

# KBbox: A Toolbox of Computational Methods for Studying the Kinetics of Molecular Binding

Neil J. Bruce,<sup>\*,†,‡</sup> Gaurav K. Ganotra,<sup>†,‡,§</sup> Stefan Richter,<sup>†</sup> and Rebecca C. Wade<sup>\*,†,§,||</sup>

<sup>†</sup>Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany

<sup>‡</sup>Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp), Im Neuenheimer Feld 205, 69120 Heidelberg, Germany

<sup>§</sup>Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany

<sup>||</sup>Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany

**ABSTRACT:** The past few years have seen increasing recognition of the importance of understanding molecular binding kinetics. This has led to the development of myriad computational methods for studying the kinetics of binding processes and predicting their associated rate constants that show varying ranges of application, degrees of accuracy, and computational requirements. In order to help researchers decide which method might be suitable for their projects, we have developed KBbox, a web server that guides users in choosing the methods they should consider on the basis of the information they wish to obtain, the data they currently have available, and the computational resources to which they have access. KBbox provides information on the toolbox of available methods, their associated software tools, an expanding list of curated examples of published applications, and tutorials explaining how to apply some of the methods. It has been designed to allow the easy addition of new methods, tools, and examples as they are developed and published. KBbox is available at <https://kbbox.h-its.org/>.



→ Association Rates  
→ Residence Times  
→ Binding Pathways  
→ Pharmacokinetics

## INTRODUCTION

In recent years there has been increasing interest in understanding the processes by which molecular binding partners associate and dissociate and in predicting the rates at which these processes occur. This is particularly true in pharmaceutical research, where the *in vivo* pharmacological activity of candidate drug compounds is increasingly seen to correlate better with target binding kinetics than target affinity.<sup>1</sup> As a result of this interest, the Kinetics for Drug Discovery (K4DD) Project began in 2012 (<https://www.k4dd.eu>), bringing together researchers from the pharmaceutical industry and academia with the goal of creating new strategies for optimizing drug binding kinetics.<sup>2</sup> Molecular binding kinetics is also important in systems biology modeling, where kinetic data are combined with omics and molecular network data in order to understand and model processes such as cellular signaling and metabolism.

While molecular simulation techniques are well-placed to offer insights into the kinetics of molecular binding, their application is hindered by the requirement for accurate modeling of the transition states of the binding and unbinding processes. As these are by definition high-energy states, they are difficult to characterize experimentally and are poorly sampled in molecular simulations. This has led to the development of a large number of advanced sampling simulation approaches that seek to accelerate the crossing of

these transition state barriers and quantitative structure–kinetics relationship (QSKR) methods that aim to relate kinetics to the structures of macromolecular targets, small molecules, or the complexes they form.<sup>3–8</sup>

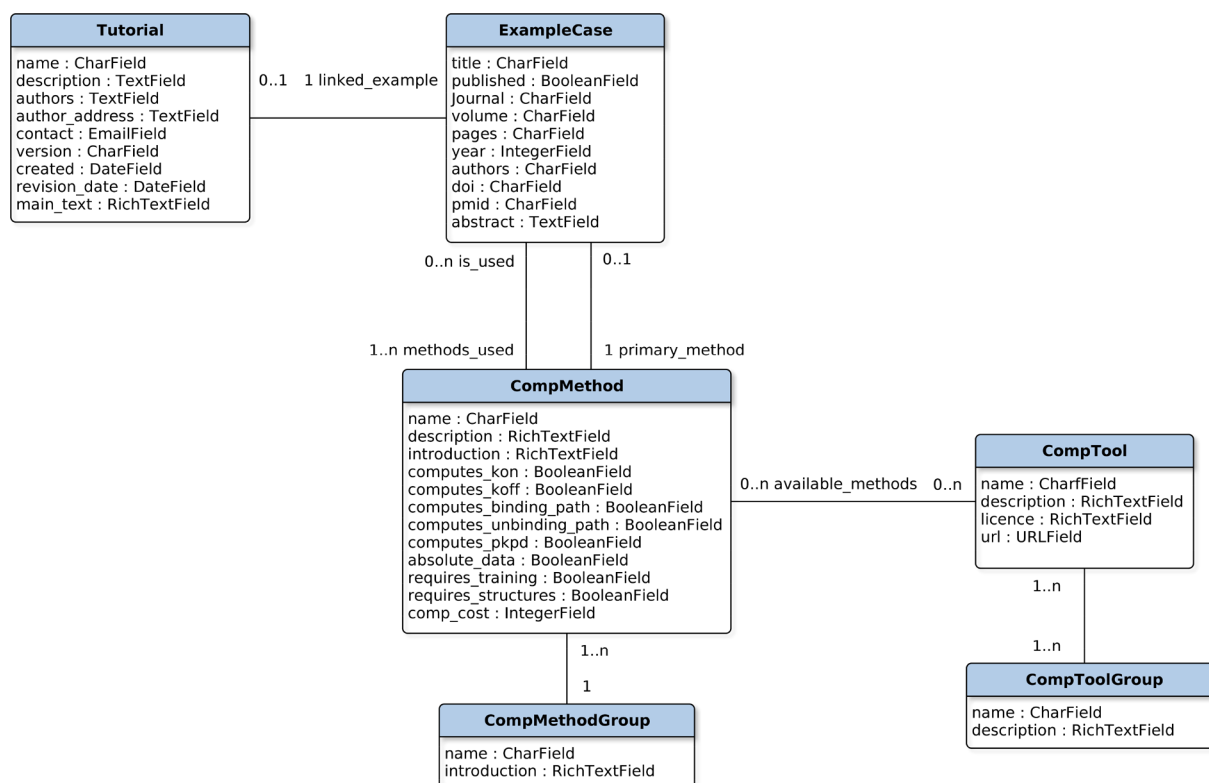
As the number of computational methods for studying binding kinetics increases, it becomes increasingly difficult for researchers to keep up to date with the new methods being developed. It is also difficult for newcomers to the field to decide which methods may be suitable to apply to address their research questions, as different methods have cost/accuracy trade-offs and different requirements in terms of the prior knowledge needed for their successful application. To help solve this problem, we here introduce KBbox, a web server designed to guide users in choosing the correct method for their studies while also providing information on the various methods along with a curated list of examples of previous applications.

## TECHNICAL DESCRIPTION

**Features.** KBbox provides a guide to the toolbox of computational methods that can be used to study molecular binding kinetics. The web server is easily extendable, allowing new methods to be added to the toolbox as they are developed

**Received:** June 12, 2019

**Published:** August 5, 2019



**Figure 1.** Entity relationship diagram showing the data representation used in KBbox.

and published. KBbox also provides a curated list of published applications of the methods, providing users with an easy-to-find reference list. For a number of methods, tutorials that give the user an introduction to how to run the calculations are also provided. For users who need advice on choosing which method to use in their projects, KBbox provides a query interface that asks a series of questions relating to the data they have available and the data they wish to calculate and provides them with a list of methods found in the toolbox that match their query, sorted approximately by the computational resources required for their application. Finally, KBbox provides a list of computational tools that can be used in studies of binding kinetics, along with links pointing the user to the locations where they can obtain the software. All of the information available on the web server is searchable via a Google search interface.

**Implementation.** KBbox is implemented in Python 3 using the Django Framework version 2.2<sup>9</sup> for serving dynamic HTML content created from information stored in its database (see [Database Structure](#)). The responsive web interface is provided through the use of the Bootstrap 3 css framework.<sup>10</sup> The Entrez module from the Biopython package<sup>11</sup> is used to populate the database with citation data and abstracts for previously published applications taken from PubMed, with the data parsed using the BeautifulSoup parser. The webpage content is created using the CKEditor rich text editor in the Django-CKEditor Python package.

**Database Structure.** The data structure used in KBbox is shown in [Figure 1](#). The main data table contains the computational methods that constitute the toolbox (class `CompMethod`). Each entry in the table contains a short summary description of the method and a more detailed introduction to the method. It also contains a number of

parameters that are used in querying the methods in the toolbox to find which methods match the user's needs. These relate to the data that the user wishes to obtain (binding or unbinding pathway sampling, association rate constants ( $k_{on}$ ), dissociation rate constants ( $k_{off}$ ), pharmacokinetic/pharmacodynamics (PKPD) predictions), whether training data or atomic-detail structural data are required by the method, and whether the method is able to provide absolute or only relative kinetic data. Finally, the table contains an integer describing the approximate computational cost of the method, ranging from 1 (least expensive) to 5 (most expensive). A broad five-point score was chosen, as the overall computational cost of a method can be highly application-dependent and some methods lend themselves more easily to parallelization or GPU acceleration. Therefore, the cost associated with each score should be seen as the cost of a typical use case on today's typical CPU-based hardware. Methods that can be run in a few minutes per compound on a single core of a desktop computer are assigned a value of 1, while those requiring several hours on a desktop computer are assigned a value of 2. Values in the range 3–5 correspond to longer simulation-based methods that require longer compute times and/or parallel processing. A value of 3 is assigned where a method can be run on a small cluster (up to ca. 200 cores) in within 1–2 days per compound, while a value of 4 suggests that the method would require 1–2 weeks on a similar-sized cluster. Methods that require longer than this are given a cost value of 5.

Each method is sorted into a particular group (class `CompMethodGroup`) to enable KBbox to organize the methods into different classes, allowing users to search for methods more easily. Currently, these groups are molecular modeling, molecular simulation, PKPD modeling, and QSKR approaches.



**Figure 2.** Screenshots illustrating two of the example use cases of KBbox. In the first use case (cyan boxes, counterclockwise), upon arriving at the KBbox homepage (black box), the user is guided in choosing a method that meets the indicated needs on the basis of a series of questions. The user can find information describing the suggested methods and is presented with examples of previously published work employing these methods, tutorials that give introductions to the methods' application, and a list of software packages where the method has been implemented. In the second use case (magenta boxes, clockwise), the user is interested in finding out more information about the use of a specific molecular simulation method, for example metadynamics, for studying binding kinetics. The KBbox logo is reproduced with permission. Copyright 2016–2019 Heidelberg Institute for Theoretical Studies gGmbH.

Each method is linked to one or more examples of previously published research or a report of currently

unpublished data in which that method was either the primary method used or one of a set of methods applied (class



ExampleCase). Each row in the table describing these examples contains author information and a flag indicating whether the work is published or unpublished. For published examples, citation data are also provided. For some of the examples, tutorials are also provided, and these are recorded in an additional table (class Tutorial).

A separate table is used to populate the list of computational tools described by KBbox (class CompTool). Each row of this table contains a description of the tool, its license, and a URL to its web page. Each tool is also linked to the various computational methods (class CompMethod) that can be employed using that tool. In a similar manner to the methods, the tools are sorted into groups (class CompToolGroup). Currently, these are data analysis tools, PKPD modeling tools, preparation and general modeling tools, simulation tools, and structure visualization tools. A particular tool can be a member of more than one group.

## ■ EXAMPLE USE CASES

KBbox is intended to be useful to a wide range of users with different needs and experiences. Here we outline three potential use cases of the KBbox web server.

**What Method Should I Use for a Given Project?** A Ph.D. student with some basic experience in computational modeling and simulation is starting a new project. The student wants to use computational modeling to predict which of a set of compounds might show a long residence time for a given protein target and has crystallographic data for these compounds bound to the target. For some of the compounds, experimentally derived residence times for their target are also available. The student arrives at the KBbox home page (Figure 2, black box), clicks on the button “Not sure what method to use for a particular problem? Click here!”, and is then given advice and information based on the indicated needs (Figure 2, cyan boxes, counterclockwise). First, the user is asked a series of questions relating to the data available and the data to be obtained. After these questions are answered, KBbox checks the database and presents a list of methods that match the query, sorted approximately by the computational resources required for the calculation. The student clicks on each entry in this list and is taken to a page that gives an overview of the particular method along with a curated list of examples of previous applications of that method, with links to the relevant journal articles, and a list of computational tools where that method has been implemented. The student selects  $\tau$ -RAMD<sup>12</sup> as the method of interest and follows the link to the tutorial that describes how to perform such a calculation.

**Where Can I Find Information on Previous Applications of a Particular Method for Studying Kinetics?** A postdoctoral researcher is interested in using metadynamics<sup>13,14</sup> to study the binding and unbinding of a compound from a target protein. This researcher is already aware of the method but does not have experience with running these types of simulations and is looking for more information on the method and examples of previously published research using metadynamics. The postdoc arrives at the KBbox homepage (Figure 2, black box), clicks on the “Methods” button at the top of the page, and then selects “Molecular Simulation” from the menu on the left-hand side of the page and finds the entry for “Metadynamics” (Figure 2, magenta boxes, clockwise). The next page presents an overview of the method and a list of published examples of its application to binding kinetics

studies. The researcher follows the links for the examples and is taken to the relevant journal pages.

**Where Can I Find Information on the Computational Tools Available for My Research?** A researcher is interested in running molecular simulations to study binding kinetics and wants more information on what software tools are available. The researcher arrives on the KBbox homepage and clicks on the “Tools” button. From here, the user selects “Simulation Tools” from the menu on the left-hand side of the page, which expands to give a list of tools. Clicking on the different tool names takes the researcher to pages giving overviews of the software along with links to the relevant web pages.

## ■ CONCLUSIONS AND OUTLOOK

KBbox provides an important resource for researchers studying molecular binding kinetics, guiding users in choosing which computational methods are suitable for a given task. It is also an important source of information on the various methods and computational tools and their application. KBbox is an active web server that will continue to be developed. As such, we will update the web server with new examples of binding kinetics research as they are published and with new methods, software tools, and tutorials as they are developed. We also encourage the community to contact us with submissions for novel methods, software, or new tutorials describing how they can be used. This can be done via the “Contact” section of KBbox, where templates are available for providing data, or via email (mcmsoft@h-its.org).

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: [Neil.Bruce@h-its.org](mailto:Neil.Bruce@h-its.org).

\*E-mail: [Rebecca.Wade@h-its.org](mailto:Rebecca.Wade@h-its.org).

### ORCID

Neil J. Bruce: 0000-0002-2735-3215

Gaurav K. Ganotra: 0000-0002-1791-1113

Rebecca C. Wade: 0000-0001-5951-8670

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank Daria Kokh, Yin-Cheong Wong, and Lars Richter for providing some of the initial tutorials in KBbox and other members of the K4DD Consortium for helpful discussions during the development of KBbox. This work was supported by the EU/EFPIA Innovative Medicines Initiative (IMI) Joint Undertaking K4DD (Grant 115366), the European Union's Horizon 2020 Framework Programme for Research and Innovation under Grant Agreements 720270 (Human Brain Project SGA1) and 785907 (Human Brain Project SGA2), and the Klaus Tschira Foundation. This paper reflects only the authors' views, and neither the IMI nor the European Commission is liable for any use that may be made of the information contained herein.

## ■ REFERENCES

- (1) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug–Target Residence Time and Its Implications for Lead Optimization. *Nat. Rev. Drug Discovery* **2006**, *5*, 730–739.
- (2) Schuetz, D. A.; de Witte, W. E. A.; Wong, Y. C.; Knasmueller, B.; Richter, L.; Kokh, D. B.; Sadiq, S. K.; Bosma, R.; Nederpelt, L.; Heitman, L. H.; Segala, E.; Amaral, M.; Guo, D.; Andres, D.; Georgi,

V.; Stoddart, L. A.; Hill, S.; Cooke, R. M.; De Graaf, C.; Leurs, R.; Frech, M.; Wade, R. C.; de Lange, E. C. M.; Ijzerman, A. P.; Müller-Fahrnow, A.; Ecker, G. F. Kinetics for Drug Discovery: An Industry-Driven Effort to Target Drug Residence Time. *Drug Discovery Today* **2017**, *22*, 896–911.

(3) Patel, H. M.; Noolvi, M. N.; Sharma, P.; Jaiswal, V.; Bansal, S.; Lohan, S.; Kumar, S. S.; Abbot, V.; Dhiman, S.; Bhardwaj, V. Quantitative Structure-Activity Relationship (QSAR) Studies as Strategic Approach in Drug Discovery. *Med. Chem. Res.* **2014**, *23*, 4991–5007.

(4) Lu, H.; Iuliano, J. N.; Tonge, P. J. Structure–Kinetic Relationships That Control the Residence Time of Drug–Target Complexes: Insights from Molecular Structure and Dynamics. *Curr. Opin. Chem. Biol.* **2018**, *44*, 101–109.

(5) De Benedetti, P. G.; Fanelli, F. Computational Modeling Approaches to Quantitative Structure–Binding Kinetics Relationships in Drug Discovery. *Drug Discovery Today* **2018**, *23*, 1396–1406.

(6) Bruce, N. J.; Ganotra, G. K.; Kokh, D. B.; Sadiq, S. K.; Wade, R. C. New Approaches for Computing Ligand–Receptor Binding Kinetics. *Curr. Opin. Struct. Biol.* **2018**, *49*, 1–10.

(7) Ribeiro, J. M. L.; Tsai, S. T.; Pramanik, D.; Wang, Y.; Tiwary, P. Kinetics of Ligand-Protein Dissociation from All-Atom Simulations: Are We There Yet? *Biochemistry* **2019**, *58*, 156–165.

(8) Bernetti, M.; Masetti, M.; Rocchia, W.; Cavalli, A. Kinetics of Drug Binding and Residence Time. *Annu. Rev. Phys. Chem.* **2019**, *70*, 143–171.

(9) *Django*, version 2.2; Django Software Foundation: Lawrence, KS, 2019.

(10) *Bootstrap*, version 3.3; Twitter, Inc.: San Francisco, CA, 2016.

(11) Cock, P. J. A.; Antao, T.; Chang, J. T.; Chapman, B. A.; Cox, C. J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; De Hoon, M. J. L. Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423.

(12) Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H. P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; Bianciotto, M.; Rak, A.; Wade, R. C. Estimation of Drug-Target Residence Times by  $\tau$ -Random Acceleration Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2018**, *14*, 3859–3869.

(13) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562–12566.

(14) Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of Protein-Ligand Unbinding: Predicting Pathways, Rates, and Rate-Limiting Steps. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E386–E391.