

2026 Helmholtz – OCPC – Program

for the involvement of postdocs in bilateral collaboration projects

Title of the project:

Chemotion4Bio: Development of a research data infrastructure for chemists and biologists

Helmholtz Centre and/or institute:

Karlsruhe Institute of Technology (KIT)

Project leader:

Dr. Nicole Jung/ Prof. Dr. Stefan Bräse

Contact Information of Project Supervisor: (Email, telephone)

nicole.jung@kit.edu; +49721-608-24697

Web-address:

<https://fms.ibcs.kit.edu/>

Department: (at the Helmholtz centre or Institute)

Institute of Biological and Chemical Systems – Functional Molecular Systems (IBCS-FMS)

Program Coordinator (Email, telephone)

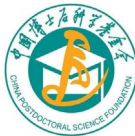
Name: Oliver Kaas

Phone: +49-721-608-45323

Email: oliver.kaas@kit.edu

Description of the project (max. 1 page):

The electronic lab notebook (ELN) Chemotion has become a widely adopted open-source platform for documenting, managing, and publishing chemical research data. It includes various chemistry-specific functions for handling chemistry data, such as chemical structures, chemical reactions, and analytical data. Its relevance in chemistry led to its integration into the National Research Data Infrastructure (NFDI4Chem). However, modern life science research increasingly combines chemistry with molecular biology, biotechnology, and biomedical sciences. To support interdisciplinary research and ensure FAIR (Findable, Accessible, Interoperable, Reusable) data principles across domains, we propose the systematic integration of biology-specific functionalities into Chemotion. This project will expand Chemotion's capabilities to include documenting biological experiments, handling structured biological data, and supporting established bioinformatics standards. The result will be a unified, open, and sustainable ELN platform for chemistry–biology research that will allow a flexible extension to a comprehensive biology platform in the long run. The work will consist of four phases that will build on each other. The work will be done by the postdoc in collaboration with the developers and biologists working at IBCS-FMS. Another proposal



planned for the extension of Chemotion into the research field of biology will be submitted to the German Research Foundation, and to further support the planned work described in here to achieve the required personnel. The work will be further coordinated with other developers working in the team Chemotion (working on DFG-based funding complementing the code contributions by the postdoc), allowing a comprehensive coverage of the relevant entities and tools for biological sciences.

Phase I - Implementation of Biology Standards and Tools: The developer will include selected data models for synthetic biology, such as SBOL, MIAME, and FASTA to Chemotion. Wherever available, tools for the visual representation of biological entities should be considered for integration into the environment of Chemotion ELN.

Phase II - Design and Implementation of Biology Data Models: The implementation will include structured documentation for the biological entities, plasmids, strains, cell lines, DNA/RNA (proteins already exist). Metadata schemas will be aligned with community standards.

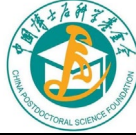
Phase III - Workflow Modules for Molecular Biology: To allow the comprehensive use of the standards, tools, and entities, the developer will establish templates for cloning, PCR, transfection, and CRISPR workflows with the biologists at IBCS-FMS.

Phase IV - User Testing and Community Integration: For all developed and implemented features, training materials, and documentation will be generated and provided. The implementation will be tested in the interdisciplinary research environment at KIT with different biology user groups at IBCS-FMS. The feedback will be collected, followed by iterative improvement of the implemented functions.

The developments will be published as Open Source Code under an AGPL license. After final integration into the main Chemotion ELN source code, the developments will expand the current research data management landscape in chemistry and biology. As such, the code extensions will be maintained in the long run by the National Research Data Infrastructure.

Description of existing or sought Chinese collaboration partner institute (max. half page):

The software Chemotion is already used by individual scientists and a few labs in China. We would like to extend the use of Chemotion, building software that is used and supported internationally and further developed. Besides the software itself, standardizing data and workflows is of primary interest. A core collaboration goal would be co-developing standardized data models that can be shared across institutions and nations. This will include the inclusion of metadata standards for biological entities (cell lines, plasmids, sequences) that align with international frameworks and are supported by German and Chinese research institutions. One of the aims is to harmonize biological workflow templates (e.g., PCR, cloning, CRISPR) to ensure interoperability across organizations. Through this collaboration, we will ensure that Chemotion's biology modules comply with Chinese data standards and regulatory requirements, potentially enhancing adoption in Chinese labs. Ideally, the Chinese partner will host pilot deployments of the biology-enabled Chemotion ELN in their own research groups, but it is not required. If the Chinese partner hosts an instance, sharing of feedback on usability, local requirements (e.g., language, regulatory), and integration with local infrastructure (e.g., LIMS and HIS systems in Chinese institutions) could help for the development of Chemotion and other open source infrastructure. We hope that the collaboration will strengthen scientific work and accelerate scientific progress by fostering the use of research data management systems and by sharing data for published



information. We aim to establish common databases across nations and to develop concepts to achieve this in collaboration with a Chinese partner. We intend to strengthen the open-source and open-science ecosystem -especially given China's growing role in life sciences and data infrastructure.

Required qualification of the postdoc:

- PhD in computer science, bioinformatics, or any field in natural sciences.
- Experience in software development (front-end, back-end): ReactJS, Ruby (if possible), HTML
- Additional skills in Python; a developer with a background in biology preferred
- Language requirement: English