

# BioModels Database, a curated and annotated resource of published quantitative kinetic models

Chen Li<sup>1,\*</sup>, Marco Donizelli<sup>1</sup>, Nicolas Le Novère<sup>1</sup>, Harish Dharuri<sup>2</sup>, Enuo He<sup>2</sup>,  
Lu Li<sup>1</sup>, Bruce Shapiro<sup>3</sup>, Melanie Stefan<sup>1</sup>, Rainer Machne<sup>4</sup>, Camille Laibe<sup>1</sup>,  
Dominic Tolle<sup>1</sup>, Renaud Schiappa<sup>1</sup>, Michael Hucka<sup>2</sup>

1. European Bioinformatics Institute, Hinxton, CB10 1SD, UK.
  2. Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, USA.
  3. Biological Network Modeling Center, California Institute of Technology, Pasadena, CA 91125, USA.
  4. Theoretical Biochemistry Group, Institute for Theoretical Chemistry, University of Vienna, Währingerstrasse 17, 1090 Vienna, Austria.
- \*E-mail: chenli@ebi.ac.uk

## Introduction

The development of quantitative models of biochemical and cellular systems, once the domain of specialists, is now part of mainstream biology. Those models need to be verified and distributed as for any other data in life science. BioModels Database (<http://www.ebi.ac.uk/biomodels/>) is a free resource for storing, viewing and retrieving published, peer-reviewed quantitative models of biochemical and cellular systems [1].

Models are thoroughly curated and available in SBML [2], CellML [3], BioPAX, SciLab and XPP etc. In addition, model elements are carefully annotated with terms from controlled vocabularies and links to relevant data resources. And when viewing models, users can efficiently navigate through them by clicking on the model elements in the reactions graphs.

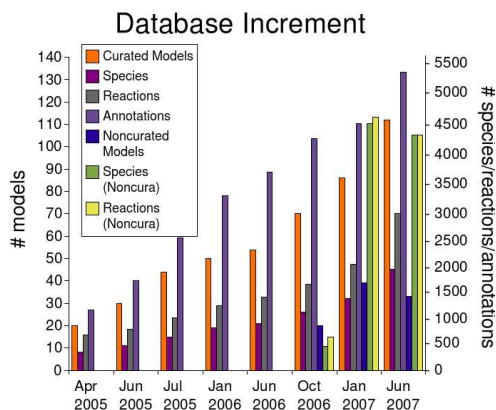


Figure 1: BioModels Database was launched in April 2005 and immediately gained recognition [4]. The eighth public release in June 2007 contained 112 curated and 33 non-curated models. These models represent 6498 species and 7684 reactions, and comprise 5330 cross-references to other data resources. And since this release, BioModels Database collaborated with JWS Online [5] to start offering on-line simulation for several models.

## Database curation pipeline

Models are accepted in either SBML or CellML, when being submitted into BioModels Database. Forthcoming versions will accept other formats, such as VCML(<http://www.nrcam.uchc.edu/>). In order to increase user's confidence, the models have gone through the curation pipeline before becoming public. Overall, the pipeline could be divided into three phases(see Figure 2).

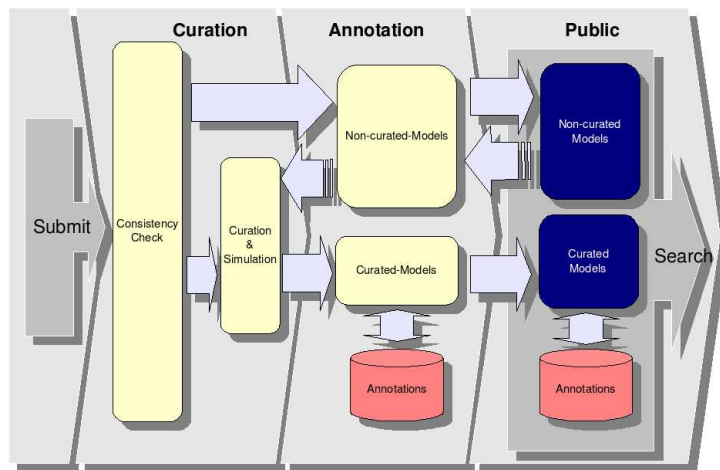


Figure 2: BioModels Database Pipeline, including Curation, Annotation and Public

- Curation phase

Firstly, models are checked for syntax, self-consistency and referential integrity. The curators then carefully read the original paper describing the model, and verify that the model follows the Minimal Information Requested In the Annotation of biochemical Models [6]. That requires to ensure that the model's simulation results are in agreement with the results described in the paper.

- Annotation phase

Well curated models are then moved to the Annotation phase. Model's elements are cross-referenced to terms of controlled vocabularies in other data resources. This annotation permits to unambiguously identify the model elements, avoiding the confusion brought by alternative naming of the same biological entity in different models.

- Public phase

On a periodical basis, models are delivered to the public. Biologists can freely browse, download the models and reuse them.

## Curated and Non-curated branches of BioModels Database

After moving to the Annotation phase, models are split into two branches. In the curated branch, all the models are verified to be MIRIAM compliant, and annotated. Alternatively, models can be distributed through the non-curated-branch. Various reasons may explain the decision to put a model in the non-curated-branch: the models cannot be simulated (Flux Balance Analysis), cannot be entirely encoded in SBML (spatial models) or the published results cannot be reproduced. Nonetheless, even those models have passed both the syntax and consistency checks.

## Access to the models in the BioModels Database

The web interface of BioModels Database is mostly used by biologists to view and download models for simulating etc. In comparison, the Web Services give more freedom to third-party software that need to directly access up-to-date data from BioModels Database.

With the rise of Systems Biology, BioModels Database community is rapidly growing. In 2007, BioModels Database was ranked first data resource for Systems Biology [7]. And deposition of models upon publication is currently supported by Nature Publishing Group, PLoS and BioMed Central.

## References

- [1] N. Le Novère et al. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res*, 34(Database issue):D689–D691, Jan 2006.
- [2] M. Hucka et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 2003.
- [3] CM Lloyd et al. CellML: its future, present and past. *Prog Biophys Mol Biol*, 85:433–450, 2004.
- [4] In pursuit of systems. *Nature*, 435, 5 2005. doi:10.1038/435001a.
- [5] B.G. Olivier and J.L. Snoep. Web-based kinetic modelling using jws online. *Bioinformatics*, 20(13):2143–2144, Sep 2004. PMID:15072998.
- [6] N. Le Novère et al. Minimum information requested in the annotation of biochemical models (MIRIAM). *Nature Biotechnology*, 23(12):1509–1515, December 2005.
- [7] E. Klipp et al. Systems biology standards the community speaks. *Nature Biotechnology*, 25:390–391, 2007. doi:10.1038/nbt0407-390.