**University of Leicester**

**Future 50 PhD Scholarship**

|  |  |
| --- | --- |
| **Project Reference** | RI LISCB Schwabe |

|  |  |  |  |
| --- | --- | --- | --- |
| **First Supervisor** | Prof. John Schwabe | | |
| **School/Department** | Molecular and Cell Biology /LISCB | | |
| **Email** | [js336@leicester.ac.uk](mailto:js336@leicester.ac.uk) | **Telephone Ext** | 1287 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Second Supervisor** | Dr. Timothy Ragan | | |
| **School/Department** | LISCB | | |
| **Email** | [tj.ragan@leicester.ac.uk](mailto:tj.ragan@leicester.ac.uk) | **Telephone Ext** |  |

|  |  |
| --- | --- |
| **Additional Supervisor** | Dr. Katie Cunnea, Evotec Ltd. |

**Section 2 – *Project Information***

|  |  |  |
| --- | --- | --- |
| **Project Title** | Machine Learning and Artificial Intelligence tools for rapid structure determination and analysis via CryoEM | |
| **Project Highlights:** | 1. | Develop tools to increase speed and quality of CryoEM sample preparation techniques. |
| 2. | Improve decision making in CryoEM data processing by capturing expert decision making heuristics. |
| 3. | Enhance understanding of flexible biological macromolecules using AI based representations. |
| **Project Summary** | | |
| Over the past decade, Electron Cryo-microscopy (cryoEM) has experienced a ‘resolution revolution’ due to both hardware and software advances, and is now capable of producing structures that can rival those produced by X-ray crystallography. In many instances, including large complexes and flexible proteins, cryoEM is the only approach capable of producing biologically useful structural information. Unlike crystallography, however, techniques to produce vitrified samples and analyse the large volume of data from these samples are still under heavy development, and accepted ‘best practices’ in these areas are evolving rapidly.  During the early stages of a cryoEM project, a significant amount of time can be spent optimising protein production & buffer conditions as well as grid freezing conditions. A systematic approach to cryoEM sample screening that can quickly assess multiple conditions, on separate grids and with small datasets, and predict which should be used for full data collection and analysis is essential for the field to progress.  Once suitable samples have been obtained and data collected, the process of converting that data into interpretable maps still requires considerable expert judgement. This makes successful rapid analysis dependent on the availability of experienced experts. Machine learning systems able to capture and communicate the heuristics of those experts are needed not only to improve the extraction of information from the experimental data, but also improve the transmission of skill from more experienced to less experienced workers.  The final maps produced from CryoEM analyses are static, with the considerable variety of flexibility and dynamic information often lost. New methods for interpreting these data, using artificial intelligence based embeddings, are starting to show the true information content available in these experiments. Making that information accessible, and translating it to biological insight, is one of the most exciting areas of development in the field right now.  This project will expose the student to the entire pipeline from sample preparation through to structural insight. The project will involve some wet-lab work and microscopy. The primary focus, however, will be on the computational aspects, including framework, high-performance computing, and development and application of machine learning / artificial intelligence systems. | | |