Grant agreement no. 823712

CompBioMed2

Research and Innovation Action
H2020-INFRAEDI-2018-1
Topic: Centres of Excellence in computing applications

D2.1 – First Report on Fast Track Application Readiness

Work Package: 2

Due date of deliverable: Month 12

Actual submission date: 01 October 2020

Start date of project: 01 October 2019 Duration: 48 months

Lead beneficiary for this deliverable: *UvA*
Contributors: *BSC, UNIBO, UNIGE, UCL, UPF, USFD, UOXF, ACE*

Disclaimer

Project co-funded by the European Commission within the H2020 Programme (2014-2020)		
Dissemination Level		
PU	Public	YES
CO	Confidential, only for members of the consortium (including the Commission Services)	
CI	Classified, as referred to in Commission Decision 2001/844/EC	

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

Table of Contents

1 Version Log 3
2 Contributors 3
3 Definition and Acronyms 4
4 Public Summary 6
5 Introduction 6
6 Activities Carried out 8
6.1 Task 2.1 – Cardiovascular Research Exemplar 8
6.1.1 Subtask 2.1.1 – Whole body blood flow modelling 8
6.1.2 Subtask 2.1.2 – Improving diagnosis of cerebral vasospasm 10
6.1.3 Subtask 2.1.3 – Cell based Blood flow simulations 12
6.1.4 Subtask 2.1.4 – An *in silico* approach to guide clinical intervention in treatment of stroke 14
6.1.5 Subtask 2.1.5 – Developing modelling techniques for in silico trials 16
6.2 Task 2.2 – Molecularly-based Medicine Research Exemplar 19
6.2.1 Subtask 2.2.1 – Testing virtual candidate libraries against novel targets 19
6.2.2 Subtask 2.2.2 – Assessing sequence influence on drug binding using free energy calculations 20
6.2.3 Subtask 2.2.3 – Combining machine learning and molecular modelling to assess drug binding 22
6.3 Task 2.3 – Neuro-musculoskeletal Research Exemplar 23
6.3.1 Introduction 23
6.3.2 Subtask 2.3.1 – Growth and Adaptation 24
6.3.3 Subtask 2.3.2 - ARF10: Multiscale model of hip fracture 25
6.3.4 Subtask 2.3.3 – Virtual patients' expansion 26
7 Risk Management 27
8 Conclusions 28
9 Bibliography/References 28

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

1 Version Log

Version	Date	Released by	Nature of Change
V0.1	14/08/2020	Gabor Zavodszky	First Draft - Template
V0.5	10/09/2020	Gabor Zavodszky	Submitted draft for internal review
V1.0	1/10/2020	Gabor Zavodszky	Final Draft, submitted to the EC

2 Contributors

Name	Institution	Role
Gabor Zavodszky	UvA	Principal Author
Jazmin Aguado-Sierra	BSC	Contributor
Alejandro Varela Rial	Acellera	Contributor
Adrià Pérez Culubret	UPF	Contributor
Xinshan Li	USFD	Contributor
Ivan Benemerito	USFD	Contributor
Pinaki Bhattacharya	USFD	Contributor
Marco Viceconti	UNIBO	Contributor
Antonino Amedeo La Mattina	UNIBO	Contributor
Jon McCullough	UCL	Contributor
Shunzhou Wan	UCL	Contributor
Mat Bieniek	UCL	Contributor
Christos Kotsalos	UNIGE	Contributor
Blanca Rodriguez	UOXF	Contributor
Wang Lei	UOXF	Contributor
Gary Tresaden	JAN	Reviewer
Raimondas Galvelis	ACE	Reviewer
Peter Coveneey	UCL	Reviewer
Emily Lumley	UCL	Reviewer

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

3 Definition and Acronyms

Acronyms	Definitions
0/1/2/3D	0/1/2/3 Dimension
aBMD	Areal Bone Mineral Density
APDL	Algorithmic Processor Description Language
ARFn	Absolute Risk of Fracture at year n
BAC	Binding Affinity Calculator
CHU	Centre Hospitalier Universitaire
CMR	Cardiovascular Magnetic Resonance
CoE	Centre of Excellence
COPD	Chronic Obstructive Pulmonary Disease
CPU	Central Processing Unit
CT	Computed Tomography
CT2S	Computed Tomography to Strength
DEXA	Dual Energy X-ray Absorptiometry
DLB	Dynamic Load Balancing
DT	Deep Track
ECG	Electrocardiogram
ESMACS	Enhanced Sampling of Molecular dynamics with Approximation of Continuum Solvent
EU	European Union
FBDD	Fragment Based Drug Discovery
FE	Finite Element
FPGA	Field Programmable Gate Arrays
FSi	Fluid/Solid interaction
FT	Fast Track
FTI	Fault Tolerance Interface
GPCR	G-protein Coupled Receptor
GPU	Graphics Processing Unit
HMM	Heterogeneous Multiscale Model
HPC	High Performance Computing
I/O	Input/Output
LB	Latttice Boltzmann

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

LRZ	Leibniz Supercomputing Centre
MD	Molecular Dynamics
MPI	Message Passing Interface
MPIO	MultiPath Input Output
PCA	Principal Component Analysis
PV loops	Pressure-Volume Loops
RBC	Red Blood Cell
SED	Strain Energy Density
TALP	Termination Analysis of Logic Programs
TIES	Thermodynamic Integration with Enhanced Sampling
UEABS	The Unified European Application Benchmark Suite
UQ	Uncertainty Quantification
vBMD	Volumetric Bone Mineral Density
VVUQ	Validation, Verification, and Uncertainty Quantification
WP	Work Package

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

4 Public Summary

Work Package 2 (WP2) in the CompBioMed 2 project aggregates the computational biomedical research activities of the project. The leader of this work package is the University of Amsterdam (UvA). Computational Biomedicine is highly diverse, subsuming a large range of modelling and simulation methods, each applicable to a specific level of human bodily organisation. State of the art mathematical models in Computational Biomedicine need to take into account the inherent multiphysics and multiscale character of human physiology. This means that, typically, different types of models, representing different scales within a given (patient specific) biomedical system, often need to be taken into account, including the details as to how to correctly connect these different models.

The main objective of this work package is to lay the groundwork for the application of High Performance Computer (HPC)-based Computational Biomedicine approaches to a great number of therapeutic areas. The HPC requirements of our users are as diverse as the communities we represent; therefore, we must support both monolithic codes, potentially scaling to the exascale, and complex workflows requiring support for advanced execution patterns on a range of diverse and geographically distributed platforms.

5 Introduction

In Deliverable 2.1:First Report on Fast Track Application Readiness, we describe and present the deployment of the applications and the related research within CompBioMed2.

Building on these applications the objectives of WP2 include:

- To advance the state-of-the-art in simulation based biomedical science, from the desktop towards highly scalable, optimised codes for the most powerful existing multi-petaflops HPC systems;
- To prepare appropriate and high impact CompBioMed2 applications for the emerging exascale;
- To develop verification, validation and uncertainty quantification techniques that will permit CompBioMed2 applications to be used in *in silico* clinical trials and personal health forecasting applications;
- To engage with the broader CompBioMed community, advising on deployment of codes, need for parallelization and development of workflows on a range of HPC platforms;
- To develop cross-application systems models that link CompBioMed2 applications in multiscale and multiphysics scenarios, in combination with high performance data analytics where appropriate.

WP2 operates in strong collaboration with the other WPs, in particular, WP3 for data management, long term data storage, and high performance data analytic techniques, and WP4 for compute and data services, to optimize the access and usage of current and emerging e-infrastructures.

In terms of the structure of our work in CompBioMed2, we maintain the concepts of Fast and Deep Track, established in CompBioMed1 with the following definitions:

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

- The Fast Track (FT) supports the exploitation and enhancement of existing codes, services and workflows to produce novel scientific results. Notably, CompBioMed2's Fast Track incorporates developments emanating from CompBioMed1. The Fast Track also encompasses the seamless integration of component codes within the complex workflows necessary to produce integrated high fidelity personalised human models. By linking together codes developed by different partners, the Centre can provide a much wider range of capabilities than by merely supporting its individual partners' own particular research areas. The Fast Track is where all users reside and relies on technologies and tools ready or almost ready for use in production; however, many problems can be identified where existing solutions do not provide the features or performance necessary to adequately understand a system or produce results in a timely manner. In these cases, development in the Deep Track will be vigorously performed.
- The Deep Track (DT) aims to develop new capabilities and optimise existing software towards the exploitation of future exascale HPC environments. As well as helping partners (inside CompBioMed and in the wider community) improve the efficient use of massive computational resources, the Deep Track will provide integrated data analytics and Validation, Verification and Uncertainty Quantification (VVUQ) capabilities.

These application areas belong to one of the major research exemplars: cardiovascular research, molecularly-based medicine research, and neuro-musculoskeletal research. Furthermore, the applications target one of the three following major deployment groups:

- HPC for applied and theoretical research
- HPC for industrial application
- HPC for in silico medical applications and clinical trials

The way a given application is implemented for HPC execution is dictated by the requirements of the particular research area and the employed numerical solution. In the following we will refer to the publication [Introduction/1], which was produced as part of a previous exascale EU project (ComPat). We regard this publication as a clear guiding description for the computational pattern "building blocks" that represent the possible fundamental execution patterns for HPC applications and workflows.

For clarity, the description of three of these fundamental building blocks are duplicated herein from [Introduction/1, Introduction/2]. While they are developed for multi-scale coupled codes, it is straightforward to generalize the patterns to be applicable for single scale codes as well.

• **Extreme Scaling**, where a single or very few (typically single scale, monolithic) models require large-scale performance, these can be potentially coupled to other, less costly models.
• **Heterogeneous Multiscale Computing**, where a very large number of compute intensive microscale models are coupled to, typically, a single macroscale model.
• **Replica Computing**, where a large number of copies (replicas) are executed, that may or may not exchange information.

It is important to note that these categories are not necessarily mutually exclusive for a given code, as they are meant to describe a code in a given workflow, which is associated with a specific application. The codes themselves may be used as part of more than one workflow that might belong to different categories; we report here the one that corresponds to the primary

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

application of the code within CompBioMed. Furthermore, complex workflows may also present situations where different parts of the workflow belong to different categories.

Each and every one of these categories have the possibility to reach the exascale. The mechanism to achieve good scaling, however, differs among the categories. We supplement the code readiness and deployment descriptions with the categorization which is indicative of the appropriate scaling mechanism and where possible we also supply available information on the current scaling measures and the expected outlook within specific workflows.

6 Activities Carried out

In this section we summarize the current status of the research tasks and their results so far. The application codes are also described in terms of their deployment status, computational pattern structure with emphasis on the current scaling information and outlook.

6.1 Task 2.1 – Cardiovascular Research Exemplar

6.1.1 Subtask 2.1.1 – Whole body blood flow modelling

Aim

The main focus for subtask 2.1.1 is the study of macroscopic blood flow at full human scale. In particular, this makes use of the open source lattice Boltzmann code HemeLB to study flows in complex arterial and venous geometries in 3D. HemeLB has been optimized to efficiently simulate blood flow through the sparse and complex geometries that are typical of large-scale vascular structures. Simulating blood flow in 3D at this scale is an ambitious challenge that pushes the limits of current computational capacity. The outcomes of this work will provide deeper insight into cardiovascular flow and disease than can be provided by 0D/1D models.

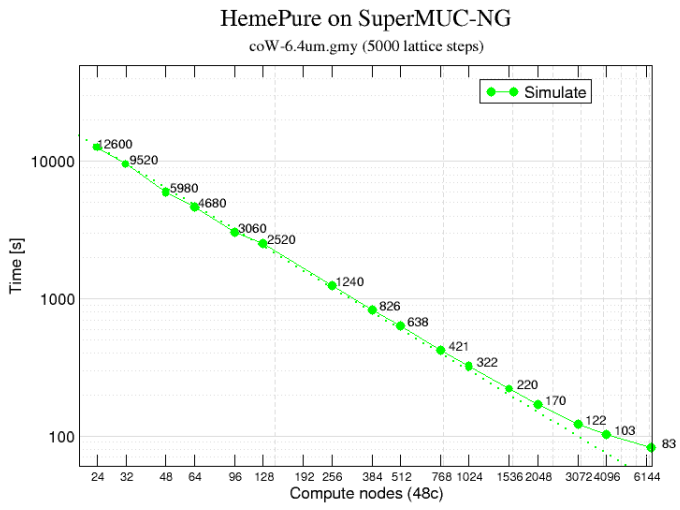

Figure 2.1.1/1 - Scaling of HemeLB up to 310,000 cores.

Current progress in subtask 2.1.1 can be illustrated in a number of critical areas. One of these has been demonstrating the scaling performance of HemeLB in realistic vascular geometries. In collaboration with the POP Centre of Excellence, we have demonstrated the strong scalability of an uncoupled version of HemeLB to 310,000 cores (full machine scale) on SuperMUC-NG [2.1.1/1, 2.1.1/2]. We have also demonstrated strong scaling to 30,000 cores for a self-coupled version of HemeLB with full human arterial and venous geometries.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

This is an excellent result for a first version of a coupled code and we plan to work with the POP CoE to identify further areas of improvement.

Geometry generation is a significant challenge for simulating 3D vessels at full-human scale. By pushing the reasonable wall-clock and memory capacity limits of SuperMUC-NG we have been able to resolve the full-human arterial and venous geometries (Fig. 1) provided by the IT'IS Foundation to 60µm (a total of approximately $2x10^9$ sites). To efficiently improve this domain we have worked with LRZ to develop a refinement tool that can halve the resolution of a geometry whilst avoiding the most memory intensive processes. In ongoing work we will be using these improved geometries to produce higher quality flow simulations of full-human blood flow.

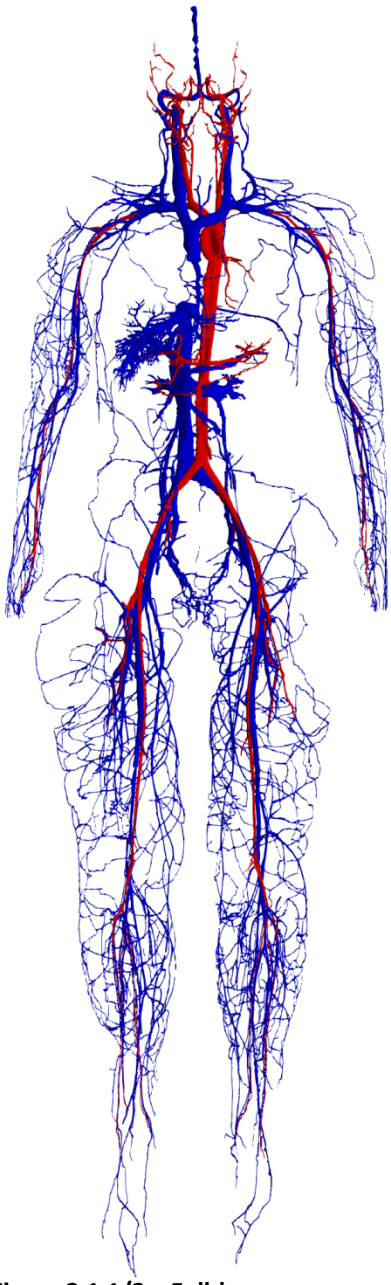

Figure 2.1.1/2 – Full-human arterial and venous domains of study

Investigation of sub-scale models for representing the capillary beds between the arterial and venous geometries has been an area of significant work. We have developed two approaches to this problem: the first assigning velocity scaling factors between coupled boundary locations and the second implementing a resistor-capacitor model similar to that used in lower dimensional studies. Publications discussing this work are currently either under review or in preparation. In both cases, the parameters for the sub-scale model could be adjusted to represent individuals with healthy or diseased vasculatures.

A key focus of ongoing work in this subtask will be the coupling of the HemeLB framework to that of the multiphysics heart model developed at the Barcelona Supercomputing Center, Alya. The coupling of these two codes will allow the full circulation of systemic blood flow to be captured in 3D. Visualisation of results at the scale of this work is an equally challenging domain. We will be working with LRZ to develop efficient methods for presenting the large datasets generated by HemeLB. Furthermore, we are working towards developing a GPU version of HemeLB to make use of these accelerators which are becoming a key performance feature of HPC infrastructure.

Software stack description

***Main code*:** HemeLB

***Components*:** Standalone and self-coupled versions of HemeLB for simulating single and coupled vascular geometries.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

***Libraries*:** MPI, TinyXML, Parmetis, CTemplate, Boost, ZLib

***Computational pattern(s)*:** Simulation studies conducted by HemeLB are monolithic and belong to the "Extreme Scaling" classification.

Current scaling information and outlook: The single version of HemeLB has been demonstrated to scale well up to 310,000 cores (full machine scale) on SuperMUC-NG, see [2.1.1/1] and Fig. 2.2.1/1. The self-coupled version has demonstrated strong scaling in a complex arterial-venous geometry up to approximately 30,000 cores on SuperMUC-NG, these results are currently unpublished.

Current deployment status (list of HPCs): HPC facilities where HemeLB is deployed include: SuperMUC-NG (Germany), Archer (UK), Cirrus (UK), MareNostrum (Spain), Juwels (Germany), Longhorn (USA), Piz-Daint (Switzerland). Deployment on Summit (USA) is anticipated in the near future. SuperMUC-NG and Archer represent the current main deployments for CompBioMed2.

6.1.2 Subtask 2.1.2 – Improving diagnosis of cerebral vasospasm

Aim

This subtask aims to unravel mechanism of phenotype variability in post myocardial infraction of the human heart via high performance computing simulations. Myocardial infarction involves death of myocyte and structural remodelling of myocardium following narrowing of the coronary artery. This structural remodelling leads to arrhythmia and mechanical dysfunction, which can be measured by Electrocardiogram (ECG) and cardiovascular magnetic resonance (CMR) in clinic.

During the first phase, we have already studied the ion channels for arrhythmia risk assessment [2.1.2/1], effect of size, transmural extent and location of ischemia on arrhythmia vulnerability and ECG alterations [2.1.2/2], performed a sensitivity analysis of a strongly-coupled human-based electromechanical cardiac model [2.1.2/3], and validated human electro-mechanical ventricular modelling and simulation [2.1.2/4].

In [2.1.2/3], the detailed description of a human-based physiologically-based, and fully-coupled ventricular electromechanical modelling and simulation framework is presented. It includes the coupling of state-of-the-art human-based electrophysiology membrane kinetics, excitation–contraction and active contraction models. Through high-performance computing simulations, a sensitivity analysis focused on its mechanical properties is performed.

In [2.1.2/4], the most recent models of human ventricular electrophysiology and active contraction are coupled and calibrated using experimental data. Simulations are conducted to quantify the effect of electro-mechanical coupling on the action potential, calcium dynamics, and active contraction in healthy virtual human ventricular cardiomyocytes and tissue and under drug exposure. Simulation results correctly predicted the inotropic response of different multichannel action reference compounds and demonstrated that the electro-mechanical coupling improves the robustness of repolarisation under drug exposure compared to electrophysiology-only models. They also generated additional evidence to explain the partial mismatch between *in silico* and *in vitro* experiments on drug-induced electrophysiology changes, see Figure 2.1.2-1.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

Currently, we are further developing the simulation tool. This is Alya, which is used in [2.1.2/3] and [2.1.2/4] for performing clinically-relevant simulations. For example, we have built a biventricular model for the human heart, which outputs Pressure/Volume (P/V) loops and ECG (Figure 2.1.2-2). This model will be improved and calibrated with the clinical data, and then used for simulation of post myocardial infarcted heart.

Software stack description

***Main code*:** Alya, Chaste, Virtual Assay
***Components*:** Alya modules for electrophysiology and solid mechanics
***Libraries*:** MPI, VTK
Computational pattern(s): Sensitivity analysis needs a lot of simulations with different model parameters, belonging to the "Replica Computing" pattern. Each simulation needs large scale performance, belonging to the "Extreme Scaling" computing pattern.

Current scaling information and outlook: In an example of sensitivity analysis, we run 350 simulations, each of them used 4 nodes, i.e. 1400 nodes for this example. This is limited by the maximum number of nodes.
Current deployment status (list of HPCs): Alya is currently deployed on MareNostrum (Spain), SuperMUC-NG (Germany) and Piz Daint (Switzerland).

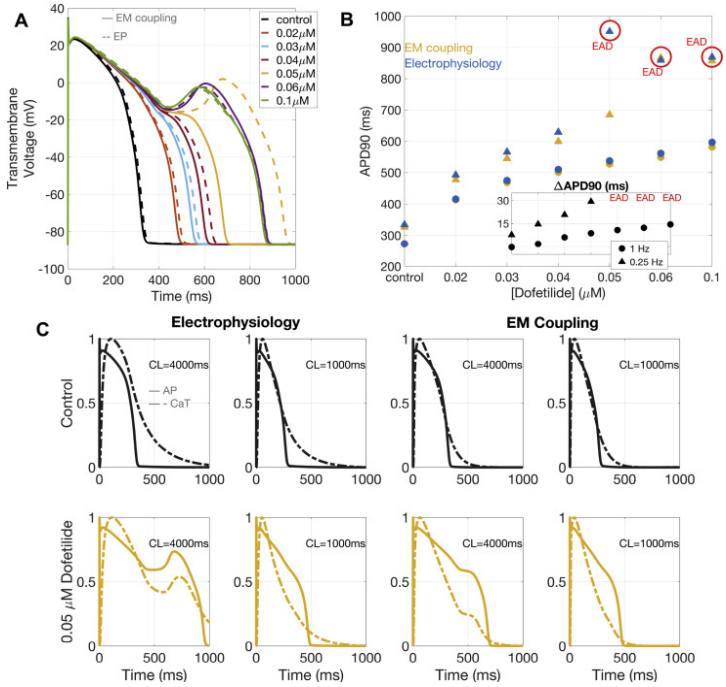

Figure 2.1.2-1: Electro-mechanical coupling and repolarisation. Electro-mechanical coupling (EM) mitigates the APD prolongation induced by Dofetilide, and delays EADs onset (A-B). This is mediated by a faster CaT decay under EM coupling (C), which recovers almost entirely during AP repolarisation. Results obtained with the ToR-ORd+Land model. CL: cycle length.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

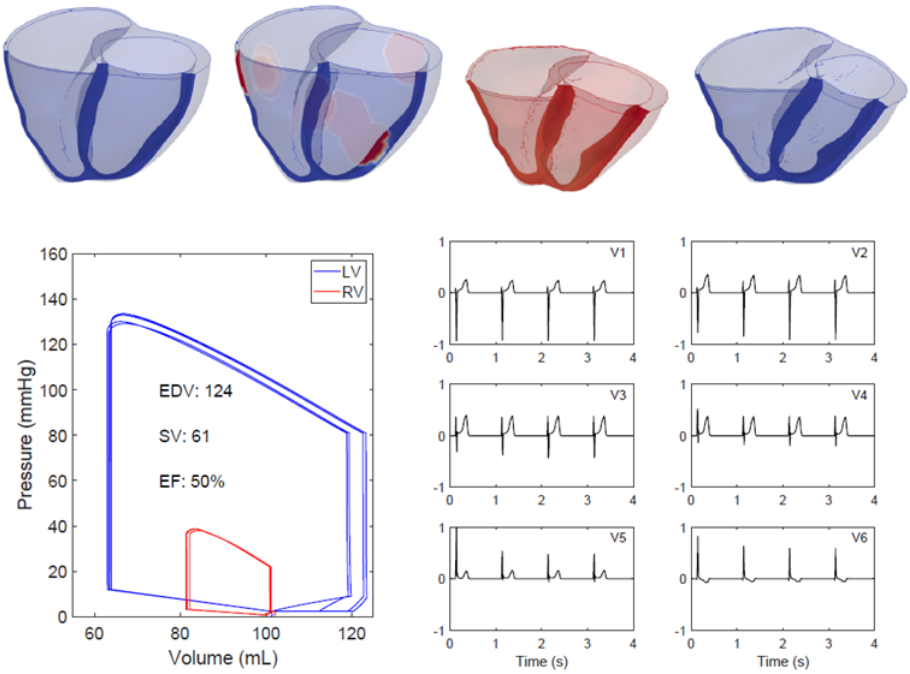

Figure 2.1.2-2: electromechanical simulations of the biventricular human heart: deformation, PV loops and ECG

6.1.3 Subtask 2.1.3 – Cell based Blood flow simulations

Aim

This subtask represents the main research effort in the investigation of blood flow on the cellular level. The main tool of the investigation is the open-source HemoCell code. Cellular resolution is necessary to be able to resolve many biological processes that play a significant role in various vascular diseases. This level of resolution, however, is quite demanding in term of computational resources, therefore we put great emphasis on improving the computational efficiency and reliability of our computations.

During the first phase of the project we have demonstrated a heterogeneous multiscale model of blood flow, where the larger scale of a whole vessel sections is strongly coupled to local microscale simulations representing detailed events from the cellular scale. This work is now published in [2.1.3/1]. The schematics of the coupling is shown in Fig. 1, representing the two spatial scales of interest.

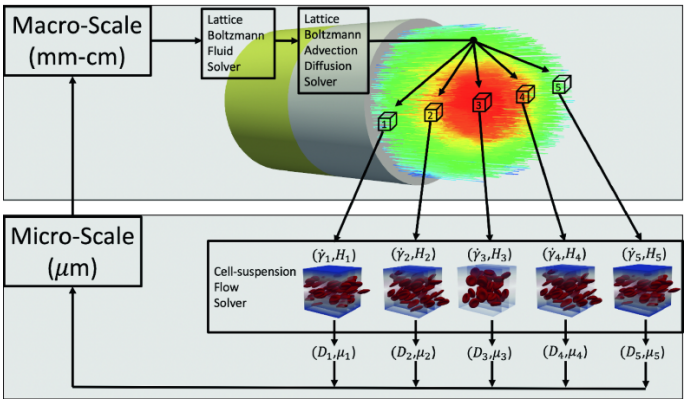

Figure 1 - Heterogeneous multiscale model schematic

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

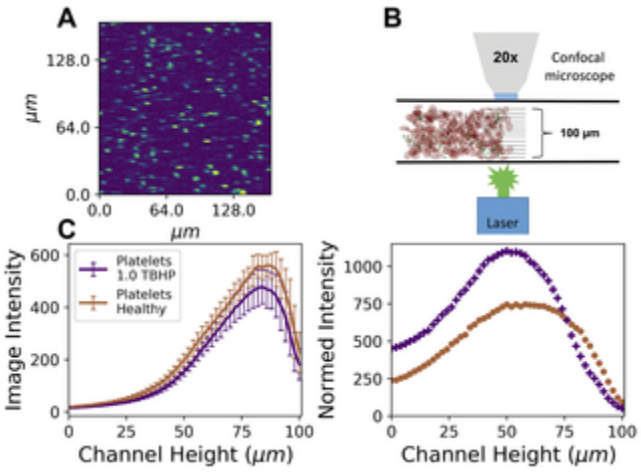

Figure 2 - Fluorescent measurement of stiffened cell distributions.

Furthermore, we have investigated the effect of changed cellular stiffness on the transport physics of blood. Such effects are associated with numerous diseases (e.g. diabetes, malaria, sickle cell disease), and their implications are poorly understood. Our work explores the results of the presence of diabetic cells by mimicking their changed mechanical properties. The results are validated through microfluidic measurements, and a snapshot of such measurement is shown in Fig. 2. The findings are published in [2.1.3/2]. The numerical model representing the diseased cells contain a set of parameters (that are fitted to experimental data), and to ensure the robustness and extensive validity of the model we have supplemented the investigation with thorough VVUQ analysis, where due to the high cost of the computational model we devised new methodologies for model parameter UQ. The findings are published in [2.1.3/3].

The currently ongoing efforts include work towards understanding cellular transport in the micro-vessels of the retina, inside micro-aneurysms that form in relation to the wide-spread disease of retinopathy. They also include more fundamental research questions such as general cellular transport and platelet margination in complex geometries.

All this happens in tandem with efforts to keep improving the computational efficiency by optimizations on the implementation level as well as through high-level load-balancing algorithms. We are also investigating heterogeneous execution, that is, porting the fluid dynamics part of the code to GPU to take better advantage of upcoming heterogeneous HPC architectures.

Software stack description

***Main code*:** HemoCell

***Components*:** Cellular mechanics code of HemoCell coupled to Palabos as the fluid flow solver

***Libraries*:** MPI, VTK, HDF5

Computational pattern(s): Most executions are monolithic, belonging to the "Extreme Scaling" pattern, whereas the HMM model executes as a "Heterogeneous Multiscale Computing" pattern.

Current scaling information and outlook: The monolithic execution scales well up to 8 000 computing cores, and the relevant efforts and results are documented in [2.1.3/4]. The heterogeneous multiscale execution is yet to be evaluated, but we expect exceptional scaling given it builds on multiple instances of the monolithic code.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

Current deployment status (list of HPCs): HemoCell is currently deployed at several HPC centres around the World (MareNostrum, Barcelona; Eagle, Poland; SuperMUC, Germany; Aspire, Singapore; Sanam, KACST; SGI Altrix, Australia; Lomonosov, Russia; Superman, Hungary; Cartesius, Lisa, Netherlands). The main deployment is on Cartesius (Netherlands).

6.1.4 Subtask 2.1.4 – An *in silico* approach to guide clinical intervention in treatment of stroke

Aim

This subtask represents the main research effort in the investigation of fully resolved 3D cellular blood flow simulations, deploying high fidelity models without compromising extreme performance. Towards this direction, we have developed Palabos-npFEM (Fig. 2.1.4-1) which is a computational framework for the simulation of blood flow with fully resolved constituents. The software resolves the trajectories and deformed state of blood cells, such as red blood cells and platelets, and the complex interaction between them. The tool combines the lattice Boltzmann solver Palabos for the simulation of blood plasma (fluid phase), a finite element method (FEM) [2.1.4/1] solver for the resolution of blood cells (solid phase), and an immersed boundary method (IBM) for the coupling of the two phases (FSI – Fluid/Solid interaction). Palabos-npFEM provides, on top of a CPU-only version, the option to simulate the deformable bodies on GPUs (CUDA implementation), thus the code is tailored for the fastest supercomputers. The software is integrated in the Palabos core library and is available on the Git repository https://gitlab.com/unigespc/palabos. It offers the possibility to simulate various setups, e.g. several geometries and blood parameters, and due to its modular design, it allows external solvers to readily replace the provided ones. The code demonstrates exceptional scaling across heterogeneous (hybrid CPU/GPU execution) infrastructure, and the benchmark can be found in [2.1.4/2].

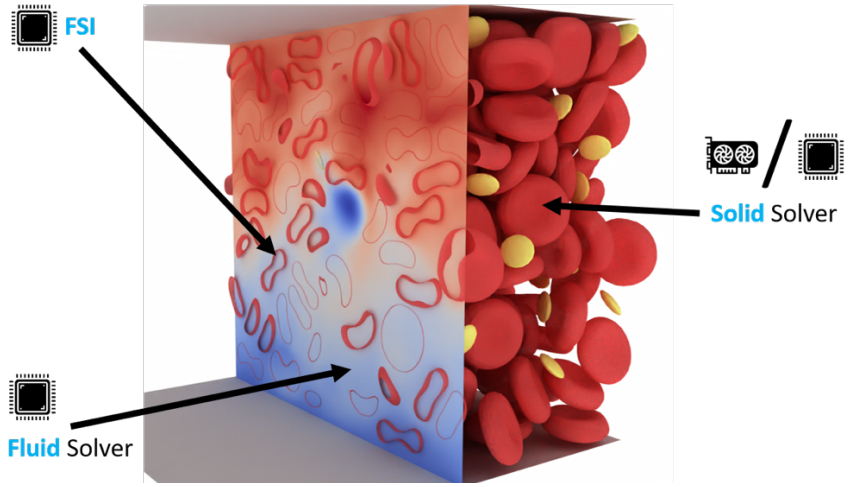

Figure 2.1.4-1 – Heterogeneous execution scheme of Palabos-npFEM.

Palabos-npFEM has been successfully deployed for the study of scientific problems from fundamental research to problems of clinical relevance. In collaboration with the Laboratory of Experimental medicine (Centre Hospitalier Universitaire (CHU) de Charleroi), we have studied the effect of red blood cell (RBC) shape-change (induced by chronic obstructive pulmonary disease-COPD) on platelet transport [2.1.4/3] (see Fig. 2.1.4-2). We have shown (via *in vitro*

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

experiments) that in COPD patients, platelets are transported faster towards the vessel walls, and following we used the *in silico* experiments to relate this observation with the RBC shape-change, ruling out the connection with the change of RBC electronegativity. This collaborative project shows the excellent complementarity between experimentations and numerical simulations to explore complex dynamic systems.

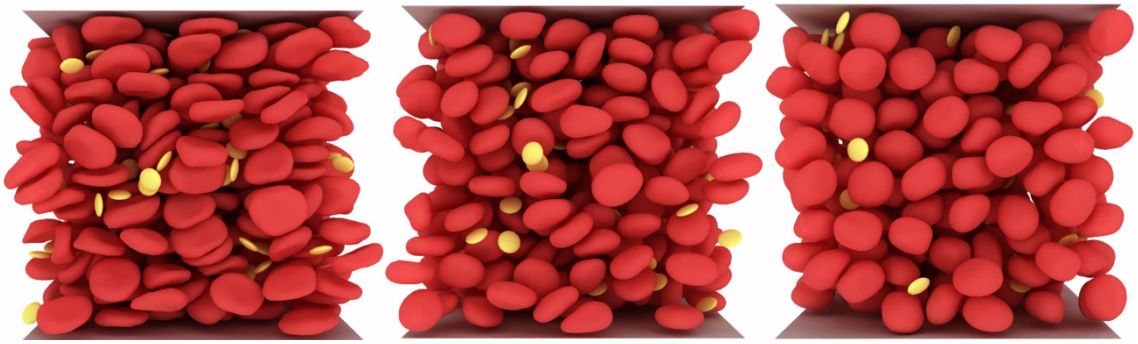

Figure 2.1.4-2 – Simulation snapshots from Palabos-npFEM. From left to right, normally shaped RBCs, spherized RBCs due to COPD (two spherized versions).

In terms of fundamental research, we proposed a disruptive view of platelet transport physics with immediate impact on devising accurate platelet function analysers (high importance in clinical practice). In more details, the transport of platelets in blood is commonly assumed to obey an advection-diffusion equation. We proposed [2.1.4/4] a disruptive view, by showing that the random part of platelets velocity is governed by a fat-tailed probability distribution, usually referred to as a Lévy flight. Although for small spatio-temporal scales, it is hard to distinguish it from the generally accepted "red blood cell enhanced" Brownian motion, for larger systems this effect is dramatic as the standard approach may underestimate the flux of platelets by several orders of magnitude, compromising in particular the validity of current platelet function tests.

Towards the direction of introducing models of higher fidelity, we are currently working on a more efficient and accurate fluid-solid interaction scheme, where our novel method exhibits second order accuracy (versus the first order accuracy of the immersed boundary method currently used) and a GPU-friendly form [publication in preparation].

Given the current trend of supercomputers to use accelerators, e.g. GPUs, we are building in parallel to Palabos-npFEM a GPU-only version. The new library, called CuBoltz, is capable of simulating multiphase flows in a multi-GPU environment, delivering impressive performance compared to the CPU counterpart. The framework will soon be open-sourced [publication in preparation].

Our research endeavours consist of using the aforementioned computational frameworks in collaborative projects with clinicians and a constant effort to target upcoming exascale supercomputers.

Software stack description

***Main code*:** Palabos-npFEM (https://gitlab.com/unigespc/palabos). The npFEM library is integrated into the Palabos core library as a coupled simulator [software metapaper submitted]. We provide a fully documented library, with installation instructions and multiple examples for blood flow simulations under various setups.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

***Components*:** Palabos (Lattice Boltzmann Method) for the resolution of blood plasma and npFEM (Finite Elements Method) for the resolution of the trajectories/deformations of blood cells. The two components are coupled in a modular way (plug and play approach).

***Libraries*:** MPI, CUDA

***Computational pattern(s)*:** Heterogeneous Multiscale Computing (HMM)

Current scaling information and outlook: Palabos utilizes CPU-only hardware, while npFEM supports both CPU & GPU execution. In more details, the deformable blood cells can be resolved either on CPUs (along Palabos-fluid resolution) or on GPUs using the CUDA-capable branch of the code. The code presents great scaling for up to O(500) GPUs and O(5000) CPU cores, while a complete performance benchmark can be found in [2.1.4/2].

Current deployment status (list of HPCs): Palabos-npFEM is currently deployed at the following HPC centres: Swiss National Supercomputing Centre (CSCS, Piz Daint), National Supercomputing Centre in the Netherlands (Surfsara, Cartesius), and the HPC Facilities of the University of Geneva (Baobab cluster).

6.1.5 Subtask 2.1.5 – Developing modelling techniques for in silico trials

Aim

Within this subtask, an important number of developments have occurred within the last few months that enables Alya to be employed and extended towards the advance of the state-of-the-art in simulations in the direction of the objectives of this work package. Categorised by the specific objectives:

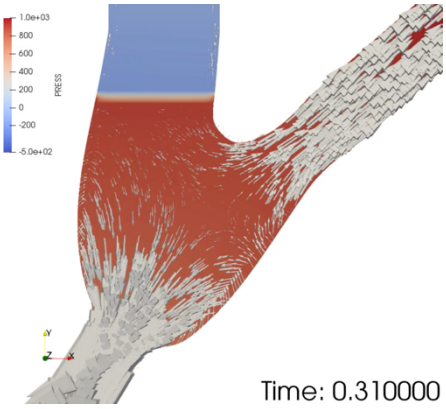

Figure 2.1.5-1 LVAD geometry and simulation results

- To develop verification, validation and uncertainty quantification techniques that will permit CompBioMed2 applications to be used in *in silico* clinical trials and personal health forecasting applications;
- To engage with the broader CompBioMed community, advising on deployment of codes, need for parallelization and development of workflows on a range of HPC platforms;
- To prepare appropriate and high impact CompBioMed2 applications for the emerging exascale;

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

The road to exascale not only implies to enable Alya for millions of parallel processes, but also to develop and implement efficient algorithms, at the sequential and parallel levels. During the past years, we have been focusing on the core of the simulation, mainly consisting in solving partial differential equations using numerical methods. Over the last years, we have broadened our expertise to implement pre and post process techniques to achieve a fully parallel simulation code, embracing all the required steps of a complete simulation cycle. The next figure (Fig. 2.1.5-2) summarizes the different topics we are working on to achieve efficient exascale simulations.

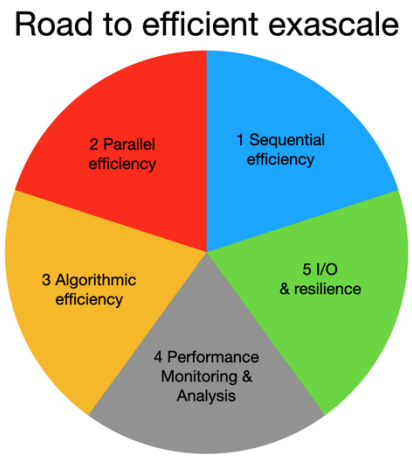

Figure 2.1.5-2 Domains of improvements to increase efficiency towards exascale

a) **Sequential efficiency**

Sequential efficiency refers here to the good usage of an individual core by an algorithm. In this aim, we have worked both with CPUs and GPU accelerators. On the one hand, for CPU, we have developed a specific matrix and right-hand side assembly technique to enable vectorisation as well as better memory access [2.1.5/1, 2.1.5/2]. Another important development consists in implementing Adaptive Mesh Refinement strategies in order to optimise the number of elements required by a simulation to obtain a given error threshold (PRACE6IP WP8 project: ParSec).

On the other hand, this assembly has also been ported to GPU. This has been carried out in collaboration with IBM. The porting to FPGA has been carried out in the context of LEGATO project. Iterative algebraic solvers have also been ported to GPU to enable co-execution as well as some specific kernels to FPGA.

At the highest level, one of the directions of development of Alya is the introduction of Auto-tuning capabilities [2.1.5/3]. The robustness of a simulation code on an exascale supercomputer requires automatic runtime adaptivity to both the computing and numerical conditions which become unpredictable at such scale. Examples of it can be the decision of the vectorisation length on different architectures, the decision of the solver parameters, the adaptation of the mesh or, at a parallel level, the distribution of resources. Huge amount of resources can be saved by automatic decisions based on in situ measurements, included in the software.

b) **Parallel efficiency**

For parallel efficiency, four different directions have been considered. In the exascale era, the overall simulation workflow needs to be parallelised. Any sequential part would otherwise become a critical bottleneck in terms of computing or memory resources requirements. For this reason, we have worked on the parallelization of the overall simulation workflow including the pre- and post- processing stages. In this regard, highly parallel geometric mesh partitioning strategies have been developed and tested on the domain decomposition of meshes of billions of elements [2.1.5/4, 2.1.5/5]. These algorithms are a building block of the inter-node load balancing in-situ process.

The second direction considers the efficient use of resources. Firstly, load balance strategies have been implemented both at the intra node and extra node levels, through the use of DLB

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

library in the context of an MPI+X paradigm [2.1.5/6] and redistribution techniques, respectively. Secondly, to further enhance the use of resources in the context of an heterogeneous computer, we have developed a co-execution strategy [2.1.5/1], where some MPI tasks are pinned to CPU cores and others to GPUs. Together with the dynamic load balance strategy, these techniques enable one to use almost all the available resources offered by a computing node. In the context of multiphysics coupling, different uses of DLB library have been proposed, for example to enable asynchronous particle tracking in a fluid [2.1.5/7], or to enhance the parallel efficiency of Gauss-Seidel algorithm in the context of fluid structure interactions [2.1.5/8, 2.1.5/9]. Part of this work has been carried out in the context of CompBioMed project.

The third direction consists in developing an automatic and runtime mechanism for selecting the number of cores required by a simulation and a given objective. Different objectives have been implemented: a minimum parallel efficiency to ensure that the application is using the available resources efficiently, and a timing objective to select automatically the number of cores to achieve the end of the simulation before the maximum queue time is reached. This work is being done in the context of a BSC internal project, CAT.

The fourth direction consists in selecting better iterative solvers. For strong scalability purposes, we have been working on communication hiding solvers (e.g. pipelined conjugate gradient). For weak scalability purposes, we have been implementing external solvers, mainly in the context of EOCOE and EOCOE-II projects (AGMS, Maphys, PSBLAS, MUMPS, etc.).

c) **Algorithmic efficiency**

To solve a given problem, usually many options are available. The selection of the most efficient algorithm depends on the required accuracy, the cost in terms of computational resources and the computational resources available to users. Therefore, algorithmic efficiency is intimately linked to sequential and parallel efficiency. The idea is to find the best algorithm for a given problem while maintaining the sequential and parallel efficiency at a desired threshold.

In the context of domain decomposition for coupling purpose, an original implicit coupling has been developed to enable the solution of moving parts in a fluid, without altering the convergence due to the coupling. This implicit coupling is achieved via a generalisation of the parallel sparse-matrix vector product [2.1.5/10, 2.1.5/11]. In the context of multiphysics coupling, the main algorithms are the Jacobi and Gauss-Seidel methods. While the Gauss-Seidel method has a better algorithmic efficiency than the Jacobi method, it suffers from very poor parallel efficiency. In this case, we decided in [2.1.5/8] to act on the implementation side to enhance the parallel efficiency of the Gauss-Seidel method to make it as good as that of the Jacobi method.

d) **Performance monitoring and analysis**

A continuous performance analysis of the code is necessary to know where the bottlenecks are and to ensure that the code is evolving in the right direction. For this, continuous performance analysis is being carried out. Firstly, we have monitored the code to measure runtime load balance efficiency and communication efficiency measures using TALP library. Together with different timings of the main kernels of Alya, we are developing a performance suite to visualisation the performance, monitor its history and provide alarms when the performance degrades.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

Also, different performance analyses are being carried out in the context of the UEABS of PRACE [2.1.5/12], PoP project [2.1.5/13] as well as with SURFsara in CompBioMed where an automated test suite based on Reframe has been implemented.

e) **I/O & resilience**

An efficient parallel I/O module based on MPIO has been developed for the parallel I/O operations in Alya. In the context of the Exellerate CoE this tool has proved its performance up to hundreds of thousands of MPI-processes on meshes of billions of nodes. The I/O operations are required at the beginning of the execution to load the problem data (geometry description, initial conditions, simulation parameters, etc) and during the simulation process for checkpointing and storing of results. We have developed a redistribution module that allows change the mesh partition during the simulation, this can be used for load balancing purposes or to change the amount of resources employed when a simulation is restarted. *In situ* processing of results is a key aspect to avoid massive requirements of disk memory, nonetheless the post-processing tools of Alya can accept data generated on different partitions or geometric discretisations. Finally, currently we are integrating the FTI library into Alya for I/O operations. FTI supports asynchronous I/O operations at different levels and includes resilience tools. This has been done during the EOCOE-II project

Software stack description

***Main code*:** Alya

***Components*:** High performance computing, multi-scale, multi-physics finite element solver for cardiac and vascular applications

***Libraries*:** Metis, MPI.

***Computational pattern(s)*:** Most executions are monolithic, which belongs to the "Extreme Scaling" pattern. Alya multiphysics coupling can also be executed as "Heterogeneous Multiscale Computing", by running various instances of Alya to solve a coupled multiphysics model.

Current scaling information and outlook: Scalability to 100K cores was shown in: Mariano Vazquez, et al. Alya: Towards Exascale for Engineering Simulation Codes. pages 1–20, 2014.

Current deployment status (list of HPCs): Alya is currently deployed in various HPC centres in the world including: Marenostrum, Barcelona; Archer, UK, Cartesius, Netherlands, Joliot-Curie Rome, France, SuperMuc, Germany.

6.2 Task 2.2 – Molecularly-based Medicine Research Exemplar

6.2.1 Subtask 2.2.1 – Testing virtual candidate libraries against novel targets

Aim

Virtual screening is a computational method which can be used to prioritize a set of compounds coming from a large library based on a metric of choice, most typically, the predicted affinity of

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

these compounds to a given enzyme or receptor (target). One of the most popular approaches in virtual screening is molecular docking, a computational method which tries to predict the binding mode of a compound in the binding pocket of the target of interest, together with its affinity.

Lately, we have developed and published SkeleDock [2.2.1/1], an algorithm that deals with docking scenarios where the binding mode of a fragment of the molecule of interest is known. Due to the popularity of the Fragment Based Drug Discovery (FBDD) paradigm, this scenario is rather common, and, because the binding mode of similar molecules is usually conserved, one can use the binding mode of a fragment to predict the binding mode of a molecule containing that same fragment.

SkeleDock was tested in the D3R GC4 challenge, a blind competition where the objective was predicting the binding mode and affinity of a series of compounds. A protocol based on SkeleDock and KDeep [2.2.1/2] (a 3D convolutional neural network) achieved competitive performance in several sub-challenges, and was the best performer in two of them, showing the advantage of using previous structural information to bias the pose prediction.

Software stack description

***Main code*:** SkeleDock

***Components*:** Scaffold docking code coupled with rDock**.**

***Libraries*:** HTMD, RDKit, rDock

Computational pattern(s): Extreme scaling.

Current scaling information and outlook: The scalability of SkeleDock across multiple cores was studied in its main article [2.2.1/1], showing linear improvement with the number of cores.

Current deployment status (list of HPCs): SkeleDock is only deployed on Acellera's HPC resources, but made available to the public trough a web application (https://playmolecule.org/SkeleDock/)

6.2.2 Subtask 2.2.2 – Assessing sequence influence on drug binding using free energy calculations

Aim

The aim of this subtask is to rank binding affinities of small molecules to their target proteins. We use binding affinity calculator (BAC) for free energy calculations, in which two protocols have been introduced: "enhanced sampling of molecular dynamics with approximation of continuum solvent" (ESMACS) and "thermodynamic integration with enhanced sampling" (TIES). To deliver an accurate ranking of drug binding affinities quickly and reproducibly, ensemble-based approaches have been applied. The automatic workflow reduces the burden of coordination of the many tasks in the high-throughput binding affinity calculator. The ability to perform such

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

calculations fast and reliably is beginning to have an important impact on the use of such methods in areas such as drug discovery and clinical decision making.

We have used the ESMACS approach to investigate drug-like small molecules bound to different therapeutic targets. We have recently studied a series of ligands binding to G protein-coupled receptors (GPCR), the most frequently exploited drug target class. The findings are published in [2.2.2/1]. We have assessed and evaluated the ensemble-based approaches in the field of molecular dynamics simulations [2.2.2/2]. We also focus on the software infrastructure and interfaces, and have automated the overall workflow and executed it on commodity cloud platforms [2.2.2/3]. The results show that BAC is able to reliably predict their binding affinities on time scales relevant to the domains of application.

TIES is a more accurate method for calculation of relative binding free energies, but it uses significantly more computational resources. We improved the underlying superimposition algorithm in TIES to utilise the transformation of partial rings. In effect, we reduced the sampling space in molecules where large regions were a part of the alchemical region, improving the reliability of the results while using fewer computational cycles. We validated our approach by computing transformations of around 100 ligands across 5 proteins obtaining on average a very good agreement with the experimental data. The mean signed error was less than 1 kcal/mol. We measured how the confidence intervals are affected by the number of samples taken (ensemble size), showing how the use of sufficient sampling is necessary in order to obtain reliable and reproducible results [publication in preparation].

The currently ongoing efforts include large scale application of our binding affinity calculator to a vast number of compounds for key protein targets in COVID-19. We are predicting binding properties of a set of compounds, consisting of drugs currently on the market and undergoing clinical trials, compounds from chemical libraries, and from static or dynamic virtual libraries generated from machine learning approaches. To improve the computational efficiency, we have used job farming approaches to run multiple simulations in parallel. In addition, for flexible task-level parallelism we also employ RADICAL-Cybertools, which is a powerful middleware environment which provides the ability to create and execute ensemble-based applications with complex coordination and communication abstraction, and to support scalable and efficient launching of heterogeneous tasks across different platforms.

Software stack description

***Main code*:** NAMD, OpenMM, AmberTools

***Components*:** NAMD and OpenMM as simulation engines, and MMPBSA module of AmberTools for free energy estimation

***Libraries*:** MPI

***Computational pattern(s)*:** Simulations are executed as the “Replica Computing” pattern as there are a large number of independent runs with no/limited information exchanges.

Current scaling information and outlook: NAMD scales well on CPUs up to thousands of cores and a recent version with GPU support will soon leave the alpha-version state, while OpenMM already achieves a high performance through GPUs. Execution of an ensemble simulation is an embarrassingly parallel workload. A big job, consisting of many independent

*This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement No **823712**“*

jobs, will run concurrently using a single submission. The performance of the big job will be identical to each individual run.

Current deployment status (list of HPCs): BAC has been deployed at several HPC centres around the world. Recently we mainly use SuperMUC-NG (Germany), Summit (USA), Frontera (USA), Longhorn (USA), Hartree (UK), Cartesius (Netherlands), Archer (UK).

6.2.3 Subtask 2.2.3 – Combining machine learning and molecular modelling to assess drug binding

Aim

Molecular simulations represent an important scientific and technological methodology to understand macroscopic phenomena from a microscopic point of view. Technologies facilitating molecular dynamics (MD) simulations, such as distributed computing and GPU MD software, have made great strides in terms of the time- and length-scales accessible *in silico*. However, even the longest protein simulations still fail to reach total times exceeding milliseconds, and dedicated analysis methods are required to infer dynamics at longer timescales [2.2.3/1]. In this subtask, we aim to solve the sampling issues of MD simulations by combining it with state-of-the-art machine learning techniques.

In order to make this combination possible, we have developed torchMD, a new molecular simulation application based on the high-performance deep learning python library PyTorch [2.2.3/2]. TorchMD allows us to perform all the operations to train, simulate and analyze a machine learned potential and its potential priors using a PyTorch backend. All computations are expressed as Torch arrays, including bonds, angles, dihedrals, Lennard-Jones and Coulomb interactions. Also, it is possible to run standard all-atom simulations using the Amber forcefield. We expect TorchMD to be very useful in the more recent applications where simulations and machine learning models are used.

One example application of such code is to be able to learn coarse-grained models of atomistic systems. As part of task 2.2.3, we have started applying TorchMD specifically in the case of protein folding and coarse-grained representations for two exemplar proteins. In the present contribution, we focus on the coarse graining model formulated as a supervised machine learning problem to reproduce the thermodynamics of small biomolecular systems. Particularly, we modify the architecture of the recently introduced CGnet framework [2.2.3/3] such that the molecular features it requires are learned instead of hand-selected as in the original formulation. By leveraging the inherently transferable SchNet scheme [2.2.3/4] to learn features, we render the entire CGnet framework transferable across molecular systems. SchNet is a scalable, transferable framework that employs neural networks and representation learning to predict the properties and behavior of small organic molecules. The SchNet architecture is implemented in TorchMD through the SchNetPack library [2.2.3/5]. In addition, TorchMD also makes use of PyTorch Lightning [2.2.3/6], a Python library that provides a high-level interface for PyTorch that makes the neural network training code in TorchMD scalable.

The ongoing work on this task is centered around reproducing the thermodynamics of previous protein folding MD simulations, in order to prove the accuracy of our coarse-grained approach. By training (and subsequent simulation) of a neural-network based coarse-grained potential with TorchMD, we have been able to successfully reproduce folding and unfolding events for

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

several fast-folding proteins, and the current efforts are focused on tackling more complicated proteins.

This represents only one possible application of TorchMD and its unique interface between machine learning and molecular simulations can play an important role in facilitating research in the field. The code is available open-source at github.com/torchmd/torchmd.

Software stack description

***Main code*:** TorchMD

Components: PyTorch-based simulation code plus a Schnet-based neural network code

Libraries: PyTorch, PyTorch-Lightning, SchNetPack

Computational pattern(s): Simulation:Replica Computing / Training: Extreme Scaling (scales with training data)

Current scaling information and outlook: Simulation code works on a single GPU (Replica computing for scaling, depends on application). Training code scales infinitely with number of GPUs

Current deployment status (list of HPCs): Code is deployed on UPF's local cluster. Available on Github (github.com/torchmd/torchmd)

6.3 Task 2.3 – Neuro-musculoskeletal Research Exemplar

6.3.1 Introduction

All the subtasks in this task contribute to the development of a single solution, aimed to predict the risk of hip fracture in patients affected by osteoporosis (digital patient technologies), and to use this information to estimate if a new drug treatment is more or less effective than those already available (using *in silico* trials technologies).

In the following we will provide the aim for each subtask, but a single software stack description, providing details for the various versions that will emerge by the completion of each task. We will assume the following versioning scheme:

- Alpha v1: ARF0 predictor (standard CT2S code)
- Alpha v2: ARF0 + growth and adaptation code
- Alpha v3: ARF10 predictor (absolute risk over 10 years, final digital patient solution)
- Alpha v4: ARF10+ Virtual cohort (final in silico trials solution)

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

6.3.2 Subtask 2.3.1 – Growth and Adaptation

Aim

The goal of this subtask is to create an efficient model to describe bone growth and adaptation, also in response to pharmaceutical treatments.

As a first step, volumetric bone mineral content (vBMD) loss due to ageing has been modeled, allowing the creation of virtual Computerised Tomography (CT) scans for future years, starting from the CT scan at time 0. Given areal bone mineral density (aBMD) and patient age at time 0, changes in aBMD up to year 10 are estimated from population-based aBMD decline data. The vBMD-specific, vBMD decline rate with respect to aBMD is obtained by regressing the 96 proximal femur CT data set with aBMD. However, we restrict to only three slopes of regression. The slopes apply to three vBMD ranges, separated by two vBMD threshold values. One separates low and medium values, while the other separates medium and high values. The above model has now been integrated into the Algorithmic Processor Description Language (APDL) which is used to conduct CT2S analysis.

In ongoing work, we are developing a vBMD- and strain energy density (SED)- specific vBMD decline rate with aBMD as well, as schematized in figure below. It will be tested whether this new disease progression model significantly changes ARF10 predictions compared to the existing model for vBMD loss due to ageing (as described above).

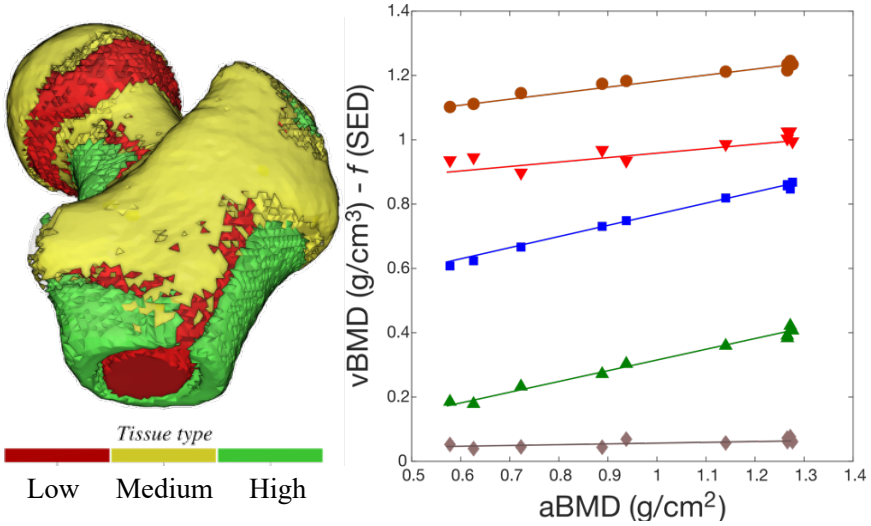

Figure 2.3.1-1 Left: Locations within the proximal femur consistently identified as either possessing low, medium or high vBMD. Right: Schematic of vBMD- and strain energy density (SED)- specific vBMD decline rate with respect to aBMD

Similarly, but from bottom up, tissue and cell-level data generated from agent-based models could also provide valuable information on bone remodelling; to inform and validate such models, HPC resources that could process large images, such as those generated from the synchrotron facility, are required.

We are also working on the creation of a scalable C/C++ library to map scalar and vector fields from a regular hexahedral grid to an unstructured tetrahedral mesh and vice versa, in order to

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

complete the CT2S pipeline automation, and make it possible to work with high resolution images and/or large anatomical region meshes.

6.3.3 Subtask 2.3.2 - ARF10: Multiscale model of hip fracture

Aim

The aim of this task is the creation of a multiscale model able to predict the clinically relevant risk of fracture at 10 years, given the following at time *t=0*: age, body height, body weight, proximal femur CT scan and femoral neck aBMD. To achieve the goal, two main tools are needed: a reliable fracture risk estimator at a given time, and an efficient disease progression model.

We firstly focused on the automation of the CT2S pipeline, allowing a complete patient simulation (non-linear finite-element (FE) models with different boundary and loading conditions), starting from a femur material-mapped mesh and a database containing patient biometric data; the simulation of a single fall angle (set of loading and boundary conditions) requires about 5 GB of RAM and one core/hour, so that for a 28 fall angles femur strength estimation around 30 core/hours are needed.

Second, the growth and adaptation pipeline, as described in Subtask 2.3.1, is being used to obtain virtual CT scans for years 1 through to 10 based on year 0 CT scan. To these virtual CT scans, the CT2S pipeline is applied to obtain femur strength distribution for 28 fall cases for years 1 through 10.

Third, a standalone pipeline called ARF10 has been created which modifies an existing ARF0 pipeline. The ARF0 pipeline takes patient body weight and height and computes a stochastic fall bone load magnitude distribution ($\geq 10^4$ samples). It then takes this load magnitude distribution, the distribution of femur strength for 28 fall angles (as provided by CT2S pipeline) and a rate of fall per year to evaluate femur fracture risk at year 0 [2.3.2/1]. The ARF10 pipeline modifies the process as follows. It repeatedly applies ARF0 pipeline to obtain from each year 1 – 10, a fracture risk using the above year-specific femur strength distribution, while assuming that body height, body weight and fall rate do not change with time. Once ARF0 is determined for years 1–10, it computes ARF10 using the combined probability principle.

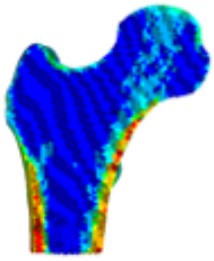

Figure 2.3.2-1 Heterogeneous elastic modulus distribution in a proximal femur, with red representing stiff cortical bones.

We are currently also working on VVUQ of the stratification accuracy of the absolute fracture risk at time 0 on a 96 postmenopausal women cohort. In parallel, we are also looking to link the bone strength model across different scales. For example, body level data from gait analysis can be used to inform force boundary conditions applied to the organ-level bone model [2.3.2/2],

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

and synchrotron/microCT-informed cell-level agent-based models can provide information on bone remodelling.

In addition, we are exploring other high-end multiscale modelling methods, such as HMM (heterogeneous multiscale modelling), which is scalable in a massively parallel manner.

6.3.4 Subtask 2.3.3 – Virtual patients' expansion

Aim

This task focused on the generation of a virtual cohort in order to simulate a phase 3 clinical trial (more than 1000 patients involved). The type of simulation intrinsically requires the running of numerous embarrassingly parallel jobs, possibly representing several different disease or treatment conditions for each virtual patient.

The first proposed application is the generation of a PCA-based statistical population, trained with active shape and appearance models of 94 patients of a pair-matched postmenopausal women cohort. An article on the population PCA used for the virtual cohort generation is being submitted [2.3.3/1].

Leveraging the CT2S pipeline automation developed in Subtask 2.3.2, simulations of the 1000-patient cohort have been set up on Cartesius cluster, requiring about 30k core/hours for a single cohort run; the simulation architecture optimization work is ongoing. We are also planning to run a Markov process to describe the cohort fracture rate evolution in the clinically relevant 10 years period, using bone remodelling algorithms developed in Subtask 2.3.1.

As a further example of patients' expansion, a virtual population for the development of a tuberculosis vaccine (immune system model from STriTuVaD European project) will be simulated, leveraging the developed statistical tools.

Software stack description for Task 2.3

***Main code*:** CT2S/ARF10

***Components*:**
Alpha v1: ANSYS Mechanical APDL for finite element simulations, Python/Matlab codes (scripts already implemented in both languages) for load distribution generation, result gathering and data elaboration.
Alpha v2: Alpha v1 + ANSYS Mechanical APDL for vBMD update, Matlab for regression slope calculation.
Alpha v3: Alpha v2 + Matlab for ARF10 calculation
Alpha v4: Alpha v3 + Matlab code to generate the Virtual Cohort, Python code for cohort orchestration and aggregate data elaboration.

***Libraries*:** NumPy, SciPy, Pandas, Scikit-learn

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

***Computational pattern(s)*:** The vast majority of the CT2S workflow uses a "Replica Computing" pattern, requiring tens of independent finite-element models solution for each patient. The grow and adaptation code can be coupled for Multiscale Computing.

Current scaling information and outlook: The embarrassingly parallel simulations leverage ANSYS Shared Memory Parallel to distribute compute load across cores within one node for every patient simulation. For each patient nearly 30 independent falling conditions are simulated, each requiring about 1 core/hour, so that for ARF10 calculation of a single patient more than 300 cores/hours are needed. A complete phase-3 (≥ 1000 patients involved) *in silico* clinical trial is expected to require at least 300k core/hours. Due to the embarrassingly parallel nature of most of the simulations, we expect very good scaling for such patient cohort simulations.

Current deployment status (list of HPCs): The codes are deployed on ShARC (University of Sheffield, England) and Cartesius (SURFsara, Netherlands).

7 Risk Management

The following possible sources of risks have been identified:

a) The internal budget of HPC core-hours might not be enough to cover all the VVUQ analyses defined in Task 2.5.

Probability	Medium
Impact	Medium
Risk assessment	Medium
Mitigation	The progress of VVUQ application deployment is monitored continuously together with the still available internal budget, and the partners report on the advancements every month. For now, the consortium is evaluating possible ways to provide access to more core-hours for VVUQ, and partners are encouraged to also apply for external budgets individually. The first deliverable on VVUQ strategy (D2.3) will reassess the situation in detail at M24 and plan ahead accordingly.

b) Delays in the proposed work due to COVID-related reduced interactions (e.g. reduced mobility, such as on-site visits between partners).

Probability	Low
Impact	Low
Risk assessment	Low
Mitigation	All research tasks using direct experimental data or relying on lab access and mobility have been coping well with the current

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

	situation. A prompt migration to online collaborative tools and the surge in the availability of such tools has helped to avoid any delays so far. We continue to keep an eye on the situation as it evolves.

c) Insufficient or non-timely input from partners delaying the deliverables.

Probability	Low
Impact	Medium
Risk assessment	Medium
Mitigation	The progress of the deliverables is checked internally on a regular basis. The regular work package meetings and the intra-workpackage teleconferences and discussions facilitate the information exchange and allow the WP leader, or if needed the Project Manager, to step in and act to mitigate the problem.

8 Conclusions

This deliverable reported on the current application readiness and deployment of the fast track applications in the CompBioMed project. We report a high level of application readiness overall, as it was expected through the continuation of CompBioMed 1. The categorization based on the HPC execution scheme of every included code was also reported (see the information and references in Chpt. 5. for the definitions), along with the list of HPC centres where the code is currently deployed and used.

Besides the readiness level, several recent advancements in each individual research task have been briefly outlined and referenced, demonstrating the considerable progress made in the first 12 month of the project. Furthermore, the most probable risk factors were reported together with the planned steps to mitigate them if they were to happen.

9 Bibliography/References

Please make use of bibliography and references rather than footnotes, these should be checked rigorously before publication.

[Introduction/1] A. G. Hoekstra, B. Chopard, D. Coster, S. Portegies Zwart, P. V. Coveney, "Multiscale Computing for Science and Engineering in the Era of Exascale Performance", Phil Trans R Soc A, 377 (2019), DOI: 10.1098/rsta.2018.0144
[Introduction/2] Alowayyed, S., Groen, D., Coveney, P. V., & Hoekstra, A. G. (2017). Multiscale computing in the exascale era. Journal of Computational Science, 22, 15-25., DOI: 10.1016/j.jocs.2017.07.004
[2.1.1/1] https://pop-coe.eu/sites/default/files/pop_files/pop2-ar-hemelb.pdf

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

[2.1.1/2] McCullough, J. W. S., Richardson, R. A., Patronis, A., Halver, R., Marshall, R., Ruefenacht, M., Wylie, B. J. N., Odaker, T., Wiedemann, M., Lloyd, B., Neufeld, E., Sutmann, G., Skjellum, A., Kranzlmuller, D. and Coveney, P.V. (2020), Towards Blood Flow in the Virtual Human: Efficient Self-Coupling of HemeLB, J R Soc Interface Focus, In Press (Accepted).
[2.1.2/1] Zhou X, Qu YS, Passini E, Bueno-Orovio A, Liu Y, Vargas HM, Rodriguez B. (2020, January) Blinded In Silico Drug Trial Reveals the Minimum Set of Ion Channels for Torsades de Pointes Risk Assessment, Frontiers in Pharmacology, 10: 1-17.
[2.1.2/2] Martinez-Navarro H, Minchole A, Bueno-Orovio A, Rodriguez B. (2019). High arrhythmic risk in antero-septal acute myocardial ischemia is explained by increased transmural reentry occurrence. Scientific Reports, 9: 16803.
[2.1.2/3] Levrero-Florencioa F, Margaraa F, Zacurb E, Bueno-Orovioa A, Wang ZJ, Santiago A, Aguado-Sierrac J, Houzeauxc G, Graub V, Kaya D, Vázquezcd M, Ruiz-Baiere R, Rodrigueza B. (2020, April). Sensitivity analysis of a strongly-coupled human-based electromechanical cardiac model: effect of mechanical parameters on physiologically relevant biomarkers. Computer Methods in Applied Mechanics and Engineering, 361, 112762.
[2.1.2/4] Margara F, Wang ZJ, Levrero-Florencio F, Santiago A, Vázquez M, Bueno-Orovio A, Rodriguez B. (2020, July) In-silico human electro-mechanical ventricular modelling and simulation for drug-induced pro-arrhythmia and inotropic risk assessment, Progress in Biophysics and Molecular Biology, In Press. DOI: 10.1016/j.pbiomolbio.2020.06.007
[2.1.3/1] Czaja, B., Závodszky, G., & Hoekstra, A. (2020, June). A Heterogeneous Multi-scale Model for Blood Flow. In International Conference on Computational Science (pp. 403-409). Springer, Cham.
[2.1.3/2] Czaja, B., Gutierrez, M., Závodszky, G., de Kanter, D., Hoekstra, A., & Eniola-Adefeso, O. (2020). The influence of red blood cell deformability on hematocrit profiles and platelet margination. PLOS Computational Biology, 16(3), e1007716.
[2.1.3/3] de Vries, K., Nikishova, A., Czaja, B., Zavodszky, A., Hoekstra, A.. (2020) Inverse Uncertainty Quantification of a cell model using a Gaussian Process metamodel, International Journal for Uncertainty Quantification, In Press. DOI: 10.1615/Int.J.UncertaintyQuantification.2020033186
[2.1.3/4] Tarksalooyeh, V. A., Závodszky, G., & Hoekstra, A. G. (2019, June). Optimizing Parallel Performance of the Cell Based Blood Flow Simulation Software HemoCell. In International Conference on Computational Science (pp. 537-547). Springer, Cham.
[2.1.4/1] Kotsalos, Christos, Jonas Latt, and Bastien Chopard (2019). “Bridging the computational gap between mesoscopic and continuum modeling of red blood cells for fully resolved blood flow”. In: Journal of Computational Physics 398. doi:10.1016/j.jcp.2019.108905
[2.1.4/2] Kotsalos, Christos et al. (2020). “Digital Blood in Massively Parallel CPU/GPU Systems for the Study of Platelet Transport”. In: Interface Focus (accepted for publication) Computational Biomedicine. url: http://arxiv.org/abs/1911.03062.
[2.1.4/3] Boudjeltia, Karim Zouaoui et al. (2020). “Spherization of red blood cells and platelets margination in COPD patients.” In: Annals of the New York Academy of Sciences. doi: 10.1111/nyas.14489
[2.1.4/4] Kotsalos, Christos et al. (2020). Anomalous Platelet Transport & Fat-Tailed Distributions. arXiv:2006.11755.
[2.1.5/1] Borrell R, Dosimont D, Garcia-Gasulla M, Houzeaux G, Lehmkuhl O, Mehta V, et al. Heterogeneous CPU/GPU co-execution of CFD simulations on the POWER9 architecture: Application to airplane aerodynamics. Future Generation Computer Systems. 2020 Jun;107:31–48.

*This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement No **823712**“*

[2.1.5/2] Houzeaux G, Borrell R, Fournier Y, Garcia-Gasulla M, Göbbert JH, Hachem E, et al. High-Performance Computing: Dos and Don'ts. In: Ionescu A, editor. Computational Fluid Dynamics - Basic Instruments and Applications in Science. InTech; 2018.
[2.1.5/3] Kannan V, Borrell R, Doyle M, Houzeaux G. Tuning Alya for Energy Eciency with READEX. Zenodo. 2019;
[2.1.5/4] Borrell R, Oyarzun G, Dosimont D, Houzeaux G. Parallel SFC-based mesh partitioning and load balancing. 2019 IEEE/ACM 10th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems (ScalA). IEEE; 2019. p. 72–8.
[2.1.5/5] Borrell R, Cajas JC, Mira D, Taha A, Koric S, Vázquez M, et al. Parallel mesh partitioning based on space filling curves. Comput Fluids. 2018 Sep;173:264–72.
[2.1.5/6] Garcia-Gasulla M, Houzeaux G, Ferrer R, Artigues A, López V, Labarta J, et al. MPI+X: task-based parallelisation and dynamic load balance of finite element assembly. Int J Comut Fluid Dyn. 2019 Mar 16;33(3):115–36.
[2.1.5/7] Houzeaux G, Garcia M, Cajas JC, Artigues A, Olivares E, Labarta J, et al. Dynamic load balance applied to particle transport in fluids. Int J Comut Fluid Dyn. 2016 Jul 2;30(6):408–18.
[2.1.5/8] Houzeaux G, Garcia-Gasulla M, Cajas JC, Borrell R, Santiago A, Moulinec C, et al. Parallel multiphysics coupling: algorithmic and computational performances. Int J Comut Fluid Dyn. 2020 Jul 2;1–22.
[2.1.5/9] Cajas JC, Houzeaux G, Vázquez M, García M, Casoni E, Calmet H, et al. Fluid-Structure Interaction Based on HPC Multicode Coupling. SIAM J Sci Comput. 2018 Jan;40(6):C677–703.
[2.1.5/10] Houzeaux G, Borrell R, Cajas JC, Vázquez M. Extension of the parallel Sparse Matrix Vector Product (SpMV) for the implicit coupling of PDEs on non-matching meshes. Comput Fluids. 2018 Mar;173:216–25.
[2.1.5/11] Houzeaux G, Cajas JC, Discacciati M, Eguzkitza B, Gargallo-Peiró A, Rivero M, et al. Domain decomposition methods for domain composition purpose: chimera, overset, gluing and sliding mesh methods. Arch Computat Methods Eng. 2016 Nov 9;1–38.
[2.1.5/12] Lioen W, Avillez M, Codreanu V, Dellis D, Dolas S, Emerson A, et al. PRACE-5IP benchmarking deliverable D7.5 "Evaluation of Accelerated and Non-accelerated Benchmarks." 2019 Apr.
[2.1.5/13] Garcia-Gasulla M, Banchelli F, Peiro K, Ramirez-Gargallo G, Houzeaux G, Ben Hassan Saïdi I, et al. A generic performance analysis technique applied to different CFD methods for HPC. Int J Comut Fluid Dyn. 2020 Jul 16;1–21.
[2.2.1/1] Varela-Rial, Alejandro, Maciej Majewski, Alberto Cuzzolin, Gerard Martinez-Rosell, and Gianni De Fabritiis. "SkeleDock: A Web Application for Scaffold Docking in PlayMolecule." Journal of Chemical Information and Modeling (2020).
[2.2.1/2] Jiménez, José, Miha Skalic, Gerard Martinez-Rosell, and Gianni De Fabritiis. "K deep: Protein–ligand absolute binding affinity prediction via 3d-convolutional neural networks." Journal of chemical information and modeling 58, no. 2 (2018): 287-296.
[2.2.2/1] Wan, S., Potterton, A., Husseini, F. S., Wright, D. W., Heifetz, A., Malawski, M., Townsend-Nicholson, A., Coveney, P. V. (2020). Hit-to-lead and lead optimization binding free energy calculations for G protein-coupled receptors. Interface Focus, 20190128. DOI: 10.1098/rsfs.2019.0128
[2.2.2/2] Wan, S., Bhati, A. P., Zasada, S. J., Coveney, P. V. (2020). Rapid, accurate, precise and reproducible ligand-protein binding free energy prediction. Interface Focus, in press.
[2.2.2/3] Zasada, S. J., Wright, D. W., Coveney, P. V. (2020). Large-scale binding affinity calculations on commodity compute clouds. Interface Focus 20190133. DOI: 10.1098/rsfs.2019.0133
[2.2.3/1] Nuria Plattner, Stefan Doerr, Gianni De Fabritiis, and Frank Noé. Complete protein–protein association kinetics in atomic detail revealed by molecular dynamics simulations and markov modelling. Nature chemistry, 9(10):1005, 2017.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*

[2.2.3/2]Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Desmaison, A. Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems, pp. 8026-8037, 2019.
[2.2.3/3] Jiang Wang, Simon Olsson, Christoph Wehmeyer, Adrià Pérez, Nicholas E. Charron, Gianni De Fabritiis, Frank Noé and Cecillia Clementi. Machine learning of coarse-grained molecular dynamics force fields. ACS Cent.Sci. 5(5) 755-767, 2019.
[2.2.3/4]K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, "SchNet–a deep learning architecture for molecules and materials," J. Chem. Phys. 148, 241722, 2018.
[2.2.3/5]Schutt, K. T., Kessel, P., Gastegger, M., Nicoli, K. A., Tkatchenko, A., & Müller, K. R. SchNetPack: A deep learning toolbox for atomistic systems. Journal of chemical theory and computation, 15(1), 448-455, 2018.
[2.2.3/6] Falcon, W. E. A. "Pytorch lightning." GitHub. Note: https://github. com/williamFalcon/pytorch-lightning, 2019.
[2.3.2/1] Bhattacharya, P., Altai, Z., Qasim, M. *et al.* A multiscale model to predict current absolute risk of femoral fracture in a postmenopausal population. *Biomech Model Mechanobiol* **18,** 301–318 (2019). https://doi.org/10.1007/s10237-018-1081-0
[2.3.2/2] Altai, Z., Montefiori, E., Veen, van B., Paggiosi, A. M., McCloskey, V. E., Viceconti, M., Mazzà, C., Li, X. (2020) Femoral neck strain prediction during level walking using a combined musculoskeletal and finite element model approach. Submitted to PLOS ONE.
[2.3.3/1] Taylor, M., Viceconti, M, Bhattacharya, P., and Li, X. (2020) Finite element analysis informed variable selection for femoral neck fracture risk prediction. In preparation.

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **823712**"*