The SMARTCyp cytochrome P450 metabolism prediction server

Patrik Rydberg, David E. Gloriam and Lars Olsen*
Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark
Associate Editor: Anna Tramontano

ABSTRACT

Summary: The SMARTCyp server is the first web application for site of metabolism prediction of cytochrome P450-mediated drug metabolism.

Availability: The SMARTCyp server is freely available for use on the web at www.farma.ku.dk/smartcyp where the SMARTCyp Java program and source code is also available for download.

The SMARTCyp cytochrome P450 metabolism prediction server

1 INTRODUCTION

The most important drug-metabolizing enzymes in humans are cytochromes P450 (CYPs). They oxidize a wide range of substrates and contribute to the metabolism of 90% of all drugs (Lewis, 2008).

While there are methods for site of metabolism prediction which are solely based on semi-empirical calculations of the substrates (Hennemann et al., 2009; Zheng et al., 2009), more accurate results should be achieved by explicitly including the reactivity of each site in a substrate. Hitherto, reactivity models for CYPs have been restricted to semi-empirical calculations of the stability of intermediates (Jones et al., 2002). Later studies have shown that more accurate activation energies for CYPs can be computed by performing very time consuming density functional theory (DFT) transition state calculations on a heme group porphyrin ring model (Olsen et al., 2006; Rydberg et al., 2008a; b; Shaik et al., 2005).

SMARTCyp, however, presents the first solution for both fast and exact prediction of CYP reactivity. Accuracy of activation energies is ensured by making very extensive DFT calculations and speed is gained by saving these in a precalculated library so that they need only be looked up at execution. SMARTCyp has been validated for the prediction of activation energies of these fragments. Rules and calculations are presented in Table 1 (overall distribution) and Supplementary Material (DFT results and SMARTS definitions).

We have analyzed the accuracy of SMARTCyp on a set of 361 drug-like CYP3A4 substrates (Table 2). The substrates were taken from the dataset in our previous work (Rydberg et al., 2010c), by extracting all molecules defined as drug-like according to the ‘lip_druglike’ descriptor in the MOE software (Chemical Computing Group Inc., version 2007.09). The results show an improved accuracy of 2 percentages compared with the original SMARTCyp implementation.

A preliminary validation against data for other CYP isoforms shows that a similar performance can be expected for the isoforms 1A2, 2A6, 2B6, 2C8, 2C19 and 2E1 (Rydberg et al., 2010a, b).

Table 1. Distribution of the transition state calculations and SMARTS rules by atom types

Table 2. Percent correctly predicted substrates

2 NEW FUNCTIONALITY AND IMPROVEMENTS TO SMARTCYP

The detailed description of the SMARTCyp algorithm, development and DFT calculations has been published elsewhere (Rydberg et al., 2010c). In this section, we highlight significant improvements as well as new features and functions in the current version 1.5.

2.1 Increased accuracy

For this version, 72 new DFT calculations were performed and added to the 139 previous to refine and create additional SMARTS rules. The library of SMARTS rules is used to look up atom activation energies in molecules by SMARTS patterns. The majority of the new calculations were performed to systematically evaluate and refine the rules for nitrogen dealkylations of peptide and acetamide groups, as well as oxidation of aromatic five-membered rings and heterocycles.

The new rules have resulted in significant improvements to the prediction of the activation energies of these fragments. Rules and calculations are presented in Table 1 (overall distribution) and Supplementary Material (DFT results and SMARTS definitions).

While there are methods for site of metabolism prediction which are solely based on semi-empirical calculations of the substrates (Hennemann et al., 2009; Zheng et al., 2009), more accurate results should be achieved by explicitly including the reactivity of each site in a substrate. Hitherto, reactivity models for CYPs have been restricted to semi-empirical calculations of the stability of intermediates (Jones et al., 2002). Later studies have shown that more accurate activation energies for CYPs can be computed by performing very time consuming density functional theory (DFT) transition state calculations on a heme group porphyrin ring model (Olsen et al., 2006; Rydberg et al., 2008a; b; Shaik et al., 2005).

SMARTCyp, however, presents the first solution for both fast and exact prediction of CYP reactivity. Accuracy of activation energies is ensured by making very extensive DFT calculations and speed is gained by saving these in a precalculated library so that they need only be looked up at execution. SMARTCyp has been validated for the prediction of metabolism by the 3A4 isoform, which contributes to the metabolism of half of all drugs (Guengerich, 1999). SMARTCyp has also been shown to work well for six other CYP isoforms (Rydberg et al., 2010a, b).

In this application note, we describe the SMARTCyp server, the first web service for prediction of CYP-mediated metabolism.

Table 1. Distribution of the transition state calculations and SMARTS rules by atom types

Table 2. Percent correctly predicted substrates

2 NEW FUNCTIONALITY AND IMPROVEMENTS TO SMARTCYP

The detailed description of the SMARTCyp algorithm, development and DFT calculations has been published elsewhere (Rydberg et al., 2010c). In this section, we highlight significant improvements as well as new features and functions in the current version 1.5.

2.1 Increased accuracy

For this version, 72 new DFT calculations were performed and added to the 139 previous to refine and create additional SMARTS rules. The library of SMARTS rules is used to look up atom activation energies in molecules by SMARTS patterns. The majority of the new calculations were performed to systematically evaluate and refine the rules for nitrogen dealkylations of peptide and acetamide groups, as well as oxidation of aromatic five-membered rings and heterocycles.

The new rules have resulted in significant improvements to the prediction of the activation energies of these fragments. Rules and calculations are presented in Table 1 (overall distribution) and Supplementary Material (DFT results and SMARTS definitions).

While there are methods for site of metabolism prediction which are solely based on semi-empirical calculations of the substrates (Hennemann et al., 2009; Zheng et al., 2009), more accurate results should be achieved by explicitly including the reactivity of each site in a substrate. Hitherto, reactivity models for CYPs have been restricted to semi-empirical calculations of the stability of intermediates (Jones et al., 2002). Later studies have shown that more accurate activation energies for CYPs can be computed by performing very time consuming density functional theory (DFT) transition state calculations on a heme group porphyrin ring model (Olsen et al., 2006; Rydberg et al., 2008a; b; Shaik et al., 2005).

SMARTCyp, however, presents the first solution for both fast and exact prediction of CYP reactivity. Accuracy of activation energies is ensured by making very extensive DFT calculations and speed is gained by saving these in a precalculated library so that they need only be looked up at execution. SMARTCyp has been validated for the prediction of metabolism by the 3A4 isoform, which contributes to the metabolism of half of all drugs (Guengerich, 1999). SMARTCyp has also been shown to work well for six other CYP isoforms (Rydberg et al., 2010a, b).

In this application note, we describe the SMARTCyp server, the first web service for prediction of CYP-mediated metabolism.

Table 1. Distribution of the transition state calculations and SMARTS rules by atom types

Table 2. Percent correctly predicted substrates

2 NEW FUNCTIONALITY AND IMPROVEMENTS TO SMARTCYP

The detailed description of the SMARTCyp algorithm, development and DFT calculations has been published elsewhere (Rydberg et al., 2010c). In this section, we highlight significant improvements as well as new features and functions in the current version 1.5.

2.1 Increased accuracy

For this version, 72 new DFT calculations were performed and added to the 139 previous to refine and create additional SMARTS rules. The library of SMARTS rules is used to look up atom activation energies in molecules by SMARTS patterns. The majority of the new calculations were performed to systematically evaluate and refine the rules for nitrogen dealkylations of peptide and acetamide groups, as well as oxidation of aromatic five-membered rings and heterocycles.

The new rules have resulted in significant improvements to the prediction of the activation energies of these fragments. Rules and calculations are presented in Table 1 (overall distribution) and Supplementary Material (DFT results and SMARTS definitions).

While there are methods for site of metabolism prediction which are solely based on semi-empirical calculations of the substrates (Hennemann et al., 2009; Zheng et al., 2009), more accurate results should be achieved by explicitly including the reactivity of each site in a substrate. Hitherto, reactivity models for CYPs have been restricted to semi-empirical calculations of the stability of intermediates (Jones et al., 2002). Later studies have shown that more accurate activation energies for CYPs can be computed by performing very time consuming density functional theory (DFT) transition state calculations on a heme group porphyrin ring model (Olsen et al., 2006; Rydberg et al., 2008a; b; Shaik et al., 2005).

SMARTCyp, however, presents the first solution for both fast and exact prediction of CYP reactivity. Accuracy of activation energies is ensured by making very extensive DFT calculations and speed is gained by saving these in a precalculated library so that they need only be looked up at execution. SMARTCyp has been validated for the prediction of metabolism by the 3A4 isoform, which contributes to the metabolism of half of all drugs (Guengerich, 1999). SMARTCyp has also been shown to work well for six other CYP isoforms (Rydberg et al., 2010a, b).

In this application note, we describe the SMARTCyp server, the first web service for prediction of CYP-mediated metabolism.
The SMARTCyp web server uses PHP code to run SMARTCyp.

The SMARTCyp server offers the user three ways to submit molecules: upload, draw a molecule, or enter SMILES strings representing molecules. The user can upload a file in any standard format, enter SMILES strings, or draw a molecule.

The results are displayed directly in the browser and include the molecular structure and an atom ranking table for each molecule as shown in Figure 1. The three top-ranked atoms are highlighted both in the structure and the table. Furthermore, all atom numbers can be shown in Figure 1. The three top-ranked atoms are highlighted both in the structure and the table. Furthermore, all atom numbers can be displayed by hovering the mouse pointer over the structure.

### 2.3 Java program output options

The command-line SMARTCyp Java program now has three optional flags, which enable the user to suppress HTML or csv output, as well as direct the output to a specific directory. Suppressing HTML and generating only csv output makes the csv output be used for integration with other softwares.

### 3 SMARTCYP WEB SERVER

#### 3.1 Interface features

The SMARTCyp server offers the user three ways to submit molecules. The user can upload a file in any standard format, enter SMILES strings representing molecules or draw a molecule.

The results are displayed directly in the browser and include the molecular structure and an atom ranking table for each molecule as shown in Figure 1. The three top-ranked atoms are highlighted both in the structure and the table. Furthermore, all atom numbers can be displayed by hovering the mouse pointer over the structure.

#### 3.2 Implementation

The SMARTCyp web server uses PHP code to run SMARTCyp and the interface functionality. To support all standard formats, uploaded files are converted by Open Babel (Guha et al., 2006) when necessary. SMARTCyp is implemented using the CDK Java library (Steinbeck et al., 2003, 2006).

**Funding**: Alfred Benzon foundation (to P.R. and D.E.G.); Danish Medical Research Council (to L.O.).

**Conflict of Interest**: none declared.

**REFERENCES**


Rydberg, P. et al. (2010a) SMARTCyp: a 2D method for prediction of cytochrome P450-mediated drug metabolism. In 5th Joint Sheffield Conference on Cheminformatics, Sheffield, UK.


