MSCA COFUND Doctoral Fellowships “Machine Learning Interatomic Potentials for Chemical Reactions”

Summary

This PhD topic is part of the INNOCHEMBIO Doctoral Programme, which is funded through the Marie Skłodowska-Curie Actions (MSCA) COFUND action. The main objective of INNOCHEMBIO is to train future experts of sustainable chemistry and biotechnology, helping Europe to take the next steps in the green transition. This project aims to extend machine learning potentials to reliably model ions, radicals, or non-equilibrium geometries, enabling rapid and accurate calculation of chemical reactivity. This PhD position will be hosted at TalTech with main supervisor Dr. Mario Öeren and will contain a secondment to co-supervisor Dr. Matthew D. Segall at OPTIBRIUM LIMITED (UK).

Research field:	Chemistry and biotechnology
Supervisor:	Dr. Mario Öeren
Availability:	This position is available.
Offered by:	School of Science Department of Chemistry and Biotechnology
Application deadline:	Applications are accepted between July 01, 2025 00:00 and August 31, 2025 23:59 (Europe/Zurich)

Description

1. General description of programme and host

The PhD fellowship is part of the Marie Skłodowska-Curie Actions (MSCA) COFUND doctoral programme INNOCHEMBIO (https://taltech.ee/en/innochembio), which is co-funded by the European Union (Grant agreement 101217295). The main objective of INNOCHEMBIO is to train future experts to help Europe take the next steps in the green transition. The solutions and trained experts can reduce the environmental impact of the chemical and agricultural industries, offer eco-friendly analytical techniques, and assess the safety of new materials. INNOCHEMBIO funding will co-finance **15 PhD positions**, for which the application process in the first call will start on the **1st of July in 2025**.

For 12 PhD positions the hosting institution will be Department of Chemistry and Biotechnology (DCB) at Tallinn University of Technology (TalTech) which combines three divisions – Chemistry, Gene Technology and Biomedicine, and Food and Biotechnology. DCB is developing solutions to the great challenges of the 21st century – climate change, environmental protection, carbon neutrality, renewable energy, and biodiversity conservation. DCB hosts the second biggest PhD programme in TalTech with nearly a hundred enrolled students.

For 3 PhD positions the hosting institution will be the implementing partner – National Institute of Chemical Physics and Biophysics (NICPB). NICPB is a public research institution that conducts both fundamental and applied research, developing novel directions in fields ranging from material sciences to informatics. NICPB houses the Laboratory of Environmental Toxicology and several laboratories focused on fundamental research in NMR technologies with expertise dating back decades. The PhD training activities conducted by NICPB are funded through TalTech.

Importantly, each PhD project has one co-supervisor from another European country, which is detailed under the specific offer (see under supervisors' section). In total, INNOCHEMBIO has **19 associate partners from 11 European countries**.

2. Description of specific PhD project

Predictive models are essential in early-stage drug development, enabling efficient screening of large compound libraries and early identification of candidates with favourable activity and properties. The models are used to help researchers guide lead discovery and lead optimisation, reducing reliance on costly laboratory experiments. The models also promote more sustainable research by minimising waste, conserving resources, and reducing animal testing—aligning modern drug discovery with the principles of green chemistry.

While different model types form a spectrum, it is bounded by two clear extremes: models that rely purely on empirical data and are built using statistical or machine learning approaches, and those that are based on physical principles, such as force field methods or ab initio calculations. In drug discovery, we often lack sufficient data or encounter high variability in said data, making it difficult to build reliable predictive models with broad applicability. Conversely, physics-based approaches are time-consuming and impractical for screening large numbers of compounds.

The sweet spot on the spectrum is a hybrid approach that combines empirical data with faster, though less accurate, physics-based methods such as semi-empirical calculations. This strategy allows researchers to train models with fewer data points while leveraging physical principles to ensure broad applicability. As a result, it offers a practical balance between accuracy and scalability, making it especially valuable in early-stage drug discovery where rapid screening is essential.

In recent years, the machine learning interatomic potentials (MLIPs) have become popular since they offer faster calculation speeds compared to semi-empirical methods while being as accurate as ab initio methods they have been trained on. However, these methods have often been trained on equilibrium or near-equilibrium geometries of neutral molecules, causing significant inaccuracies when applied to ions, radicals or transition states.

Predictive models that rely on the aforementioned non-equilibrium geometries would greatly benefit from new MLIPs. The overarching goal of this research is to advance and expand existing MLIP approaches. The role of the successful candidate will be training of MLIPs for radicals of drug-like molecules that are the result of hydrogen abstraction. The study will be initially limited to compounds that contain the H, C, N, O chemical elements with the intention to subsequently expand the scope of chemical elements. Once the MLIPs have been successfully built, they shall be integrated into models that predict the metabolism mediated by the Cytochrome P450 family of enzymes—the most important enzymes involved in xenobiotic metabolism. This integration ensures that the MLIPs have immediate practical relevance and are positioned to make a high impact in the scientific community, facilitating rapid adoption by domain experts.

Link to the project: https://taltech.ee/en/innochembio/oeren

3. Supervisory team

- Tallinn University of Technology (main supervisor): Dr. Mario Öeren
- OPTIBRIUM LIMITED (UK): Dr. Matthew D. Segall (The length of the long-term mobility will be 6 months.).
- Tallinn University of Technology: Prof. Toomas Tamm

4. Requirements

- Excellent command of written and spoken English.
- MSc degree or equivalent in chemistry or biotechnology, or equivalent.
- Compliance with the rules of INNOCHEMBIO (e.g. eligibility, adhering to MSCA mobility rules, etc.).
- The primary workplace will be in Estonia. Therefore, candidates from outside the EU must be eligible to obtain a visa. The position is expected to start in the first half of 2026.
- The candidate must be proficient in computational chemistry, cheminformatics, or a related field.
- Proficiency in at least one of the following fields: data science, machine learning, or programming is expected.
- Experience with Python and familiarity with libraries such as NumPy, Pandas, and matplotlib will be highly advantageous.
- Knowledge of Linux would also be beneficial.
- The PhD candidate should be capable of presenting their research findings in group seminars, conferences, and peer-reviewed articles. Strong communication and writing skills are essential for conveying complex concepts to both specialists and broader audiences.
- The candidate should approach their research with openness and clarity, conveying developments as they evolve.

5. Duties and Responsibilities

- Undertake postgraduate research for specific doctoral research project at TalTech or NICPB, respectively.
- Present and publish research in both academic and non-academic audiences. Attend and participate in academic and non-academic conferences, events and seminars.
- Attend and participate in all training events and supervisory meetings.
- Be seconded to the associated partner as necessary to fulfil the grant obligations.
- Prepare progress reports and similar documents on research for funding bodies, as required.
- Actively contribute to the public engagement and outreach activities of the project.

- The above job descriptions are not exhaustive, the PhD candidate may be required to undertake other tasks, which are broadly in line with the above duties and responsibilities.
- Full-time employment (40 hours per week), temporary contract for 4 years.

6. Eligibility requirements

- The applicant must be a doctoral candidate (i.e. not already in possession of a doctoral degree at the date of the recruitment).
- At the time of recruitment, the researcher must not have resided or carried out their main activity (work, studies, etc.) in Estonia for more than 12 months in the three years immediately prior to the recruitment date. Compulsory national service and/or short stays such as holidays are not taken into account.

7. Benefits

- Competitive funding scheme, with a minimum gross monthly salary of EUR 2500. Topped by additional mobility allowances as well as optional family allowances (if applicable).
- Covered tuition costs, research costs and funding for short term mobility (i.e. conference attendance).
- Interdisciplinary and international research projects.
- Early-stage researcher position, with corresponding social and medical benefits in Estonia.
- Becoming a Marie Skłodowska-Curie PhD fellow.

8. How to Apply

All applications must be sent through TalTech's official application platform Glowbase and only applications submitted here will be considered for the programme. We ask the candidates not to contact the supervisors directly, in case of questions please write at innochembio@taltech.ee. Each application must include the following material: CV, 1-page motivation letter, copies of BSc and MSc study records and diplomas, scanned copy of valid photo ID, 2 reference letters, eligibility statement.

NB! The INNOCHEMBIO programme has additional requirements compared to the standard TalTech application process. Details on the exact nature of these documents and how to insert them in Glowbase can be found at our official INNOCHEMBIO website: https://taltech.ee/en/innochembio/application-process. **If any of the required documents are missing, the candidate will not be eligible to proceed to the selection stage.**

9. Selection Process

The selection and recruitment process will be in accordance with the European Charter and Code of Conduct for the Recruitment of Researchers. The recruitment process will be open, transparent, impartial, equitable, and merit-based. There will be no overt/covert discrimination based on race, gender, sexual orientation, religion or belief, disability or age. To this end, the following selection criteria will be considered.

The application deadline is 31 August 2025. The application process will be carried out in 3 steps. In short, first an eligibility check is performed. All eligible candidates will proceed to stage 1, where they will be evaluated by independent evaluators based on the application documents. Lastly, shortlisted candidates from stage 1 will proceed to stage 2, where they will be interviewed via teleconference, which will be used to determine a candidate to whom an offer will be made. All candidates will be informed about the progress in due course after each step of the process. The selection process is described on the guide for applicants available here: https://taltech.ee/en/innochembio/application-process.

10. Disclaimer

By applying for this position, the applicants

1. give their consent to circulate their application and personal data within the INNOCHEMBIO consortium and with the evaluators;
2. confirm that the data provided is valid and accurate;
3. confirm compliance with the eligibility requirements;

4. commit to undertaking the planned secondment at the co-supervisor’s institution.

To get more information or to apply online, visit https://taltech.glowbase.com/positions/1029 or scan the the code on the left with your smartphone.