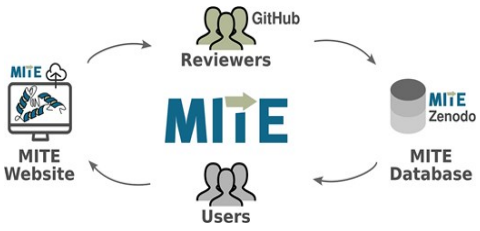
Master Thesis Project

Engineering biological databases: case studies on the MITE database

Supervisor	Ass.-Prof. **Mitja M. Zdouc**. ZdoucLab.org Division of Pharmacognosy, University of Vienna.
Topic	Data science and visualization, biological databases, software engineering, frontend/backend development
Duration	6 Months (full-time), **computational work only**
Language	English/German

Description

The biosynthesis of specialized metabolites, including many antibiotics, is governed by a specific group of enzymes. These “tailoring” enzymes are responsible for converting the nascent metabolites into their mature and often bioactive forms. To provide data on tailoring enzymes in a machine-readable form, we have recently developed **MITE**, the **M**inimum **I**nformation about a **T**ailoring **E**nzyme database [1] (https://mite.lisc.univie.ac.at/). MITE is a community-driven project, with researchers creating and reviewing entries, before they are released to the public domain. With increasing popularity, new use cases are appearing, which are yet not addressed by the web application. In this project, you will improve the MITE web app by creating functionality to cater emerging needs of the research community. Depending on your interest, you may choose between several topics and/or combine them:



- Re-engineer the frontend to improve data visualization, the submission of entries, and interactions with the database.
- Extend MITE’s REST API to support authentication, database querying, data submission, and reviewing, including support for automated data exchange with the MIBiG [1] and Rhea [2] databases.
- Explore Nanopublications [4] as an emerging format to publish MITE entries as linked open data by investigating appropriate ontologies and publication mechanisms, for instance via GitHub Actions.

These projects will significantly improve the usability of the MITE database and contribute to free and open science, accelerating the discovery of new specialized metabolites.

Learning outcomes

At the end of this Master thesis, you will be able to:

- **Apply** software engineering in distributed development, following clean code principles.
- **Analyze** needs of users of biological databases and transfer them into applications.
- **Define** and test hypotheses, design data analysis tasks, and present findings in word and writing.

How to apply

Please contact Ass.-Prof. Mitja Zdouc via E-Mail (mitja.zdouc@univie.ac.at). Include your CV and a brief summary of the progress of your studies, your research interest, and why you would like to work on which project. Indicate any prior knowledge related to the topic, as well as your preferred working style, supervision expectations, and preferred starting date. Applicants will be invited for a brief interview where the project will be discussed in detail.

References:
1 Rutz et al. 2025 Nucl Acids Res DOI: 10.1093/nar/gkaf969
2 Zdouc et al 2025 Nucl Acids Res DOI: 10.1093/nar/gkae1115
3 Bansal et al 2022 Nucl Acids Res DOI: 10.1093/nar/gkab1016
4 https://nanopub.net/