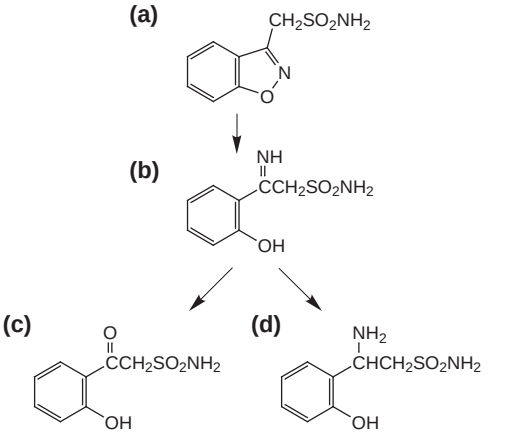

Figure 2. *Metabolites of zonisamide involving oxazole ring-cleavage.*

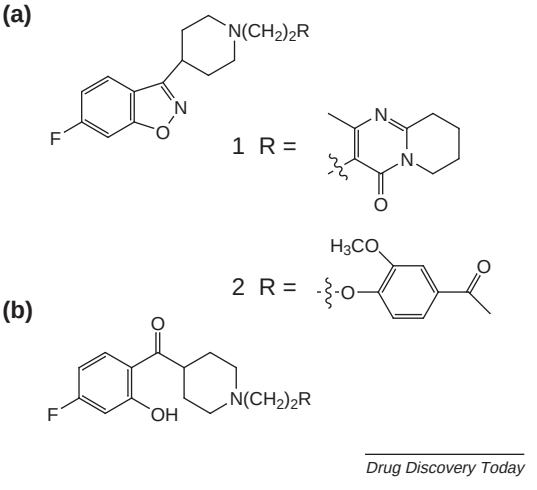

Figure 3. *Ring-cleaved metabolites (b, structures 1 and 2) of risperidone (a, structure 1) and iloperidone (a, structure 2).*

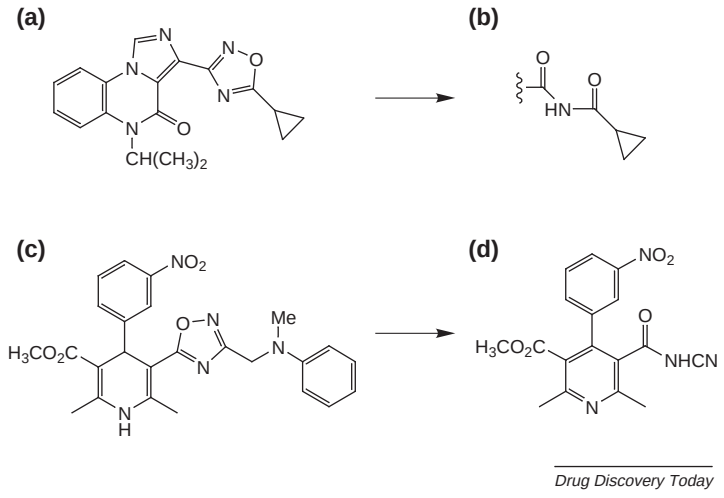

Figure 4. *Metabolites formed by cleavage of an oxadiazole ring.*

structure 1) have been identified as minor faecal metabolites in the rat, dog and human[6,7]. Similar metabolites of the related compound, iloperidone (Fig. 3a, structure 2), have been reported *in vivo* and *in vitro* in the rat and dog[8], and are apparently produced by liver enzymes.

The two compounds containing oxadiazole rings in Figs 4a and 4c might be expected to undergo reductive cleavage of the N–O bond. Interestingly, however, two very different functionalized metabolites are reported in rats, namely the imide (Fig. 4b)[9] and the cyanoamide (Fig. 4d)[10]. In fact, in the case of SM6586 (Fig. 4c), cleavage of the oxadiazole ring occurs in the absence of ester hydrolysis or methyl-group hydroxylation. The products of the ring cleavage might therefore be governed by the nature of the substituents at position 5.

Retrieval of this information indicates the inherent metabolic stability of the ring system, the nature of the metabolites formed and the structural modifications possible to enhance or limit these pathways. Cleavage of heterocyclic ring systems often produces a dramatic structural modification of the molecule, leading to the formation of metabolites with possible pharmacological implications. Novel metabolites might also be formed with structures not readily perceived, an example being the cyanoamide (Fig. 4d).

Metabolism of the indole ring

Indole rings are a common structural unit in a wide range of pharmacological agents and searching a database provides a rapid way of reviewing the pathways involved in the metabolism of this group and the impact of modifying functional substituents. Some compounds with recently reported metabolites are shown in Fig. 5. The published information does not always provide the comprehensive data on all compounds to make strict comparisons but there are often key pointers and indications of trends.

One particularly interesting observation that can be made from such a search is the relative importance of ring-hydroxylation. Of all the compounds listed in Fig. 5, sumatriptan is alone in that no ring-hydroxylation occurs, and this is partly because of the blocking substituent in position 5 and the two side-chains that provide soft targets for metabolism. In particular, the dimethylaminoethyl side-chain can be readily oxidized to a hydrophilic carboxylic acid metabolite, as shown in humans, which can also be conjugated with glucuronic acid[11]. The unsubstituted compound, DIIM, has a secondary alcoholic-group in the side-chain (which is available for direct conjugation) and the respective glucuronide is a major human metabolite with the 6-hydroxy compound being only a minor component[12]. Pindolol has a similar side-chain, although located on the phenyl ring, and while an O-glucuronide is produced, additional side-chain and ring oxidation products are formed[13]. The thiophenyl group in L734005 provides an opportunity for both S-oxidation and ring-hydroxylation, although it is accompanied by hydroxylation in position 6 (Ref. 14).

Consideration of those compounds possessing unsubstituted rings indicates that hydroxylation can occur at any of three positions. The favoured position is *meta* to the indole nitrogen while hydroxylation at the corresponding *ortho*-position only occurs in tropisetron[15] and ondansetron[16]. The side-chain of dolasetron provides the opportunity for keto-group reduction and the formation of hydrophilic conjugates, which are major pathways[17]. It is therefore apparent that structural modifications can greatly influence the extent and position of ring-hydroxylation in the indole sub-structure.

Predictive databases

Knowledge databases provide the basis for development of predictive programmes. The concept of prediction in metabolism can have different interpretations according to the perspective of the individual. There is now comprehensive

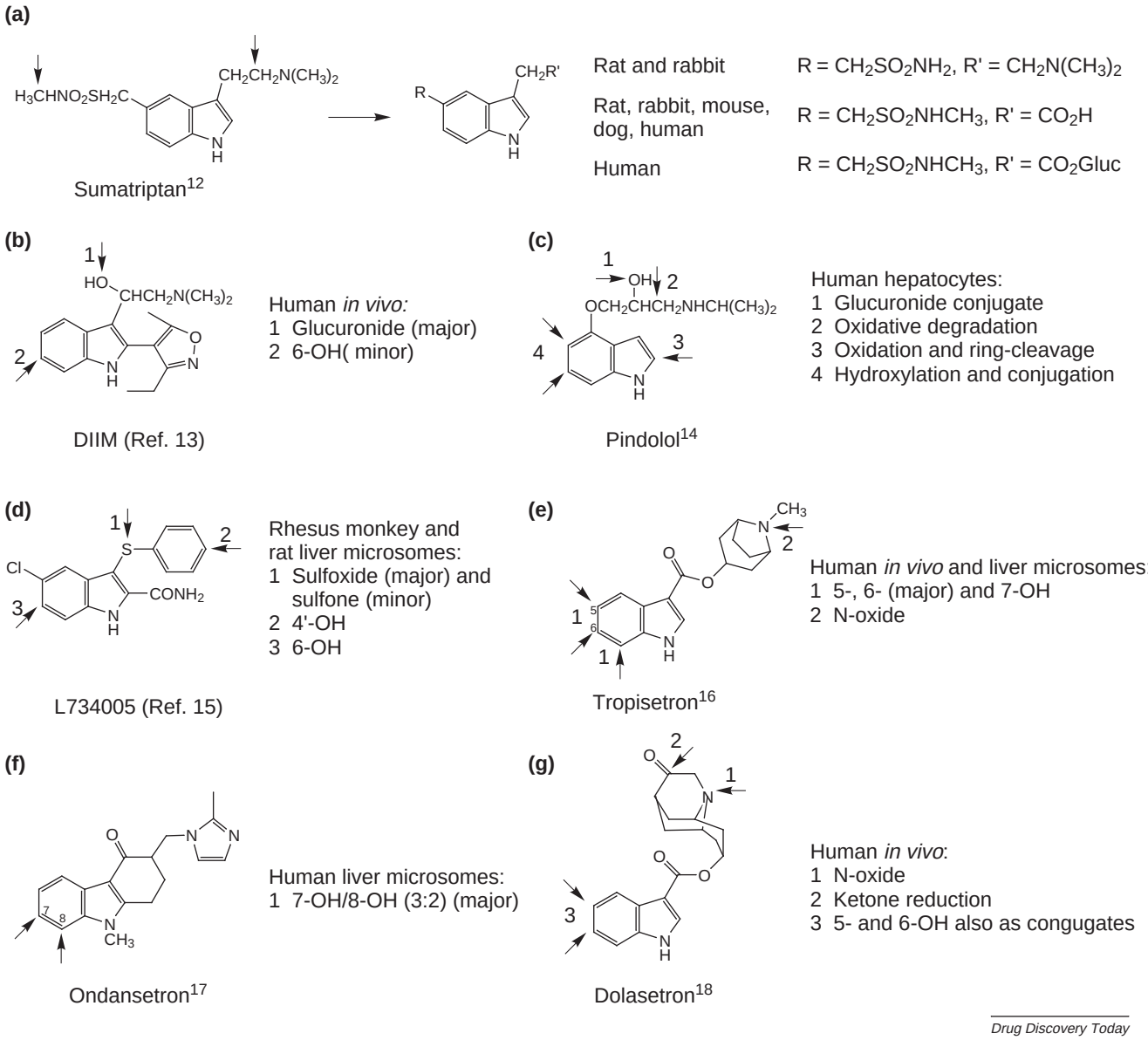

Figure 5. *Metabolism of some drugs containing an indole sub-structure.*

knowledge of the biotransformation pathways and, for any given compound, it is relatively straightforward to predict the sites where metabolism could occur and the structure of the resulting metabolites. However, even for moderately complex molecules, the number of potential metabolites from different pathways is often >100. The role of a predictive database is therefore to highlight the most likely metabolites that will be formed and any species differences in the pathways. This type of product functions on the basis of a database of prediction rules and the more sophisticated the rules, the greater the predictive ability. Two products currently exist, MetabolExpert[18] (Compu-Drug, Budapest, Hungary) and META (Ref. 19; Multicase Inc., Beachwood, OH, USA) and it is commendable that the foresight existed for these products to be developed.

Predictive databases have found some uses and, naturally, some limitations have been encountered. For prediction systems of this type, together with the developmental stage of the respective scientific understanding of the processes involved, it is not a question of whether the system will fail but when. These databases need to evolve with a continuous process of refinement as the knowledge base expands

allowing the development of increasingly sophisticated rules. It is obvious that knowledge and predictive databases need to proceed in parallel. New rules will be necessary when novel pathways are discovered such as routes involving cleavage of heterocyclic ring systems. There are now initiatives to develop a new generation of products using the larger knowledge base that now exists[20].

It is a major challenge to find innovative ways of defining rules for metabolism. For example, sufficient information exists to define rules for the most likely positions of hydroxylation in substituted aromatic rings or the metabolites of a particular heterocyclic ring system, but this does not address the question of whether these processes will occur. The structures of new active molecules have become more complex and contain combinations of diverse functional groups and sub-structures. Hence, there are various options available for metabolism and rules need to be able to establish a priority order for phase I metabolism and incorporate the intervention of phase II conjugation, which would remove availability of metabolites for further metabolism.

The current programmes generate metabolic trees with primary metabolites being considered as parents for conversion to secondary metabolites. The META programme will halt when a prescribed level of lipophilicity has been reached on the basis that these metabolites will not be substrates for lipophilic active sites in enzymes and will also be readily excreted. MetabolExpert excludes phase II metabolites from further reaction. However, this might be too simplistic, as it would exclude the array of subsequent metabolites produced from a glutathione conjugation pathway, whereby the very hydrophilic conjugate is subject to a series of phase I-type transformations. META incorporates a stability module check that recognizes unstable atom arrangements and amends a structure until a stable entity is reached. A transformation can result in the cleavage of a molecule resulting in two metabolites, both of which merit consideration either pharmacologically or toxicologically. MetabolExpert incorporates the provision to include both metabolites. Both programmes have included valuable features but further refinement using our increasing knowledge is required.

New directions

Prior to the past five years, there was limited knowledge concerning the precise isozyme (e.g. CYP450) involved in the formation of a specific oxidation product from given substrates. However, because of the importance of anticipating the potential for drug–drug interactions and genetic polymorphism in metabolism, it is now common for the isozymes involved to be reported alongside the identified metabolites[21,22]. The increased knowledge base of the substrates for specific isozymes has also contributed to an increased understanding of the nature and binding requirements of their active sites. This knowledge could provide another approach to the prediction of metabolites taking account of physicochemical parameters such as lipophilicity, electrostatic features and conformation including size, shape and steric features. The structure of oxidized metabolites can be predicted by assessing the ability of a molecule to bind to a defined site based on established parameters and considering the part of the molecule that aligns with the enzyme active site, namely the haem ferric ion. Stereoselective metabolism of enantiomers is well-known and the modelling approach could allow prediction of differential metabolism. A racemic drug should be considered as a mixture of two compounds as, increasingly, single enantiomers have been selected as development candidates. Any information to aid the selection of one enantiomer at an early stage would be valuable. For instance, the differential metabolism of *R*- and *S*-mephenytoin has been rationalized by demonstrating their interaction with the active sites of the isozymes CYP2C9 and CYP2C19, respectively[23].

There is now considerable potential for developing more sophisticated predictive databases. A further extension of this process would be to provide a link to toxicology databases with the aim of giving predictive alerts to various toxicological endpoints using structure–activity rules. A limitation of the current stand-alone systems is that the contribution of toxic reactive intermediates or toxic metabolites is not taken into consideration. This would be another tool to use during lead compound optimization and selection, and would provide feedback to enable structural modifications.

REFERENCES

1 Snyder, R.W. and Grethe, G. (1999) in *Metabolism Databases and High-Throughput Testing During Drug Design and Development* (Erhardt, P.W., ed.), pp. 277–280, International Union of Pure & Applied Chemistry and Blackwell Science

2 Hayward, J. (1999) in *Metabolism Databases and High-Throughput Testing During Drug Design and Development* (Erhardt, P.W., ed.), pp. 281–288, International Union of Pure & Applied Chemistry and Blackwell Science

3 Stiff, D.D., Robicheau, J.T. and Zemaitis, M.A. (1992) *Xenobiotica* 22, 1–11

4 Stiff, D.D. and Zemaitis, M.A. (1990) *Drug Metab. Dispos.* 18, 888–894

5 Lee, M.S. and Yost, R.A. (1988) *Biomed. Environ. Mass Spectrom.* 15, 193–204

6 Meuldermans, W. *et al.* (1994) *Drug Metab. Dispos.* 22, 129–138
7 Mannens, G. *et al.* (1993) *Drug Metab. Dispos.* 21, 1134–1141
8 Mutlib, A.E., Strupezewski, J.T. and Chesson, S.M. (1995) *Drug Metab. Dispos.* 23, 951–964
9 Speed, W. *et al.* (1994) *Biol. Mass Spectrom.* 23, 1–5
10 Yabuki, M. *et al.* (1993) *Drug Metab. Dispos.* 21, 1167–1169
11 Dixon, C.M. *et al.* (1993) *Drug Metab. Dispos.* 21, 761–769
12 Tse, F.L.S. *et al.* (1987) *Xenobiotica* 17, 1259–1267
13 Guillouzo, A. *et al.* (1988) *Xenobiotica* 18, 131–139
14 Balani, S.K. *et al.* (1993) *Drug Metab. Dispos.* 21, 598–604
15 Vickers, A.E.M. *et al.* (1996) *Eur. J. Drug Metab. Pharmacokinet.* 21, 43–50
16 Fischer, V. *et al.* (1994) *Drug Metab. Dispos.* 22, 269–274
17 Reith, M.K., Sproles, G.D. and Cheng, L.K. (1995) *Drug Metab. Dispos.* 23, 806–812
18 Darvas, F. *et al.* (1999) in *Metabolism Databases and High-Throughput Testing During Drug Design and Development* (Erhardt, P.W., ed.), pp. 237–270, International Union of Pure & Applied Chemistry and Blackwell Science
19 Klopman, G. and Tu, M. (1999) in *Metabolism Databases and High-Throughput Testing During Drug Design and Development* (Erhardt, P.W., ed.), pp. 271–276, International Union of Pure & Applied Chemistry and Blackwell Science
20 Greene, N. (1999) in *Metabolism Databases and High-Throughput Testing During Drug Design and Development* (Erhardt, P.W., ed.), pp. 289–296, International Union of Pure & Applied Chemistry and Blackwell Science
21 Spatzenegger, M. and Jaeger, W. (1995) *Drug Metab. Rev.* 270, 397–417
22 Gibson, G.G. (ed.) (1998) *Xenobiotica* 28, 1095–1273
23 Lewis, D.F.V. *et al.* (1998) *Xenobiotica* 28, 235–268

In short...

Cerep (Paris, France) and **Bristol-Myers Squibb** (BMS; Princeton, NJ, USA) have entered into a strategic collaboration that comprises a subscription for BMS to Cerep's proprietary database, BioPrint™ and a collaborative research and drug discovery program. This program will enable the use of Cerep's molecular modelling, library design and syntheses, high-throughput screening, profiling and compound optimization whilst BMS will be responsible for clinical development. In return, Cerep get access to the scientific staff, research equipment and facilities of UPSA, a wholly owned French subsidiary of BMS.