METABOLISM

Special Report

New IUPAC Book Examines Metabolism Databases

Can software systems and databases really assist in the development of new drugs? The International Union of Pure and Applied Chemistry (IUPAC) recently initiated a working party to explore this question. The working party specifically examined metabolism databases and collected case studies and contributions from academics and commercial vendors in a new book from Blackwell Science Limited, *Drug Metabolism: Databases and High-Throughput Testing During Drug Design and Development*.

The book is intended as a resource for discovery researchers, according to the book's editor, Paul Erhardt, Ph.D., who directs the Center for Drug Design and Development at the College of Pharmacy at the University of Toledo in Ohio. Two scientists from MDL, Bob Snyder and Guenter Grethe, put Metabolite through its paces as part of a chapter that shows how various metabolism systems tackle the same drug design task. "For metabolism information to be useful, the data must have statistical relevance," Erhardt said. "Large databases provide this relevance. I commend MDL for putting together such a huge database and for making it easy to access and integrate with other scientific applications."

Drug Metabolism also includes several chapters of "case studies." In these chapters, metabolism scientists describe some of the applications that they have developed to make drug metabolism data easier to access.

> "Metabolite and its Registrar are easy tools to use which can assist a scientist in drug discovery and development since they can give a user a combination of the published data along with any in-house drug metabolism data."
>
> ***Wolfgang Kanhai,*** Bayer A.G.

Their challenges are strikingly similar—in each case, the relative youth of the technology combined with the tendency of scientists to rely on intuition over archived information makes implementing effective metabolism systems difficult.

For instance, Mark Johnson in the computer-assisted drug discovery group at Pharmacia & Upjohn, Inc. writes that his company's Metabolite database is used less frequently than chemical reaction databases. He attributes the scientists' apathy to the inherent complexity of metabolic data and the culture surrounding its use. "One's expectations regarding the metabolic fate of a molecule are governed by many factors…. It all makes for a complicated issue, and some of my collaborators have asked the following question: Can we improve those expectations by systematically analyzing the data in a suitably defined metabolic database?" Through a series of pilot studies, Johnson concludes that "metabolic databases can be used to provide statistically-based prediction rules that begin to add a measure of objective quantification to our expectations concerning the metabolic fate of a molecule."

The case studies show that Metabolite's particular value lies in its ability to be queried by chemical structure as well as by other traditional text and data queries. Chapter authors also appreciate the capabilities of the Metabolite Registrar, which lets scientists create and store their own metabolic schemes. "Metabolite and its Registrar are easy tools to use which can assist a scientist in drug discovery and development since they can give a user a combination of the published data along

continued on next page

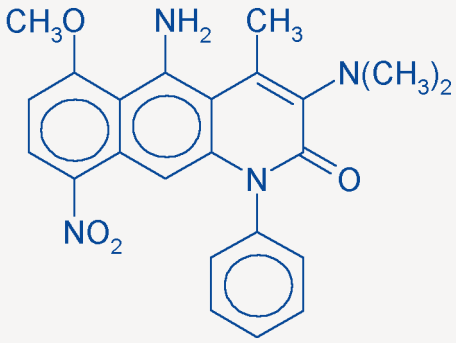

Figure 1: *Query molecule* ***1****.*

"Metabolic databases can be used to provide statistically-based prediction rules that begin to add a measure of objective quantification to our expectations concerning the metabolic fate of a molecule."

Mark Johnson, Pharmacia & Upjohn, Inc.

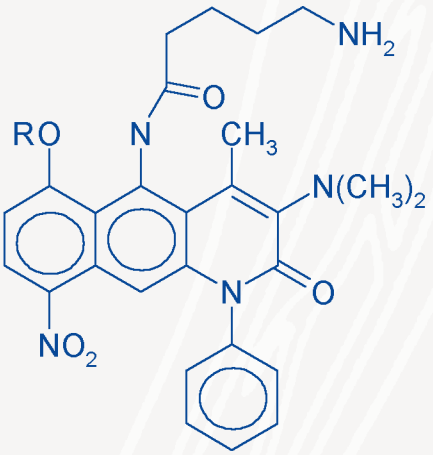

Figure 2: *Potential prodrug (R=CH3) for target molecule* ***1****. Another possibility, although probably much more remote, would be the phenolic precursor where R=H.*

with any in-house drug metabolism data," writes Wolfgang Kanhai of Bayer A.G.

Erhardt hopes that his book will inspire further interest in making metabolism information more widely available to researchers around the world. "The one thing that was abundantly clear to me at the book's completion was that the metabolism field needs a publicly available human metabolism database," Erhardt said. "Such a system would both complement and validate the information contained in commercial and proprietary systems."

The Center for Drug Design and Development has already begun collaborating with the International Union of Pharmacology to collect metabolism information from worldwide sources, including underrepresented populations. "We are simply trying to advance this field," Erhardt said. "To the extent that I can make this happen or see it happen in somebody else's hands, I will be very pleased."

The portion of the textbook illustrating Metabolite's performance has been excerpted below with permission from Erhardt, Blackwell Science, and IUPAC. The discussion of the problem, along with the summary, are the author's. The Metabolite discussion has been edited for length and some of the structures have also been edited for clarity. A full description of the searches conducted and the results can be found in the text. To order a copy of *Drug Metabolism*, visit the IUPAC Web site at *http://www.iupac.org* and navigate to their latest publications.

The Problem

To explore and compare their versatility with regard to the design of prodrugs and codrugs, each of the available database vendors was contacted to examine the same query molecule **1** (see Figure 1) relative to the following set of three questions:

1) What metabolic reactions might be relied upon to produce the query molecule from any type of prodrug precursor?

2) What metabolic reactions might be expected to occur on the query molecule itself?

3) Can anything be indicated about the associated rates for metabolic production of the query molecule versus its metabolic clearance?

Some Answers

To identify potential prodrugs of **1** the 'Metabolite Browser' was employed (1). This program allows searching the database

for molecules that contain the desired features in the metabolite but not all of them in the substrate. [Several substructure searches for potential modifications were performed, selecting only those that were indicated in the database as prodrug transformations showed] that compound **2** (see Figure 2) could potentially be an effective prodrug for the target molecule **1**.

Compound **1** was analyzed for possible metabolic fate by selecting structural fragments of the target and searching the database for examples in which the fragment has been metabolized. The target compound was broken into five (5) fragments for analysis. These queries are depicted in Figure 3.

A composite of the results from all of these fragment searches is summarized within Figure 4. Relative rate data between these possibilities is not available, although O- and N-dealkylation would seemingly be very important pathways.

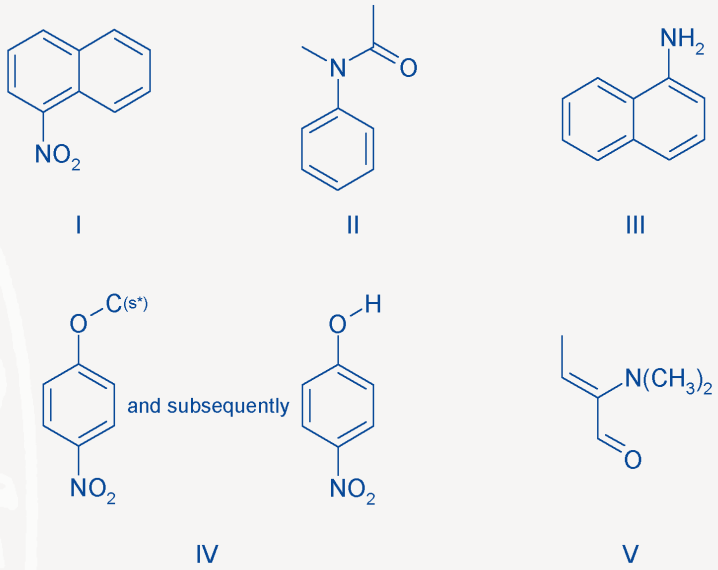

Figure 3: *Specific fragments of* ***1*** *that were searched within the Metabolite database (1). I. Represents the southwestern corner of* ***1****. II. Represents the southeastern portion of* ***1****. III. Addresses the key aromatic amino functionality located within the central region of* ***1****. IV. Addresses the northwestern corner of* ***1*** *and its immediate O-demethylated metabolite. And V. represents the northeastern corner of* ***1****.*

Summary

It is important to point out immediately that the depth of each of the answers should not be used to compare the potential value of the various databases. Even though the same query molecule and set of corresponding questions was posed to each of the corresponding database vendors, different levels of examination and rigor were then undertaken by each vendor. In general, the various results reflect a quick, first-pass examination by each of the database holders. Overall, there appears to be much more congruency of the results than disparency [sic] of the results.

Alternatively, it should be clear that the use of such databases within the setting of prodrug/codrugs has indeed revealed some possibilities beyond those which are apparent by simple visual inspection and classical medicinal chemistry logic for both target design and the eventual experimental study of metabolic consequences. Likewise, the need for **relative rate data** or ***statistically-based relative occurrence data*** (2) between the potentially competing 'metabophore' (3) pathways has been underscored ever more heavily by the model study described herein.❖

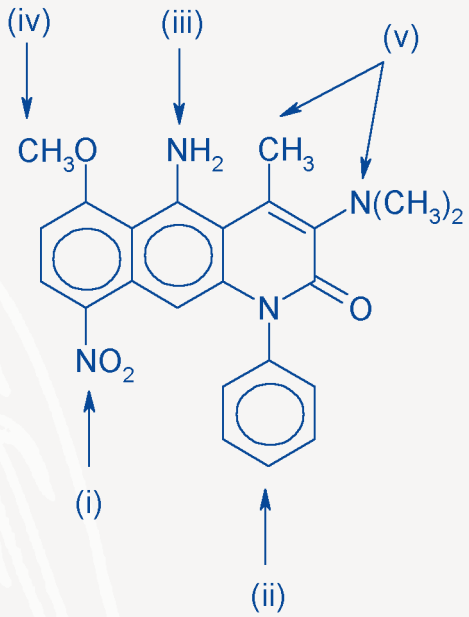

Figure 4: *Potential metabolism of* ***1*** *as assessed by Metabolite (1): (i) Formation of amino-, nitroso-, and methylnitroso-moieties with the amino-form undergoing subsequent Phase II metabolic conjugation with sulfate, glucuronic acid, and various protein and DNA binding species; (ii) Aromatic hydroxylation followed by the possibility for glucuronide and sulfate formation; (iii) Formation of N-acetyl, glucuronide, sulfate, and N-oxide metabolites as well as specific interactions involving certain protein and DNA binding species; (iv) O-demethylation followed by formation of sulfate, phosphate, and glucuronide conjugates; and, (v) Hydroxylation of the methyl group and N-demethylation, the latter followed by acetylation, urea formation or further N-dealkylation at the ring attachment position.*

References (cited in this reprint):

1) This section was contributed by G. Grethe and R. Snyder, MDL Information Systems, Inc. A detailed discussion of the searches can be found in the book on pp. 215-219. Also see the chapter entitled *Metabolite* by Snyder and Grethe.
2) For example, see the chapter entitled *Statistics-Based Probabilities Of Metabolic Possibilities* by P. Erhardt within the book.
3) P. Erhardt has adopted this term, by analogy to the term "*pharmacophore*," to describe the specific structural array associated with a given metabolic occurrence. It is also similar to G. Klopman's (4) use of the term "toxicophore" to describe the specific structural array associated with a given toxicological outcome at the biochemical level.
4) G. Klopman and H. Rosenkrantz, *Toxicity Estimation by Chemical Substructure Analysis: The Tox II Program*, Tox. Lett., **79**, 145-155 (1995).