

EnzymeML: F.A.I.R. data management in biocatalysis

Jürgen Pleiss

Institute of Biochemistry and Technical Biochemistry, University of Stuttgart

Allmandring 31, 70569 Stuttgart, Germany

Juergen.Pleiss@itb.uni-stuttgart.de, <https://www.ibtb.uni-stuttgart.de/tb/>

Key Words: *enzyme catalysis, standardization, reproducibility*

Enzyme catalysis provides a powerful toolbox for novel, sustainable synthesis routes and innovative solutions for bio-based chemistry. A comprehensive biochemical characterization of the desired enzyme-catalyzed reaction is essential and provides the basis for enzyme engineering and process development. Standardization of reporting of enzymatic data and metadata is considered as pivotal to accelerating bioprocess development and reducing costs. Meta-research studies suggest the lack of standardization to report and share experimental protocols, results, and data as one of the causes of the reproducibility crisis in the biomedical sciences. As first steps for the standardized reporting of enzyme function data, the enzymology and biocatalysis community has established the Standards for Reporting Enzymology Data (STRENDA) Guidelines, the STRENDA DB as a public database to make enzymatic data findable and accessible, and the XML-based data exchange format EnzymeML to make enzymatic data interoperable and reusable. An Application Programming Interface facilitates the integration of applications such as electronic lab notebooks, modelling platforms, or databases. Thus, EnzymeML enables a reproducible and scalable documentation of biocatalytic data according to the F.A.I.R. principles.