

pkCSM: predicting small-molecule pharmacokinetic properties using graph-based signatures

Douglas E. V. Pires*, Tom L. Blundell and David B. Ascher

*Correspondence: dpires@dcc.ufmg.br

Help - How to use pkCSM

Main page

pkCSM: small-molecule pharmacokinetics prediction and optimization with graph-based signatures

Douglas E. V. Pires, Tom L. Blundell, David B. Ascher

Abstract

Modern high throughput drug discovery approaches have increased the numbers of lead compounds being identified, and in shorter time frames than traditional medicinal chemistry, however many of these promising compounds often fail because of unsatisfactory ADMET properties. In silico screening approaches help to reduce these risks. Here we propose a novel approach to the prediction of pharmacokinetic properties, called pkCSM, which relies on graph-based signatures. These encode distance patterns between atoms and are used to represent the small molecule and to train predictive models.

The pkCSM signatures were successfully used across five main different pharmacokinetic properties classes to develop predictive regression and classification models. We show that pkCSM performs as well or better across different pharmacokinetic properties than other freely available methods. Here we present a web server to provide an integrated freely available platform to rapidly screen multiple pharmacokinetic properties.

Available Resources

- pkCSM: Small-molecule pharmacokinetics prediction
- Related: Other resources

About pkCSM

pkCSM is a machine-learning platform to predict small-molecule pharmacokinetic properties, which relies on distance/pharmacophore patterns encoded as graph-based signatures. The platform is composed of 22 regression and classification models, trained and tested on different experimental data sets encompassing a diverse and complementary set of ADMET descriptors as follows:

- **Absorption:** Caco-2 permeability, water solubility, intestinal absorption (human), P-glycoprotein substrate and inhibitor.
- **Distribution:** Volume of distribution (human), fraction unbound (human) BBB and CNS permeability
- **Metabolism:** Cytochrome P450 inhibitor and substrate
- **Excretion:** Renal OCT2 substrate
- **Toxicity:** Rat LD50, AMES toxicity, *T.Pyriformis* toxicity and Minnow toxicity

More information about the predictive models and **how to interpret the pkCSM predictions** can be accessed via the Theory menu (1).

Help - How to use pkCSM

Submission page

Pharmacokinetic properties

Step 1: Please provide a set of molecules (SMILES format)

Upload your SMILES file: No file selected

OR

Provide a SMILES string:

Step 2: Please choose the prediction mode

Prediction of pharmacokinetic properties

Disclaimer

No molecule information will be retained on the system after being uploaded by the user.

How to run a prediction

To run a prediction:

- Click on "Prediction" (1) to open the submission page.
- Provide either an input file with a list of molecules (2) in SMILES format (up to a maximum of 1000 molecules) or supply a single SMILES string (3). Users are advised to use Canonical SMILES. Syntax non-compliant molecules will be ignored.
- The next step is to choose the prediction mode (4). Users can choose between the main ADMET property classes (**A**bsorption, **D**istribution, **M**etabolism, **E**xcretion and **T**oxicity) by clicking on its corresponding button or run a systematic evaluation of all predictive models.

Help - How to use pkCSM

Results page

The screenshot shows the 'Pharmacokinetic Properties' section of the pkCSM interface. It includes a navigation bar at the top with 'pkCSM' and 'Predict' buttons, and links for 'Theory', 'Help', 'Contact', 'Acknowledgements', and 'Related Resources'. The main content is divided into three numbered callouts: (1) 'Molecule Depiction' showing a chemical structure of a molecule; (2) 'Molecule properties:' listing various descriptors and their values; and (3) a table of predicted values for different models. A 'Run another prediction' button is also visible.

Property	Model Name	Predicted Value	Unit
Absorption	Caco2	1.09	Numeric (log cm/s)
Absorption	Water solubility	-4.906	Numeric (log mol/L)
Absorption	Intestinal absorption (human)	91.125	Numeric (% Absorbed)
Absorption	P-glycoprotein substrate	No	Categorical (Yes/No)
Absorption	P-glycoprotein I inhibitor	Yes	Categorical (Yes/No)
Absorption	P-glycoprotein II inhibitor	Yes	Categorical (Yes/No)

Molecule properties:

Descriptor	Value
Molecular Weight	375.871
LogP	4.3172
#Rotatable Bonds	6
#Acceptors	3
#Donors	1
Surface Area	157.952

Results

For a **single molecule** prediction, your results can be displayed as follows:

- A **depiction** of the uploaded molecule will be shown in (1). Make sure the depiction is what you are expecting for your molecule.
- A list of **molecule properties** will be calculated and shown in (2).
- The prediction will be displayed in a tabular format as presented in (3). The information shown include the **ADMET property** being predicted, **model name**, the actual **predicted value** and whether the prediction is **numeric** (for regression models, including the unit of the predicted value) or **categorical** (for classification models).

Results for **multiple molecules** will be shown in a tabular format, without molecule depiction, which can be downloaded as a CSV file.

Help - How to use pkCSM

Contact page

The screenshot shows the 'Contact' page of the pkCSM interface. It features a navigation bar at the top with 'pkCSM' and 'Predict' buttons, and links for 'Theory', 'Help', 'Contact', 'Acknowledgements', and 'Related Resources'. The main content is titled 'Contact' and includes a 3D character holding an envelope icon. There are two main sections: 'Mailing address' and 'Get in touch'. The 'Mailing address' section provides contact information for Dr. Douglas E. V. Pires. The 'Get in touch' section includes a form for reporting problems or requests, with fields for 'Name' and 'Email address' (optional), and a 'Submit' button.

Mailing address

[Dr. Douglas E. V. Pires](#)
University of Cambridge
Department of Biochemistry
80 Tennis Court Road
Cambridge UK
CB2 1GA

E-mail addresses

dpires@doc.umg.br
dep2@cam.ac.uk

Get in touch

Have you come across a problem on the website or have any requests or suggestions? Please report it here!

Name Email address

Getting in touch

In case you experience any trouble using pkCSM or have any suggestions or comments, please do not hesitate in contacting us (1) either via e-mail or through the online form.