An Auto-Meshing Pipeline for Biosimulation at the Exascale

Megan Ratcliffe, James Gebbie-Rayet, Charles Moulinec
STFC Daresbury Laboratory, Sci-Tech Daresbury, Keckwick Lane, Daresbury, Warrington, WA4 4AD, UK

code_saturne

UKRI Science and Technology Facilities Council

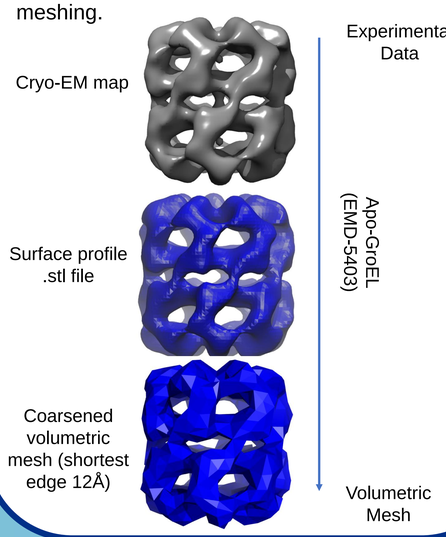
UNIVERSITY OF LEEDS

Current Biosimulation Methodology

Biosimulations are pivotal in understanding protein-protein interactions and developing new healthcare technologies. Molecular dynamics are a particle-based example of such simulations and have developed a close relationship with X-Ray crystallography. However, the rapid emergence of cryo-electron microscopy (cryo-EM) has led to imaging of structures many orders of magnitude higher than can be viewed using X-ray crystallography. Thus posing challenges for particle-based simulations which are highly computationally expensive at this scale.

FFEA Approach

The Fluctuating Finite Element Analysis (FFEA) approach uses meshes over atomic co-ordinates, therefore lending itself to data produced by cryo-EM. The original software was developed by a team at Leeds University and utilises continuum physics to model biomolecules as a finite element tetrahedral mesh and thus the volumetric data is ideal for meshing.

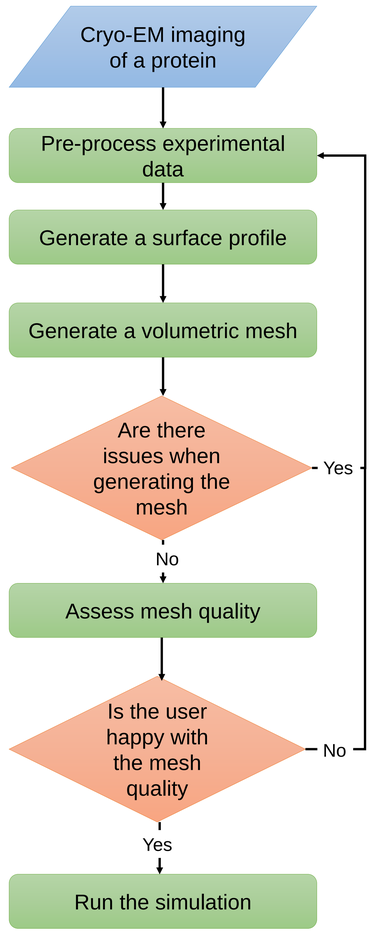


The Project

The automated pipeline is a part of a wider collaboration project between multi-disciplinary scientists at STFC and Leeds University researchers to re-implement the physics of FFEA into Code_Saturne to create a new biosimulation tool called Bio_Saturne. The pipeline aims to solve usability issues with the initial software surrounding mesh generation from noisy experimental data. These issues include:

- Holes in the mesh
- Widely heterogenous tetrahedra
- Overlapping surface faces

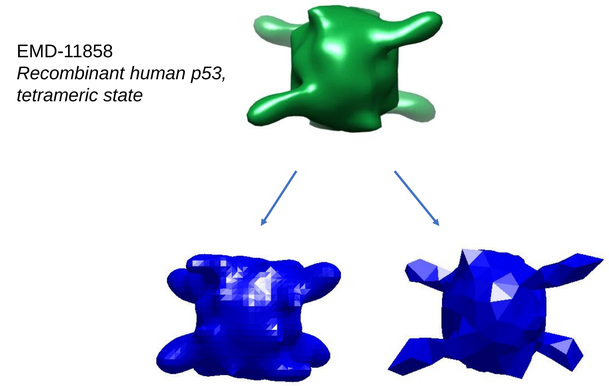
Software Pipeline



Automation

The software pipeline will be automated to enhance usability of the Bio_Saturne toolkit. The program will interact with the command line interface to run the required tools to build a mesh and extract any error messages which may arise.

EMD-11858
Recombinant human p53, tetrameric state



The functionality of the pipeline will allow users to access mesh quality scoring and provide both automated and user-defined adjustments. Intuitive messages will be displayed to the user from processed command line messages to allow them to apply finer modifications.

Running at the Exascale

Combined with the pre-exascale readiness of Code_Saturne this framework is capable of tackling extremely large biological problems using next generation exascale computing in a way currently unfeasible with other methodologies.

Example of CFD for 57B Tetra Cell Mesh (HAWK – AMD ROME EPYC)

MPI Tasks	Time in Solver	Efficiency
131,072	68.959 s	100%
262,144	34.769 s	99%
524,288	**18.677 s**	**92%**

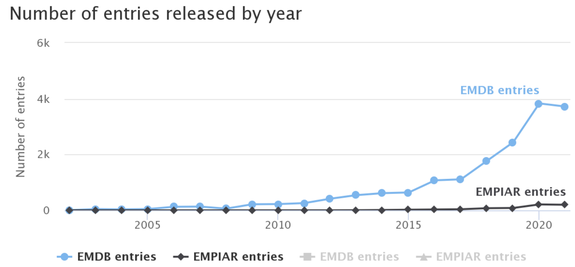
Biosimulations at the end of the pipeline could be ran on an exascale machine in several ways

- Complex protein simulations, requiring 100 billion elements.
- Several complex proteins interacting with each other.

In which cases the pipeline would be used once and several times respectively.

Future Application

The automated pipeline will make Bio_Saturne more accessible to the wider biomolecular simulation and experimental structural biology communities, this will lower the barrier to access the wealth of experimental data emerging from new microscopy sources.



Data from: https://www.ebi.ac.uk/emdb/

This includes both current and next generation cryo-EM microscopes, such as those at Diamond's Electro-Bioimaging Centre (e-BIC).

References and Acknowledgements

FFEA

Solernou A., Hanson B. S., Richardson R. A., Welch R., Harris S. A., Read D. J., Harlen O. G. "Fluctuating Finite Element Analysis (FFEA): A continuum mechanics software tool for mesoscale simulation of biomolecules" (2018), PLoS Comput. Biol. 14(3): e1005897.

Project Collaborators

- Leeds University
- Science Technology and Facilities Council
- Code Saturne Developers Team