PhD Program between Freie Universität Berlin (FUB) and the China Scholarship Council (CSC)

Open PhD Position at Freie Universität Berlin, offered only to Chinese CSC scholarship candidates 2027

Department/Institute: **Department of Biology, Chemistry, Pharmacy, Institute of Chemistry and Biochemistry, Takustr. 3, 14195 Berlin**

Subject area: **Cheminformatics / NMR Spectroscopy / Artificial Intelligence in Chemistry**

Name of Supervisor: **Prof. Mathias Christmann**

Number of open PhD positions: *1*

Type of the PhD Study: Full-time or Sandwich-Model

Project title: Automated High-Throughput Analysis of NMR Data for Internal Consistency

PhD Project description:

This PhD project will develop an automated high-throughput platform that extracts NMR characterization data from publications, Supporting Information and laboratory reports and checks them for internal consistency. Extending a Python workflow demonstrated for accurate-mass data across more than 3000 Supporting Information files, the project will combine deterministic chemical rules with large language models to recognize heterogeneous 1H, 13C and 2D NMR reports, normalize the data and flag contradictions among molecular formulas, signal counts, integrations, multiplicities, coupling constants and assignments.

The candidate will build privacy-preserving pipelines with agentic coding workflows and locally deployed LLMs so unpublished data can remain on institutional hardware. Performance will be benchmarked against curated examples, with explicit measurement of extraction accuracy, false positives and failure modes. The goal is a reproducible tool that improves data quality and reduces the workload of researchers, reviewers and editors while keeping expert review in the loop.

[1] M. Christmann, What I Learned from Analyzing Accurate Mass Data of 3000 Supporting Information Files, *Org. Lett.* 2025, 27, 4-7, DOI: 10.1021/acs.orglett.4c03458.

Language requirements:

- IELTS: 6,5 oder TOEFL: 95 ibt

Academic requirements:

Master's degree in chemistry, cheminformatics, computer science, data science or a closely related field. Strong Python programming skills are required, together with practical experience in agentic coding, large language models (LLMs) and locally deployed LLMs. Familiarity with organic chemistry and NMR spectroscopy is highly desirable. Experience with document parsing, data pipelines, software testing, version control, benchmarking or privacy-aware local computation is an advantage.

Information of the professor or research group leader (website, awards etc.):

Prof. Dr. Mathias Christmann is Full Professor of Organic Chemistry at Freie Universität Berlin and head of the Christmann Research Group at the Institute of Chemistry and Biochemistry. His research focuses on synthetic organic chemistry, natural product total synthesis, the development of new synthetic methods, organocatalysis and sustainable chemistry. He received the Carl Duisberg Memorial Prize (2013), the DECHEMA Young Scientist Award for Natural Products Research (2012), as well as fellowships from the Fonds der Chemischen Industrie.

Website:
https://www.bcp.fu-berlin.de/en/chemie/chemie/forschung/OrgChem/christmann/index.html

Publications of the group:
https://www.bcp.fu-berlin.de/en/chemie/chemie/forschung/OrgChem/christmann/Publications/index.html

Information on FU Berlin:
https://www.fu-berlin.de/universitaet/leitbegriffe/filmportraet/filmportraet_englisch_lang/index.html

Please note:
In a first step, the complete application should be uploaded to the online portal (https://fuberlin.moveon4.de/form/60acfece5d328710e40bdbd5/eng) for evaluation by January 15th, 2027.