Identification/Selection of E-CAM meso and multi-scale modelling codes for development

E-CAM Deliverable 4.1
Deliverable Type: Report
Delivered in Month 6– April 2016

E-CAM
The European Centre of Excellence for
Software, Training and Consultancy
in Simulation and Modelling

Funded by the European Union under grant agreement 676531

Project and Deliverable Information

Project Title	E-CAM: An e-infrastructure for software, training and discussion in simulation and modelling
Project Ref.	Grant Agreement 676531
Project Website	https://www.e-cam2020.eu
EC Project Officer	Dimitrios Axiotis
Deliverable ID	D4.1
Deliverable Nature	Report
Dissemination Level	Public
Contractual Date of Delivery	Project Month 6(April 2016)
Actual Date of Delivery	31.10.2016

Document Control Information

Document	Title:	Identification/Selection of E-CAM meso and multi-scale modelling codes for development
	ID:	D4.1
	Version:	As of October 27, 2016
	Status:	Accepted by WP leader
	Available at:	https://www.e-cam2020.eu/deliverables
	Document history:	Internal Project Management Link
Review	Review Status:	Reviewed
	Action Requested:	Submit
Authorship	Written by:	Ignacio Pagonabarraga(University of Barcelona)
	Contributors:	B. Dünweg and T. Stühn (Mainz), V. Lobaskin and D. MacKernan (NUI UCD), G. Sutmann (Jülich), C. Hartmann and L. Delle Site (Berlin), and L. Petit and M. Seaton (STFC)
	Reviewed by:	D. Tildesley (EPFL)
	Approved by:	I.Pagonabarraga (UB)

Document Keywords

Keywords:	E-CAM, HPC, CECAM, Materials, ...

October 27, 2016

***Disclaimer**:This deliverable has been prepared by the responsible Work Package of the Project in accordance with the Consortium Agreement and the Grant Agreement. It solely reflects the opinion of the parties to such agreements on a collective basis in the context of the Project and to the extent foreseen in such agreements.*

[1] ipagonabarraga@ub.edu

Contents

List of Figures

Executive Summary

The present report has analyzed a variety of existing modeling approaches to analyze the equilibrium and non-equilibrium properties of complex systems at a coarse grained and multiscale level.

The analysis has shown the viability to exploit existing expertise within E-CAM and use this expertise to select a number of multi scale and coarse grained codes. Developing these codes, and producing new modules associated to them and that can also work transversally with other computational packages provides a fruitful perspective to make progress.

E-CAM activities, such as the state of the art workshops, will constitute the natural forum to expose and confront the selected codes to complementary approaches. The outcome of such activities will help to decide the need to enlarge the palette of codes and/or promote the development of transverse modules that can interface selected codes with complementary coarse-grained and multiscale software packages.

The report also shows that all existing codes are amenable to be run on supercomputing environments. We will exploit E-CAM activities to probe more accurately their scalability within Partnership for Advanced Computing in Europe (PRACE). These scaling tests will serve to provide the community with a deeper understanding of the potential of the software within High Performance Computing (HPC). Not all codes are equally supported on CPUs and GPUs. We identify a general need to be able to use both type of hardware. This fact will be taken into account when developing modules. The report shows that the selected codes will be developed for specific new uses according to industrial interest. The outcomes of these projects in terms of modules associated to the mentioned codes will be of wider use.

1 Introduction

The Work Package 4 (WP4) of the E-CAM Centre of Excellence focuses on support, development, optimization and parallelization of software components and modules in the field of mesoscopic and multiscale methods mainly in materials science, soft matter and biophysics applications. Due to the diversity of time and length scales and a broad field of applications, there has a number of different codes been developed, sometimes centered around new developments of algorithms. Within the consortium of E-CAM there exist a number of codes in the field of mesoscopic methods, which will be applied and further developed. Furthermore, modules which are developed within E-CAM will be mainly tested along existing software from the project. To get on overview of the landscape of existing software, WP4 has also carried out a search for other existing software and computational approaches to deal with materials at the mesoscale. In particular, we have evaluated existing expertise in Kinetic Monte Carlo (KMC) and Dynamic mean-field density functional theory (MESODYN) [1], as well as alternative computational approaches based on coupling phase fields, following the general principles of dynamic density functional theory [2].

KMC is a Monte Carlo simulation method (MC) method that was developed to analyze the dynamics in systems characterized by energetic barriers [3]. As such, it includes both the difference in energy between different states and the energy barriers that atoms/molecules have to overcome to move form one state to another. KMC has been a very useful method to study a variety of physico-chemical proceses and regimes in solids, where atom motion is characterized by the presence of energy barriers. Surface diffusion, surface growth and defect mobility are examples of relevant phenomena where KMC has shown its capabilities. KMC requires as input the corresponding rates associate to the barriers. If the transitions rates satisfy detailed balance, KMC can recover the appropriate thermodynamic behavior of the system. However, since the rates are input parameters, if detailed balance is broken KMC can also be used to analyze kinetic processes far from equilibrium, such as heterogeneous catalysis and chemical reactions [4], nucleation processes on solids, or coarsening. The jump rates that characterize KMC can be postulated or derived from microscopic or *ab-initio* simulations. Therefore, KMC stands as a natural candidate for a mesoscopic or multiscale method that focuses on the time scales of atom diffusion or reaction but with atomistic detail encoded in the transition rates (derived from the corresponding energy barriers). KMC can combine easily with results from WP1, WP2 and WP3 in ECAM, and this is the reason why it was mentioned as one of the potential methods to be developed within WP4. KMC has been combined with electronic structure calculations to derive atomistically quantitative jump rates for solid diffusion of adatoms [5]. In particular, KMC has also been used to study viscoelasticity in polymer melts and gels [6], one of the areas where E-CAM teams have experience in mesoscopic modeling and of potential interest to some of the industrial partner.

MESODYN is a computational mesoscopic model based on a dynamical mean field density functional theory for polymers [7]. The molecular structure is captured by treating the polymer architecture and the interaction among chains from a Self-Consistent Field Theory (SCFT). In this way it is possible to capture the specificity of the polymer chains that constitute a polymer suspension, melt or gels. MESODYN embeds this SCFT into a Model B dynamics that accounts for the conservation of diffusing chains. SCFT is the most accurate theory that, at mean field level, can provide configurational information on polymeric systems keeping topological information about the polymer chains and also details about chain length and segment interaction (hence accounting for chemical specificity). MESODYN has proved to be a flexible computational tool that has allowed to study the dynamics of complex polymeric materials at the mesoscale, such as the kinetics of a variety of block copolymers [8] in and out of equilibrium. The SCFT also allows easily to address more complex scenarios, where polymers interact with generic solid surfaces. This flexibility has been exploited to study thin film block copolymers, mapping the complex phase diagram that characterizes these systems [9]. MESODYN has been a valuable tool because long simulations are required to sample the morphological changes and wide relaxation time spectrum of block copolymer melts under confinement. MESODYN has been the core of the development of Chemistry Unified Language Interface (CULGI) B.V. [10], an international IT services company, specializing in computational chemistry, data management and lab automation. CULGI B.V. unifies a diverse set of computational tools in one platform. In particular, it has extended the dynamic self consistent treatment of polymer chains to extend the molecular information required in some applications.

These are methods that have a value on their own, and are complementary of other mesoscospic methods, such as Dissipative Particle Dynamics (DPD) [11], Lattice Boltzmann (LB) [12], Multiparticle Collision Dynamics (MPC) [13]. In fact, in some cases there is an intimate underlying conceptual connection. This is the case, for example, between some approaches to treat complex fluids within LB, based on free energy functionals, and Dynamic Functional Theory (DFT) or phase field models. These approaches nowadays can deal with polymeric systems in a degree of detail comparable to that captured by MESODYN, with the advantage that they treat appropriate the hydrodynamic coupling. There are also computational emerging strategies that take advantage of the advances in dealing with Big Data, or the learning ability of neural networks, to design new schemes to develop force fields, effective interactions and coarse grained models. We have identified that the use of nerual networks w can take place more naturally in the context of material exploration in WP1. Neural network derived potentials, for example, will develop more naturally at the boundary between *ab-initio* potential calculations and atomistic force fields. We expect then these approaches

to develop in t connection with collaborations between WP1 and WP2.

All these different approaches make the area of coarse graining and multi scaling very rich and dynamic. In order to develop and provide new modules that can be of use to a wide community, we plan to star from existing knowledge within E-CAM, together with the interests and requests received from the industrial partners involved in E-CAM. Specifically, we have not identified a direct, short.-time strong interest from none of our industrial partners or academic beneficiaries in the area of KMC and MESODYN. On the contrary, both Michelin and Unilever have shown strong interest in DPD, MD and alternative mesoscopic models for polymer composites. E-CAM groups pursue computational soft-matter research on a quite broad basis. Therefore, we take advantage of this versatility to focus on the relevant expertise of WP4 teams that better ft to the specific goals identified by interacting with the industrial partners. In the following section we describe the relevant scientific background of the academic partners. This scientific perspective, although not exhaustive, provides the motivation to identify the relevant software we will develop within E-CAM

We will exploit and take advantage of E-CAM activities, such as the State of the Art workshops, to bring external experts in the complementary methods and strategies mentioned earlier in this section. Exposing E-CAM partners, both academic and industrial, to leading developers of alternative software and model developers will help to identify new generalizations of existing software, or the need to make an effort and expand the existing expertise.

2 Relevant scientific background of WP4 partners

Partner groups in WP4 originate from different directions with different scientific motivation and background. The common strength is a solid background in method and algorithm development and efficient implementation of software, which forms a strong team for a broader class of methods and software components in E-CAM. In the following, partner groups are briefly introduced with their interest and strengths.

The Berlin partner is composed by three different groups linked directly to the CECAM-DE-MMS node: The group of Prof. Luigi Delle Site (former node director), the group of Prof. Carsten Hartmann (current node director) and the group of Prof. Christof Schütte (scientific advisor of CECAM-DE-MMS for the Freie Universität Berlin and the Zuse Institute Berlin). They have expertise in open systems quantum/path integral adaptive resolution simulation technique, Particle Insertion Grand Canonical Adaptive-Resolution-Simulation (PI-GC-AdResS)[14, 15, 16, 17], efficient numerics for the optimal control of (open) molecular systems or Model Order Reduction (MOR) [18],[19, 20],[21, 22], and discretisation of high-dimensional Molecular Dynamics (MD) using Markov State Models (MSM)[19, 20]. The Berlin team will take advantage of this complementary expertise with the objective to turn these complementary pieces of software and numerical skills into functional numerical algorithms. The development of a software platform which ensures a consistent exchange of energy and matter with the environment constitutes a challenge. Such a platform will be useful to a wide community of researchers and computer scientists. It will establish a computational approach to open boundaries. The group in Berlin has the capabillity to remodel the principles of MSM and MOR within the perspective of a Grand Ensemble (so far these approach have been designed only for a constant number of particles). The new computational platform will be assessed in a variety of contexts. We will take advantage of activities that will be developed in WP3 to explore how to adapt the new software platform for mixed dynamics, involving Quantum and Classical systems.

The partner from Jülich contributes to WP4 with a code for particle based hydrodynamics simulations. In fact, hydrodynamic interactions between solutes in solution often play an important role for transport properties, structure formation and collective properties of a variety of complex, highly heterogeneous systems [23]. There exist a variety of approaches explored previously by the scientific community to take into account this type of interaction, which is mediated between the solutes by the solvent and which incorporates many-body contributions. The inclusion of many-body contributions, which prohibit the application of a linear superposition principle, makes it difficult or even impossible to model hydrodynamic interactions by a force-field description (as it is often done for direct atomistic interactions). Therefore, a variety of methods (e.g. LB, DPD, MPC) have been proposed for a solvent model, which include hydrodynamic properties and hydrodynamic interactions between solvated particles as a result of collective properties, resulting from the abidance of conservation laws. The software platform based on MPC developed in Jülich, called MP2C, constitutes a flavour of MPC where the coupling between solvent and solute particles is realised by the hybrid implementation of MPC and MD.

The partner from Mainz has expertise in the quantitative prediction of material properties of macromolecular systems (and beyond), based upon a fundamental understanding of the underlying molecular structures and processes. This interest fits into the needs of the scientific and industrial community. In particular, it has developed simulation approaches corresponding to very different length and time scales, plus a team composed of scientists with complementary expertise that correspond to these scales. Therefore simulations need to tackle questions as different as calculating electron transport in organic semiconductors on the one hand, and nonlinear rheology of polymer melts on the other. Both fundamental understanding and practical industrial applications quite often require the consideration of the same system on different scales. It is therefore highly desirable to have software packages available that offer a very broad host of (complementary) methods, are easily extensible, computationally efficient, well-documented, and user-friendly. These demands contradict each other to a certain extent (in particular efficiency vs. flexibility), and therefore different answers have been found in the community, depending on the weight that is put on these various aspects.

The Barcelona partner has expertise in the development and use of coarse-grained, dynamically consistent computational codes to study soft materials. As a result, it has developed flexible computational tools that can be used to address a wide variety of heterogeneous materials. Specifically, the developments in the foundations of DPD and the proposal of new DPD algorithms, and its contribution to the development of a LB platform for complex fluids is of scientific interest to the community, as well as for the industrial partners involved.

The Dublin partner has expertise in the development of modules that combine different degrees of corse-graining. Understanding of the interactions between living organisms and nanomaterials require addressing how to model bio-nano interfaces as well as a number of the processes, such as hydration, and their sensitivity to thermodynamic conditions (solvent pH, salt, temperature, among others). The Dublin partner has proposed new mesoscale and multiscale models to address these new situations. The teams brings to the consortium a number of force fields, and different proposals for coarse graining. these new computational approaches are flexible and can be combined with existing packages, such as Groningen Machine for Chemical Simulations (GROMACS). This expertise and existing codes

provide a solid basis to develop modules and codes to model at a mesosocpic scale effects such as polarizability, systematic proposals for new potentials of mean force (PMF), various constant stress conditions (going beyond isotopic barostats), and the development of models that operate at constant pH.

The team at STFC provides support to the scientific community by developing and sustaining a number of software packages. Regarding the purpose of WP4, STFC has created, developed and is maintaining DL_MESO[2], a general-purpose software package for modelling complex chemical systems using mesoscale modelling techniques [24]. DL_MESO was conceived by and created for UK Collaborative Computational Project for computer simulation of condensed phases (CCP5) [25] – UK Engineering and Physical Sciences Research Council (EPSRC)'s Collaborative Computational Project for computer simulation of condensed matter – as a companion package to its classical molecular dynamics code DL_POLY[3] [26]. The package consists of two highly-parallelisable simulation codes: DL_MESO_LBE based in the LB method [27] and DL_MESO_DPD based on DPD [28, 11]. These methodologies are capable of modelling complex fluids at length and time scales approaching those of interest for engineering and industry.

[2] A general purpose mesoscale simulation package developed at Daresbury Laboratory
[3] A general purpose MD simulation package developed at Daresbury Laboratory

3 Description of the relevant software to address the scientific challenges

WP4 is characterized by a heterogeneous set of computational platforms and tools that will be developed in parallel by the involved teams. The modular structure of new pieces of software developed within this WorkPackage will ensure they can be useful for the different types of software that need to be developed. The description of the different relevant software, and their structure will be a reference tool as the project unravels and each team develops modules. This effort to provide, as much as possible, standard modules will require the strong involvement of E-CAM software programmers.

In most cases, the software identified is of public access and maintained. However, due to the versatility of coarse-grained and multiscale methods, in some cases the software described is being actively developed as discussed appropriately below.

3.1 PI-GC-AdResS

The aim of the E-CAM project is twofold: (a) to provide an efficient computational platform for PI-GC-AdResS, MSM discretizations and MOR, including easy-to-use interfaces and algorithms that are as flexible and general as possible so as to allow for easy portability and transferability between different MD codes, (b) to rigorously study connections between control theory and MSM-based decision problems and the physics of open systems, principles of which are the backbone of PI-GC-AdResS.

3.1.1 What can the software do? What are its limitations?

So far the open source code in which PI-GC-AdResS has been implemented is GROMACS. Simulations using this code have proven the solid nature of the theoretical principles of the method, however its implementation is not as efficient as it could. E-CAM will search for possible alternatives to GROMACS. The other two approaches (MSM, MOR) do not need a specific code, but, because they use trajectories, they need an optimal interface with the MD code.

3.2 MP2C

MP2C implements the particle based hydrodynamics method MPC, which can be coupled in the same program with MD. The method is based on a stochastic collision model, where collisions between particles are considered within a local environment. Solvent particles are considered to have a cartesian coordinate, momentum and mass, but are considered as point-like particles with no excluded volume. Collisions between solvent particles are therefore taken into account on a more abstract level of description. To make the momentum transport across the system more efficient, multi-body collisions are considered, which change simultaneously the momentum of N_c particles together, while obeying the conservation laws for momentum and energy. Angular momentum conservation can be implemented additionally, which induces additional computational costs and slows the method down by approximatively a factor of 2-3. The total system is subdivided into small cells of length a, where on average $N_c = \rho\, a^3$ particles are located. A simple collision rule that obeys local mass, momentum and energy conservation is imposed [29]. If combined with MD, there is a simple way to take momentum transfer onto solutes into account.

3.2.1 What can the software do? What are its limitations?

MP2C has been used in a number of soft matter simulations, primarily to study the solvent properties or statistical properties of large polymer systems. The efficient parallelization has especially allowed for the study of large systems and the long-range effects in hydrodynamic systems in the small wavenumber regime. Non-equilibrium simulations in Couette and Poiseuille flow have been performed to study, e.g., concentration effects in semi-dilute polymer systems [30] or dynamic effects of polymers [31].

The model compressibility constiutes a deficiency, as it consists essentially of particles, which do not interact via a force field parametrization. Especially there are no repelling forces, which allows to compress the liquid. This is not a problem in situations of low velocities and simulations in thermal equilibrium, but comes into play when large velocity gradients appear. This currently restricts the method essentially to small Mach numbers. Also the inclusion of multi-component solvents or mixture of fluids has not yet been fully solved for the three-dimensional case. The model, however, also allows for more general flow field simulations, e.g. which includes engineering applications

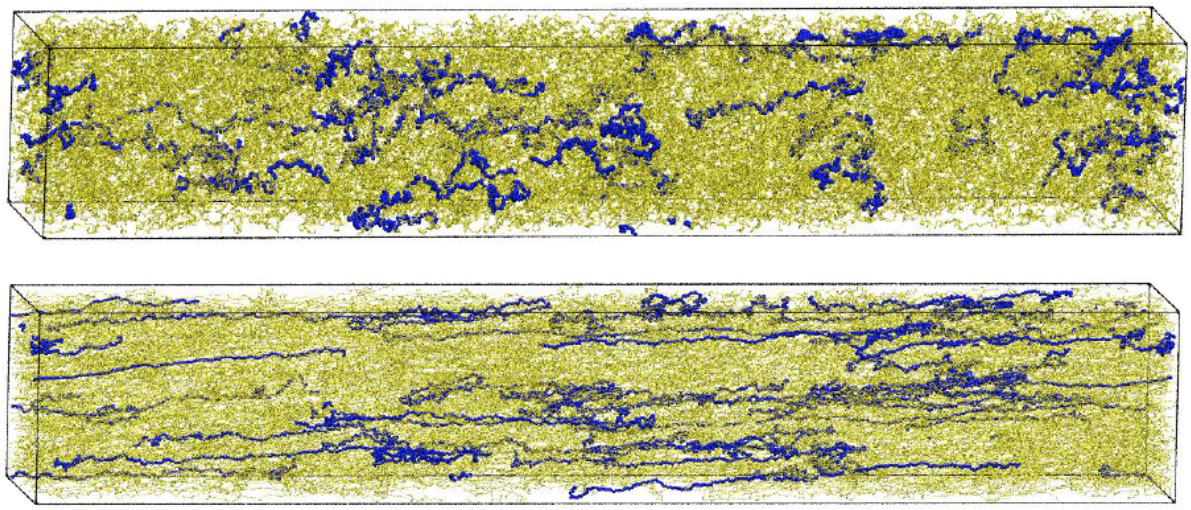

Figure 1: Semidilute polymer systems, simulated with MP2C, in Couette flow at two different shear rates: $\dot{\gamma} = 10^{-4}$ (top) and $\dot{\gamma} = 10^{-1}$ (bottom). Solvent particles not shown. Different colours only highlight separate polymer chains to illustrate the structure change from low to high shear rate conditions.

in the field of gas diffusion layers of fuel cells. A recent development of the model has considered an adaptive refinement strategy, where the nature of an atom has been gradually been transformed from an atomistically resolved particle into an MPC particle [32]. This kind of extension is useful in the study of solvation, if the structure of solvent particles around a solute has to be considered in greater detail. The solvent properties are then relevant on a larger scale, where the coarse-grained description of MPC can be applied. It could be shown, thereby, that the essential hydrodynamic properties of the fluid are conserved across the transition zone between explicit solvent atoms and MPC fluid [32].

3.3 ESPResSo++

Extensible Simulation Package for Research on Soft matter, ESPResSo++ [33, 34] is a software package that has evolved from the still existing older package "ESPResSo" [35, 36], is a general purpose computational package to simulate complex fluids. It has been developed by the Mainz team. ESPResSo++ is essentially a package for doing MD in various “flavors”, including DPD [11], and features also a LB part that can be coupled to a particle system [12], such that soft-matter systems with hydrodynamic interactions can be simulated efficiently. Other packages (partly competing, partly complementary) are the standard MD packages Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [37, 38], GROMACS [39, 40, 41] and DL_POLY [42, 43], as well as `DL_MESO` [44, 24].

ESPResSo++ is designed as a flexible simulation tool that builds upon the data structures of both particles (MD, DPD) and lattices (LB). It is also clearly designed to run on high-end supercomputers. Its scalability is very good (slightly worse but well comparable with LAMMPS, see Ref. [34]). It is reasonably well documented but this aspect could be improved. The Mainz team regularly organizes one CE-CAM tutorial per year to teach the use of the software. Easy exchange of data with other packages (LAMMPS, GROMACS, Versatile Object-oriented Toolkit for Coarse-graining Applications (VOTCA)) is implemented. Due to its implementation as a Python module, full interoperability with other scientific Python packages is automatically implied.

3.3.1 What can the software do? What are its limitations?

ESPresSo++ has implemented a wide variety of functionalities. Among them: Standard MD, Various thermo- and barostats, DPD (still needs testing and optimization), LB (ideal gas), Coupling LB — MD, Electrostatics, Adaptive-Resolution-Simulation (AdResS) and H-AdResS, Parallel tempering, Advanced methods for equilibrating polymer melts, Online coupling to Visual Molecular Dynamics (VMD), a molecular visualization program for displaying, animating, and analyzing large biomolecular systems [45], for visualization, Parallel I/O, and Load-balancing for inhomogeneous systems

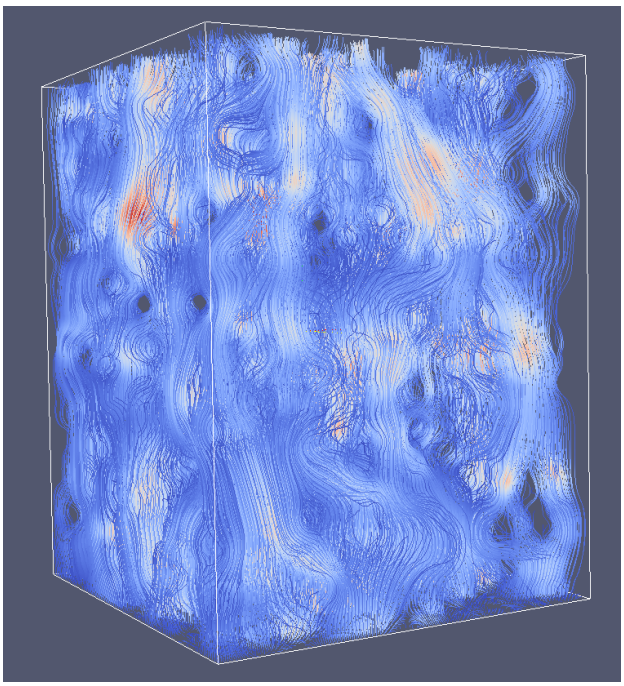

Figure 2: Flowfield of a gas stream in a random geometry, representing the gas diffusion layer in a fuel cell.

3.4 Ludwig

Ludwig is a versatile LB platform, which exists in serial, Message Passing Interface (MPI) and Open-MP versions, and which has been developed in an academic context [46]. *Ludwig* has increased in capabilities since its beginning, in 1998, at the university of Edinburgh, to become competitive at an international level. The initial version was developed to study the dynamics of binary mixtures based on a free energy implementation following the pioneering work of the Oxford group [47]. The code has evolved to account for wetting [48], colloidal suspensions [49, 50], liquid crystals [51], surfactants, active suspensions and active fluids [52] and electrokinetics [53]. This versatility shows the potential of this mesoscopic technique to study a wide variety of soft materials. As a result of these technical advances associated to the new scientific needs, the code description is distributed in a series of publications. In some cases the new methods are described jointly with new scientific results, as the examples provided above show. The Barcelona team has been involved in the development of new modules within *Ludwig*. Specifically, it has generalized colloidal features and have included metabolic processes in a hybrid, continuum model which merges LB with a continuum description of the microorganism populations.

In parallel to this platform, the Barcelona team will exploit the in-house codes developed for DPD with many-body potentials [54] to extend current capabilities. The goal is to develop appropriate effective potentials that account for polarization effects in the presence of heterogeneous systems. Such many-body potentials will be complementary to other initiatives within E-CAM and will contribute to the generation of a new family of effective, soft potentials, for a large variety of materials. The corresponding module will contribute to the E-CAM library.

3.4.1 What can the software do? What are its limitations?

The model is very versatile because of its modular structure. *Ludwig* clearly distinguishes the core associated to model the solvent from the different free energy models implemented to simulate a variety of complex fluids. As a result, *Ludwig* can be easily adapted to new kinds of materials once the appropriate free energy models identified. Ludwig accounts consistently for thermal fluctuations [55, 12] and can be easily adapted to include thermal fluctuations for the different free energy models that have been already implemented [56].

Ludwig is intensive in memory and generates large data files. Data management is a challenge when scaling the code to run on a large number of processors. The requirement to deal with large amounts of data to post process the simulations outcomes also will require careful scrutiny. When dealing with suspended particles, load unbalance may develop. In electrokinetic simulations one needs to solve the Poisson equation in real space (due to the electrostatic heterogeneity of the systems of interest). This may become a bottleneck for the simulations and efficient and highly scalable methods to solve the Poisson equation will be required.

3.5 Coarse-grained models for nanoparticles

The Dublin team has developed a variety of effective potentials to model rigid protein globules and lipids, where each aminoacid or part of a lipid molecule is replaced by a CG bead The model preserves the size and shape of the molecules, location of hydrophobic and hydrophilic sites and charges, and therefore is suitable for studying the specificity of materials in the interaction. UCD partner has also developed a CG model for nanoparticles (NPs), based on effective beads that separate the interior from the nNP surface, which can be adapted to simulate arbitrary materials. The material-BM interaction can account for the interaction of surface beads with the solvent, while the inner beads are responsible for the van der Waals interaction between the bulk material of the NP and the BM beads.

These new forces and nanoparticle models can interface with standard software packages, such as ESPResSo++ (and tcl scripting), GROMACS (including modified versions made either in-house or elsewhere), MagiC, and possibly LAMMPS or DL_POLY. Specifically, the Dublin partner has already developed tcl and python scripts that convert the all-atom PDB crystal structures of proteins into configuration files for ESPResSo package. Additional ESPResSo scripts have been developed to be able to compute a number of relevant thermodynamic properties, such as calculate the mean adsorption energy for proteins on NPs of different size. The existing pieces of software can be made more systematic to reach a broader community.

Since GROMACS and LAMMPS have different strengths, it is important that newly developed force fields and different coarse grained models are structured in a way that can interface with both packages. For example, the treatment of long range forces in GROMACS is not completely clear, the treatment of stress is rather restricted. LAMMPS is much more flexible to use isotropic or anisotropic and triclinic barostats, but it is not great for biological simulations as its version of Chemistry at Harvard Macromolecular Mechanics (CHARMM) is ancient (for example it does not include CMAP corrections).

Regarding constant pH simulations - several approaches are currently used including: *ab-initio*, Path integral MD, Quantum mechanics/Molecular Mechanics (QM-MM), empirical valence bond, hybrid Poisson Boltzmann-MC-MD, λ dynamics (available in specialized versions of GROMACS from Donnin and Grosmuller) , and standard MD with prior PK estimates of appropriate charge distributions over titratable sites. Path integral MD is arguable the most accurate, as standard *ab-initio* does not take into account zero point motion - but far too expensive for large systems. For large systems, only Poisson Boltzmann- MC- MD (using modified versions of GROMACS and Assisted Model Building with Energy Refinement (AMBER)) , λ dynamics, and standard MD with prior PK estimates of appropriate charge distributions over titratable sites have any chance of being practical for both large scale MD and mesoscale simulations.

The Dublin partner has expertise in exploiting the mentioned packages, to perform both regular atomistic MD and mesoscale coarse-grained MD simulations, involving around a quarter million atoms, and in some cases more than a million. The greatest software/hardware challenges are expected to come from multiscale modelling - for example, in order to account properly for the effects of salt, pH and hydration.

3.5.1 What can the software do? What are its limitations?

The existing software combined with ESPResSo is very flexible to compute a wide variety of thermodynamic and mechanical properties for complex molecules, such as adsorption energies and entropies on inorganic surfaces. The methods can be exploited to analyse kinetic processes for systems containing nanoparticles at a variety of interfaces (liquid, solid, membranes).

The main limitations of the model and the software include the fact that the model, and hence the corresponding software, neglect protein unfolding and change of conformation, the change of hydration state upon contact of the CG beads, as well as the counterion release effects upon contact of the CG beads.

3.6 DL_MESO_DPD

WP4 will make use of the DPD code, DL_MESO_DPD, which models particles along with a pairwise thermostat to ensure Galilean invariance and thus correct fluid behaviour. The particles frequently interact with soft pairwise potentials[57] that allow for larger time steps than those available in classical molecular dynamics while still retaining sufficient chemical detail for mesophase-based structures.

DL_MESO_DPD is capable of modelling a wide range of interactions between particles: various pairwise conservative (van der Waals-like) interactions, bonded interactions and electrostatic interactions with smeared charges. For pairwise interactions, the code uses the linked-cell list strategy to find relevant particle pairs efficiently with linear scalability.

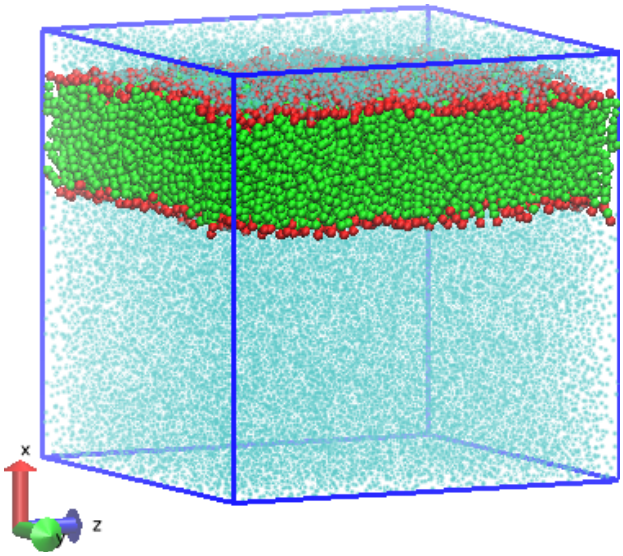

Figure 3: Lipid bilayer formed from self-assembly of amphiphilic molecules in solution, as modelled using DL_MESO_DPD.

The code is modular in structure: similarly functioning subroutines (e.g. pairwise interaction calculations) are grouped together, making it straightforward to find particular functionalities in the code and to add new features. It is well documented: both a user manual and a descriptive article[24] are available.

3.6.1 What can the software do? What are its limitations?

DL_MESO_DPD can model a variety of non-bonded interaction potentials between particles. Specifically, it has implemented a Standard DPD model (Groot-Warren) [57] as well as a version of the many-body (density-dependent) DPD [54, 58, 59] , Lennard-Jones[60] and Weeks-Chandler-Andersen [61] potentials, along with the DPD thermostat or alternative pairwise thermostats: Lowe-Andersen [62], Peters [63, 64] or Stoyanov-Groot[65]. A variety of bonded interactions (stretching, angles and dihedrals) are available for representing longer molecules as joined chains of particles. Charges can also be applied to particles and the associated electrostatics calculated using Ewald sums along with either Slater-type (exponential)[66] or Gaussian [67] charge smearing to prevent collapse of soft ions.

General limitations of DL_MESO_DPD include additional memory requirements and communication for non-DPD pairwise thermostats, less effective particle and parallel scalability for Ewald reciprocal space calculations due to the use of direct Ewald sums, and no current capability to model fixed-length constraints between particles to maximise simulation time step sizes. Considering polarisable models in DPD, while DL_MESO_DPD can model explicit dipole systems by applying oscillating bonded interactions between two oppositely charged particles, smeared charges can only be applied at the centres of particles and no capability currently exists for applying constant dipole lengths or implicit dipoles (including determining rotational movement and orientation).

4 Software structure, HPC perspectives and flexibility in technical development

All the identified codes are written in standard scientific languages and have a structure that allows them to be exploited already on supercomputing facilities. We describe here the main features of the software identified, their capabilities to exploit HPC facilities and their computational potential.

4.1 MP2C

The program is written in modern Fortran (90 to 2008) and highly efficiently parallelized, based on MPI [68]. Due to its local character, no interaction between particles in different collision cells need to be considered and the collisions can be computed completely independent in different collision cells. Therefore, the parallelization utilizes a domain decomposition, where P processors share the work over the whole system. Due to the shifted mesh of collision cells, communication between processors has to be considered to complement information of overlapping cells. Also, if solvent particles leave the domain after the integration step, information has to be transferred from one processor to the neighbor. In this case, individual particle information is completely transferred from one processor to the other and original particle information is erased. This ensures a fully data distributed implementation, which allows for simulations of very large systems.

Due to the strong imbalance between the number of MD and MPC particles, a problem of workload distribution can easily occur. This is especially the case, when small numbers of MD particles are simulated which even might be distributed inhomogeneously in the system, as it is often the case in non-equilibrium and dynamical state state simulations. To account for the imbalance, two procedures can be considered: (i) implementation of a load balance scheme, which is based on a shift of domain boundaries and (ii) a hybrid domain decomposition, which considers MD and MPC particles on their own partitions. The load balance procedure is implemented in two flavours, where one is based on a shift of parallel planes of sub-domains, which ensures a constant number of neighbour processors (i.e. 26 in a three-dimensional setting), whereas the other method also considers staggered domains, which has a better adjustment of equal particle numbers in every domain, but does not guarantee a given number of neighbour domains, which,sometimes, induces a large communication overhead between domains.

The program has been tested on a variety of different architectures, e.g. Intel based cluster architecture or IBM BlueGene/P/Q [69, 70]. In a benchmark simulation, the program could be executed on the full BlueGene/Q in Jülich (JUQUEEN [71]), i.e. on 458752 cores, which made the program enter into the so called *High-Q Club* [72, 73] in a very early stage, which comprises all codes making use of the whole JUQUEEN hardware. To enforce capabilities in input/output of data, the highly efficient library SIONlib[4] has been connected to MP2C, which enables high throughput of large data sets [70]. If electric charges are connected to the solutes, functionality of the code can be extended to electrostatic problems by linking the ScaFaCoS[5] library to the code, which implements a number of methods to treat electrostatic interactions between particles in various types of boundary conditions [74, 75, 76]. The code will also be benchmarked in PRACE.

4.2 ESPResSo++

While LAMMPS and GROMACS are slightly superior in terms of execution speed of a "standard" Molecular Dynamics simulation, ESPResSo++ is built upon a fairly abstract C++ kernel and hence is particularly useful for introducing new concepts, new methods, and new algorithms. An important feature is the interfacing of the kernel with Python, such that many high-level tasks can (and should) be done via Python scripts, without ever touching the basic simulation code. In particular, ESPResSo and ESPResSo++ have proven quite useful for the implementation of long-range interactions [77, 78] (mainly ESPResSo), and, more recently, of "AdResS" (only available in ESPResSo++), which is also further developed in this WorkPackage. The latter is an adaptive resolution scheme that simulates molecular systems at different levels of coarse-graining in different parts of the simulation box [79, 80, 81].

ESPResSo++ is designed as a flexible simulation tool that builds upon the data structures of both particles (MD, DPD) and lattices (LB). It is also clearly designed to run on high-end supercomputers. Its scalability is very good (slightly worse but well comparable with LAMMPS, see Ref. [34]). It is reasonably well documented but this aspect could be improved. The Mainz team regularly organizes one CECAM tutorial per year to teach the use of the software. Easy exchange of data with other packages (LAMMPS, GROMACS, VOTCA) is implemented. Due to its implementation as a Python module, full interoperability with other scientific Python packages is automatically implied.

[4] Scalable I/O library for parallel access to task-local files
[5] Library for Fast Coulomb Solver

4.3 Ludwig

Ludwig has been run in parallel environments using the supercomputing facilities in Europe. *Ludwig* has been benchmarked in a Cray T3D, Cray-T3E-1200, Hitachi SR-2201, Blue Gene as well as on a shared memory platform SUN HPC-3500 at EPCC (Edinburgh) and CSAR (UK), as well as in Mare Nostrum (Spain). Since 2006 it is exploited routinely by the Barcelona team in Mare Nostrum where it had showed an excellent performance, using up to 4096 processors [50]. It has been proven to show excellent scaling properties in all MPI-architectures from 16 to 4096 processors. I/O is the main bottleneck in scalability (a 512^3 system will generate 31Gb per configuration dump), and it has been optimized by performing parallel I/O. Nonetheless, the frequency of configuration dumps must be tuned appropriately on large system sizes.

Ludwig is intensive in memory and generated large data files. Data management is a challenge when scaling the code to run on a large number of processors. The requirement to deal with large amounts of data to post process the simulations outcomes also will require careful scrutiny. When dealing with suspended particles, load unbalance may develop. In electrokinetic simulations one needs to solve the Poisson equation in real space (due to the electrostatic heterogeneity of the systems of interest). This may become a bottleneck for the simulations and efficient and highly scalable methods to solve the Poisson equation will be required.

4.4 DL_MESO_DPD

DL_MESO_DPD is designed to be run on high-performance computing environments and uses domain decomposition as its main parallelism strategy by dividing system volume is divided equally among processor cores: boundary halos are constructed using MPI communication with neighbouring cores to obtain particle data. The sweep through linked cells for particle calculations can optionally be accelerated using OpenMP loop multithreading.

Depending on problem size and the associated computation/communication ratio, the code scales well up to thousands of processor cores. The code has been run on a variety of HPC environments, including supercomputers with high-powered processors such as Cray, lower-powered processors (e.g. IBM BlueGene/Q) and Intel Xeon Phi coprocessors/accelerators.

5 Industrial connections

The software identified has also taken into account the existing and foreseen contacts with industries. These contacts are relevant since they will help to identify the relevant modules that should be developed within EP4. The challenges posed by these interactions will also provide specific cases to motivate the needs to exploit HPC and address challenges associated to exploiting intensive parallelization. E-CAM postdoctoral researchers will play an instrumental role in developing these connections and identifying the software and hardware challenges.

The company *MODAL AG*, a small Berlin-based enterprise that gathers different scientific competences in the Berlin-Potsdam area (from optimisation of railway networks to biomedical applications), has expressed a genuine interest in our project. Attached you will find a letter of interest by the MODAL AG. We expect that in the course of the project and through the E-CAM network, other companies may get involved in the project which will promote the results obtained in the project and eventually foster collaborations, with the aim of applying our methods and codes to specific system or phenomena of interest (for example IBM with Tavernelli from WP3, see next section).

The planned industrial collaboration of the Mainz team with Michelin builds, to a significant extent, on software within ESPResSo++ and VOTCA that already exists. The industrial partners are interested in block copolymer systems. One major part of the project is the generalization of a recently established method for the equilibration of polymer melts to block copolymers. This new method [82] applies a hierarchy of coarse-graining and fine-graining steps in (weak) analogy to the multigrid strategy of lattice simulations. It is expected that desired material properties can be obtained by suitably tuning the chemistry and the molecular architecture, in particular in a system where one block (A) is glassy and the other (B) is not. To find the mechanical properties of such a system, it is necessary to resolve the structure and dynamics near an AB interface. In a later stage of the project, AdResS shall be applied to tackle this problem. It is expected that this will require both the development of new C++ routines in ESPResSo++, and also a host of new Python scripts to drive the simulations.

Since 2010, STFC has been involved in an industrial collaboration with Unilever, Syngenta and Infineum to parameterise and use DPD as a tool for computer-aided formulation of new products. This collaboration was subsequently strengthened by the creation of the Hartree Centre in partnership with IBM in 2012, the creation of strategic research and development partnerships between Unilever and STFC/IBM in 2013 and the award of a grant from EPSRC's Technology Strategy Board to support the collaboration.

The major software component for the collaboration is `DL_MESO_DPD` as the DPD simulation engine, with several additional utilities written by STFC and IBM to set up simulations and process results to predict the mesoscale-based behaviour (e.g. phase behaviour, critical micelle concentration) of new trial chemicals. The process of determining the requirements for the DPD model used for this project have led to additional features in `DL_MESO_DPD` to expand the range of chemical species and behaviour that can be realistically modeled. The development of polarisable water models for DPD is of great interest to all three industrial partners, since the associated changes to dielectric properties and short-range interactions akin to hydrogen bonding are likely to prove useful for relevant chemical species. The work in WP4 will include direct involvement with Unilever to develop the theoretical basis of DPD polarisable water models, as well as their practical implementation for systems of particular interest to Unilever.

MP2C has been applied and further developed in a number of EU [83] and national projects [84]. There are no industrial connections at present.

The main applications of the techniques that will be developed by the Dublin team lead naturally to a number of potential interaction with different industries. Among the companies already contacted, we can mention partners in a) Food industry: where the coarse-graining software can be used, for example, for creating models for milk (proteins + fat + sugars) and other complex fluids, in particular to study the fowling process, milk concentration, and protein adsorption on surfaces. Existing contacts include the Kerry group. b) Cosmetics and medicine, where the models can be used to predict the NP protein corona and NP uptake rates in biological fluids and tissues (L'Oreal). c) Semiconductor industry where the models of bionano interfaces can be extrapolated for coarse-graining polymers and modeling block-copolymer directed self-assembly, which is currently used in nanolithography for making sub-10 nm masks (Intel).

6 Future needs and challenges

As outlined before the software components in terms of simulation codes, analysis tools and functions comprise a set of heterogenous elements, which ideally have to be unified under a common driver software to be accessed most conveniently. However, the complexity of the codes makes it currently unrealistic to rewrite the codes under a unified programming model. This is on the one hand a matter of time and human resources and on the other hand a decision problem, which programming model would be selected one for the exascale compute facilities. As long as this is not commonly agreed between hardware vendors, software developers and the scientific community, a decision for rewrite of complex software would be on a speculative level. It is understood that there is a discrepancy between (large part of) industrial needs and current computational capabilities in terms of hardware evolution. Therefore, the project aims at providing software modules in combination with complex simulation codes which can be coupled together and offered to the industrial and scientific community. It is the flexibility of software components and modules which is needed to follow next generation hardware developments most closely and to be able to port and optimize software to varying landscape of hardware architectures. It is believed, however, that simulation codes, written in languages, e.g. C,C++, Fortran90 and higher, will be able to run on future architectures. Nevertheless, it might be expected that performance will degrade with respect to the peak performance, since the complexity in terms of compute units, memory hierarchy and network capabilities have to be considered explicitly for optimal performance of codes on next generation hardware.

Future developments should take into account this flexibility, which might be implemented on different levels: (i) use of languages, styles and communication protocols, which can be addressed via wrappers or (ii) hardware independent software languages. Especially the latter approach is appealing since with a given implementation a set of different hardware platforms can be addressed. Open Computing Language (OpenCL) is one one example for this language approach which can be combined with communication libraries, e.g. MPI. At present it might, however, be too early to focus on one of theses approaches, since the development and implementation of specific drivers is always dependent on the hardware vendors and as long as a commitment for future development and adaption to new platforms is not guaranteed, the development of new software components might consider alternative ways, e.g. wrappers and interfaces. There are several approaches to adapt to new hardware. One example is the interface library StarSs from Barcelona, which provides a number of different components for addressing shared memory devices or Graphics Processing Unit (GPU)s. Specifically for GPUs, the directive based approach Open Accelerators (OpenACC) has been proposed, which offers a simple way to couple existing codes with the power of GPU devices. For non-standard and more complex algorithms, the optimization seems to be involved and a lot of details have to be passed to the compiler level. As mentioned before, OpenCL, seems to be a most versatile language to cope with a larger number of hardware components. The use of such languages allows an immediate comparison of an algorithm on various platforms and provides a high flexibility for exploitation of next generation hardware. However, as mentioned before, the hierarchy of memory has to be taken into account which is sometimes crucial for the performance (e.g. large difference in size and latency between global and shared memory on GPUs).

Concerning the software landscape, which is relevant for E-CAM, there are a number of different challenges and needs. In the following we describe short-term challenges that will be addressed within each type of software.

6.1 PI-GC-AdResS

It is necessary to document properly the software implementation of Adaptive resolution codes. We will deploy a manual that describes the use of PI-GC-AdResS and how it can be interfaced with the most popular open source MD codes (GROMACS [39], LAMMPS [37], CP2K[6] [85], NAMD [86], OpenMM[7] [87]). We will also produce a user manual to use optimized control algorithms based on reduced-order MSM and how to plug it into the PI-GC-AdResS. We need to clarify how to make PI-GC-AdResS more useful and transferable reducing the number of degrees of freedom (without altering the accuracy). Such an advance will allow the study of systems which otherwise would not be feasible with standard tools, thus it will allow groups which are not armed with exotically large computational facilities to study systems otherwise not accessible to them.
Software activities on WP3 focused on open quantum systems will allow for the extension of PI-GC-AdResS to model mixed classical/quantum systems. Synchronization of the two WPs in this area is crucial for an efficient and productive collaboration leading to new modules and pieces of software of high potential use in the scientific community.

[6] CP2K is a quantum chemistry and solid state physics software package that can perform atomistic simulations of solid state, liquid, molecular, periodic, material, crystal, and biological systems.

[7] High performance toolkit for molecular simulation

6.2 MP2C

MP2C has been efficiently parallelized for Intel based cluster architectures. We plan to extend the code capabilities to include accelerators, e.g. GPU capabilities or Intel Xeon Phi. To better utilize cache and to improve performance, a block structure of solvent particles, combined with a sorting and block based load balancing strategy needs to be designed. In order to enhance the versatility of the code and the potential to benefit from different hardware architectures and HPC facilities, we plan to extend to a hybrid programming model, including MPI based parallelism between compute nodes and Thread based parallelism on a node, profiting from shared memory programming models, could be advantageous. Besides the parallelization approaches, functionality towards more complex geometries (e.g. porous media) or more complex solvent models, allowing for fluid mixtures and low compressibility will be addressed.

6.3 ESPResSo++

ESPResSo++ shows good scalability on HPC facilities. Even if the code will benefit from being adapted to run on a wider variety of architectures, the initial plans will focus on implementing new parts into ESPResSo++, including Smoothed DPD, Multiphase LB, together with an improved documentation and GPU support.

E-CAM will help to make ESPresSo++ more known and more popular in the community. Such a goal will benefit from the close connection with industrial partners along the development of the E-CAM project. Also the interaction with the E-CAM programmers will be helpful to clarify and decide how it will be better to extend the ability of ESPresSo++ in different computational architectures.

The Mainz team has developed yet another package called VOTCA [88, 89] that provides a host of routines to automatically map molecular models from a fine-grained to a more coarse-grained description. VOTCA obviously provides an additional solid basis for further developments in the field. The attempt to systematize these coarse-graining procedures as much as possible is in our opionion *the* major hallmark of modern mesoscale simulation studies.

6.4 Ludwig

Ludwig will be developed to implement thermal fluctuations for a generic free energy. Regarding electrokinetics efficient Poisson solvers that show good scalability must be identified. Current implementations for electrolytes will have to be generalized to account for polarizability. There exist issues of scalability of load balance when considering porous media or very dense colloidal suspensions. The former scenario will be relevant for some of the planned potential applications of the code and modules that will be developed within E-CAM. The problems of load balance are shared with MP2C; a close collaboration between these complementary packages will help to push forward new schemes to enhance load balance under general, non-equilibrium situations, which can be relevant for other computational models.

A contribution to the development of new classes of effective, soft potentials, which can be exploited within DPD is desirable within E-CAM to complement advances in new models and modules associated to charge fluids. Understanding the basis of many-body potentials in the case of charge mixtures will open new possibilities to model complex materials with variable composition and heterogeneous polarization properties. A close collaboration with other E-CAM members, such as STFC, and industrial partners, Unilever, is envisaged in the development and use of these new features.

6.5 Coarse-grained models for nanoparticles

Further development of the software for modelling bionano interfaces within E-CAM can include the development of a code for automated calculation of CG bead-bead interaction (Hamaker) constants according to Lifshitz theory. This is currently not possible, and the simulation method can only be applied to materials with known (measured) Hamaker constants.

Currently the methodology presented here can only be used for very small NPs. A quantitative simulation of coated small or large NPs in contact with a lipid bilayer is not feasible at the moment for several reasons. Most importantly, none of the currently available MD codes is able to combine an NPT simulation of a lipid bilayer with explicit charges with long-range vdW interactions. In the ESPResSo package, the NPT and bilayer compressibilities are not treated properly, while GROMACS does not allow inclusion of long-range forces. Although GROMACS in principle does allow PME treatment of VDW interactions, co-workers have reported difficulties.

6.6 DL_MESO_DPD

For the work in WP4, additional functionality can be added to DL_MESO_DPD to improve the scalability of electrostatic interaction calculations – either by alternative methods of calculating Ewald reciprocal space terms[90] or by using particle-particle particle-mesh methods[91] – and to include a constraint solver (e.g. [92]) to allow for fixed explicit dipoles. An intended output for the work package is to develop algorithms to apply charge dipoles and other multipoles to particles with soft repulsive potentials and implement these in DL_MESO_DPD. The challenge of implementing an efficient solver for electrostatic interactions that shows good scalability on large parallelization is shared by other computational approaches. We will exploit E-CAM environment and explore transversal approaches for these different mesoscopic computational approaches.

References

Acronyms Used

AdResS Adaptive-Resolution-Simulation

AMBER Assisted Model Building with Energy Refinement

CCP5 UK Collaborative Computational Project for computer simulation of condensed phases

CECAM Centre Européen de Calcul Atomique et Moléculaire

CHARMM Chemistry at Harvard Macromolecular Mechanics

CPU Central Processing Unit

CULGI Chemistry Unified Language Interface

DFT Dynamic Functional Theory

DPD Dissipative Particle Dynamics

EPSRC UK Engineering and Physical Sciences Research Council

GPU Graphics Processing Unit

GROMACS Groningen Machine for Chemical Simulations

HPC High Performance Computing

KMC Kinetic Monte Carlo

LAMMPS Large-scale Atomic/Molecular Massively Parallel Simulator

LB Lattice Boltzmann

MC Monte Carlo simulation method

MD Molecular Dynamics

MESODYN Dynamic mean-field density functional theory

MOR Model Order Reduction

MPC Multiparticle Collision Dynamics

MPI Message Passing Interface

MSM Markov State Models

OpenACC Open Accelerators

OpenCL Open Computing Language

PI-GC-AdResS Particle Insertion Grand Canonical Adaptive-Resolution-Simulation

PRACE Partnership for Advanced Computing in Europe

QM-MM Quantum mechanics/Molecular Mechanics

SCFT Self-Consistent Field Theory

VMD Visual Molecular Dynamics

VOTCA Versatile Object-oriented Toolkit for Coarse-graining Applications

URLs referenced

Page ii

https://www.e-cam2020.eu ... https://www.e-cam2020.eu
https://www.e-cam2020.eu/deliverables ... https://www.e-cam2020.eu/deliverables
Internal Project Management Link ... https://redmine.e-cam2020.eu/issues/157
ipagonabarraga@ub.edu ... mailto:ipagonabarraga@ub.edu
http://creativecommons.org/licenses/by/4.0 ... http://creativecommons.org/licenses/by/4.0

Citations

[1] J.G.E.M. Fraaije, B.A.C. van Vlimmeren, N.M. Maurits, M. Postma, O.A. Evers, C. Hoffman, P. Altevogt, and G. Goldbeck-Wood. The dynamic mean-field density functional method and its application to the mesoscopic dynamics of quenched block copolymer melt. *J. Chem. Phys.*, 106:4260, 1997.

[2] M. Müller and J.J. de Pablo. Computational approaches for the dynamics of structure formation in self-assembling polymeric materials. *Ann. Rev. of Mat. Res.*, 43:1, 2013.

[3] A.P.J. Jansen. *An Introduction to Kinetic Monte Carlo Simulations of Surface Reactions.* Lecture Notes in Physics 86, Springer Verlag, 2012.

[4] K. Reuter, D. Frenkel, and M. Scheffler. *Phys. Rev. Lett.*, 93:116105, 2004.

[5] K. Reuter and M. Scheffler. *Phys. Rev. B*, 68:045407, 2003.

[6] S.A. Baeurle, T. Usami, and A.A. Gusev. *Polymer*, 47:8604, 2006.

[7] J.G.E.M. Fraaije, B.A.C. Van Vlimmeren, N.M. Maurits, M. Postma, O.A. Evers, C. Hoffman, P. Altevogt, and G. Goldbeck-Wood. The dynamic mean-field density functional method and its application to the mesoscopic dynamics of quenched block copolymer melts. *J. Chem. Phys.*, 106:4260, 1997.

[8] A. Sevink, A.V. Zvelindovsky, B.A.C. Van Vlimmeren, N.M. Maurits, and J.G.E.M. Fraaije. Dynamics of surface directed mesophase formation in block copolymer melts. *J. Chem. Phys.*, 110:2250, 1999.

[9] A. Knoll, K.S. Lyakhova, A. Horvat, G. Krausch, G.J.A. Sevink, and A.V. Zvelindovsky. Direct imaging and mesoscale modelling of phase transitions in a nanostructured fluid. *Nature Mat.*, 3:886, 2004.

[10] http://culgi.com.

[11] P. Espanol and P. Warren. Statistical Mechanics of Dissipative Particle Dynamics. *EPL (Europhysics Letters)*, 30(4):191, May 1995.

[12] Burkhard Dünweg and Anthony J. C. Ladd. Lattice Boltzmann Simulations of Soft Matter Systems. volume 221 of *Advances in Polymer Science*, page 89. Springer Berlin Heidelberg, 2009.

[13] G. Gompper, T. Ihle, D.M. Kroll, and R.G. Winkler. Multi-particle collision dynamics: A particle-based mesoscale simulation approach to the hydrodynamics of complex fluids. *Reviews Modern Physics.*, 22:1, 2009.

[14] A.B. Poma and L. Delle Site. Classical to path-integral adaptive resolution in molecular simulation: Towards a smooth quantum-classical coupling. *Physical Review Letters*, 104:250201, 2010.

[15] H. Wang, C. Hartmann, C. Schütte, and L. Delle Site. *Physical Review X*, 3:011018, 2013.

[16] A. Agarwal, J. Zhu, C. Hartmann, H. Wang, and L. Delle Site. *New Journal Physics*, 17:083042, 2015.

[17] A. Agarwal and L. Delle Site. *Journal Chemical Physics*, 143:094102, 2015.

[18] H. Wang, C. Hartmann, and C. Schütte. *Molecular Physics*, 111:3555, 2013.

[19] C. Schütte, S. Winkelmann, and C. Hartmann. *Math. Program. (Series B)*, 134:259, 2012.

[20] R. Banish and C. Hartmann. *Math. Control Relat. F.*, XX:in press, 2016.

[21] B.. Schäfer-Bung, C. Hartmann, B. Schmidt, and C. Schütte. *Journal Chemical Physics*, 135:014112, 2011.

[22] C. Hartmann, B.. Schäfer-Bung, and A. Zueva. *SIAM Journal Control Optimization*, 51:2356, 2013.

[23] M.C. Marchetti, J.F. Joanny, S. Ramaswamy, T. Liverpool, J. Prost, R. Madan, and R.A. Simha. Hydrodynamics of soft active matter. *Reviews Modern Physics.*, 85:1143, 2013.

[24] Michael A. Seaton, Richard L. Anderson, Sebastian Metz, and William Smith. DL_meso: highly scalable mesoscale simulations. *Molecular Simulation*, 39(10):796–821, September 2013.

[25] http://www.ccp5.ac.uk/.

[26] I.T. Todorov, W. Smith, K. Trachenko, and M.T. Dove. Dl_poly_3: new dimensions in molecular dynamics simulations via massive parallelism. *Journal Methamitzal Chemistry.*, 16(20):1911, 2006.

[27] S. Succi. *The lattice Boltzmann equation for fluid dynamics and beyond.* Clarendon Press., Oxford, 2001.

[28] P. J. Hoogerbrugge and J. M. V. A. Koelman. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhysics Letters.*, 19(3):155, 1992.

[29] T. Ihle and D.M. Kroll. Stochastic rotation dynamics. a galilean invariant mesoscopic model for fluid flow. *Physical Review E.*, 63:020201, 2001.

[30] C.C. Huang, R.G. Winkler, G. Sutmann, and G. Gompper. Semidilute polymer solutions at equilibrium and under shear flow. *Macromolecules.*, 43:10107, 2010.

[31] C.C. Huang, R.G. Winkler, G. Sutmann, and G. Gompper. Tumbling of polymers in semidilute solution under shear flow. *EPL.*, 93:54004, 2011.

[32] U. Alekseeva, R.G. Winkler, and G. Sutmann. Hydrodynamics in adaptive resolution particle simulations: Multiparticle collision dynamics. *Journal Computational Physics.*, 314:14, 2016.

[33] http://www.espresso-pp.de/.

[34] Jonathan D. Halverson, Thomas Brandes, Olaf Lenz, Axel Arnold, Stas Bevc, Vitaliy Starchenko, Kurt Kremer, Torsten Stuehn, and Dirk Reith. ESPResSo++: A modern multiscale simulation package for soft matter systems. *Computer Physics Communications*, 184(4):1129–1149, April 2013.

[35] http://espressomd.org/.

[36] H.J. Limbach, A. Arnold, B.A. Mann, and C. Holm. Espreesso, an extensible simulation package for research on soft matter systems. *Computer Physics Communications*, 174(9):704–727, May 2006.

[37] http://lammps.sandia.gov/.

[38] Steve Plimpton. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics*, 117(1):1–19, March 1995.

[39] http://www.gromacs.org/.

[40] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1–3):43–56, September 1995.

[41] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(3):435, March 2008.

[42] http://www.ccp5.ac.uk/DL_POLY_CLASSIC/.

[43] W. Smith and I. T. Todorov. A short description of DL_poly. *Molecular Simulation*, 32(12-13):935–943, October 2006.

[44] http://www.cse.clrc.ac.uk/ccg/software/DL_MESO/.

[45] http://www.ks.uiuc.edu/Research/vmd/.

[46] J.C. Desplat, I. Pagonabarraga, and P. Bladon. Ludwig: A parallel lattice-boltzmann code for complex fluids. *Computer Physics Communications.*, 134:273, 2001.

[47] M.R. Swift, W.R. Osborn, and J.M. Yeomans. Lattice boltzmann simulation of nonideal fluids. *Physical Review Letters.*, 75:830, 1995.

[48] R. Ledesma-Aguilar, R. Nistal, A. Hernandez-Machado, and I. Pagonabarraga. Controlled drop emission by wetting properties in driven liquid filaments. *Nature Materias.*, 10:367, 2011.

[49] K. Stratford, R. Adhikari, I. Pagonabarraga, J.C. Desplat, and M. Cates. Colloidal jamming at interfaces: a route to fluid-bicontinuous gels. *Science.*, 309:2174, 2005.

[50] K. Stratford and I. Pagonabarraga. Parallel simulation of particle suspensions with the lattice boltzmann method. *Computers and Mathematics with Applications.*, 55:1585, 2008.

[51] K. Stratford, O. Henrich, J.S. Lintuvuori, M. Cates, and D. Marenduzzo. Self-assembly of colloid-cholesteric composites provides a possible route to switchable optical materials. *Nature Communications.*, 5:3954, 2014.

[52] E. Tjhung, D. Marenduzzo, and M. Cates. Spontaneous symmetry breaking in active droplets provides a generic route to motility. *Proceedings National Academy Sciences USA.*, 109:12381, 2012.

[53] B. Rotenberg, I. Pagonabarraga, and D. Frenkel. Coarse-grained simulations of charge, current and flow in heterogeneous media. *Proceedings National Academy Sciences USA.*, 144:223, 2010.

[54] I. Pagonabarraga and D. Frenkel. Dissipative particle dynamics for interacting systems. *Journal Chemical Physics.*, 115:5015, 2001.

[55] R. Adhikari, K. Stratford, M. Cates, and A.J. Wagner. Fluctuating lattice boltzmann. *EPL.*, 71:473, 2005.

[56] S.P. Thampi, I. Pagonabarraga, and R. Adhikari. Lattice boltzmann - langevin simulations of binary mixtures. *EPL.*, 84:046709, 2011.

[57] R. Groot and P.B. Warren. Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation. *Journal Chemical Physics.*, 107(3):4423, 1997.

[58] S. Y. Trofimov, E. L. F. Nies, and M. A. J. Michels. Thermodynamic consistency in dissipative particle dynamics simulations of strongly nonideal liquids and liquid mixtures. *Journal Chemical Physics.*, 117:9383, 2002.

[59] S. Merabia and I. Pagonabarraga. Density dependent potentials: structure and thermodynamics. *Journal Chemical Physics.*, 127:054903, 2007.

[60] J.E: Jones. On the determination of molecular fields. ii. from the equation of state of a gas. *Proceedings Royal Academy London A.*, 106(738):463, 1924.

[61] J. D. Weeks, D. Chandler, and H. C. Andersen. Role of repulsive forces in determining the equilibrium structure of simple liquids. *Journal Chemical Pysics.*, 54:5237, 1971.

[62] C.P. Lowe. An alternative approach to dissipative particle dynamics. *Europhysics Letters.*, 127:145, 1999.

[63] I. Pagonabarraga, M.H.J. Hagen, and D. Frenkel. Self-consistent dissipative particle dynamics algorithm. *Europhysics Letters.*, 42:377, 1998.

[64] E.A.J.F. Peters. Elimination of time step effects in dpd. *Europhysics Letters.*, 66:311, 2004.

[65] S.D. Stoyanov and R.D. Groot. From molecular dynamics to hydrodynamics: A novel galilean invariant thermostat. *Journal Chemical Physics.*, 122:114112, 2005.

[66] M. González-Melchor, E. Mayoral, M.E. Velázquez, and J. Alejandre. From molecular dynamics to hydrodynamics: A novel galilean invariant thermostat. *Journal Chemical Physics.*, 125:224107, 2006.

[67] P. B. Warren, A. Vlasov, L. Anton, and A. J. Masters. Screening properties of gaussian electrolyte models, with application to dissipative particle dynamics. *Journal Chemical Physics.*, 138:204907, 2013.

[68] http://www.mcs.anl.gov/research/projects/mpi/.

[69] G. Sutmann and W. Frings. Extending scalability of mp^2c to more than 250k compute cores. In W. Frings and B. Mohr, editors, *Jülich Blue Gene/P Extreme Scaling Workshop 2009.*, pages B–2010–02. IOS Press, 2010.

[70] J. Freche, W. Frings, and G. Sutmann. High throughput parallel-i/o using sionlib for mesoscopic particle dynamics simulations on massively parallel computers. In B. Chapman, F. Desprez, G.R. Joubert, A. Lichnewsky, F. Peters, and T. Priol, editors, *Parallel Computing: From Multicores and GPU's to Petascale. Advances in Parallel Computing.*, page 423. IOS Press, 2010.

[71] http://www.fz-juelich.de/ias/jsc/EN/Expertise/Supercomputers/JUQUEEN/JUQUEEN_node.html.

[72] http://www.fz-juelich.de/ias/jsc/EN/Expertise/High-Q-Club/_node.html.

[73] http://www.fz-juelich.de/ias/jsc/EN/Expertise/High-Q-Club/MP2C/_node.html.

[74] G. Sutmann, Th. Lippert, , and P. Gibbon. *Lecture Notes on Fast Methods on Long Range Interactions in Complex Systems. Volume 6.* IAS., Jülich, 2011.

[75] A. Arnold, F. Fahrenberger, C. Holm, O. Lenz, M. Bolten, H. Dachsel, R. Halver, I. Kabadshow, F. Gähler, F. Heber, J. Iseringhausen, M. Hofmann, M. Pippig, D. Potts, and G. Sutmann. Comparison of scalable fast methods for long-range interactions. *Physical Review E.*, 88:063308, 2013.

[76] https://github.com/scafacos.

[77] Markus Deserno and Christian Holm. How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines. *The Journal of Chemical Physics*, 109(18):7678–7693, November 1998.

[78] Markus Deserno and Christian Holm. How to mesh up Ewald sums. II. An accurate error estimate for the particle–particle–particle-mesh algorithm. *The Journal of Chemical Physics*, 109(18):7694–7701, November 1998.

[79] Matej Praprotnik, Luigi Delle Site, and Kurt Kremer. Multiscale Simulation of Soft Matter: From Scale Bridging to Adaptive Resolution. *Annual Review of Physical Chemistry*, 59(1):545–571, 2008.

[80] S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, and K. Kremer. Adaptive resolution molecular dynamics simulation through coupling to an internal particle reservoir. *Physical Review Letters*, 108(17):170602, April 2012.

[81] Raffaello Potestio, Sebastian Fritsch, Pep Espanol, Rafael Delgado-Buscalioni, Kurt Kremer, Ralf Everaers, and Davide Donadio. Hamiltonian Adaptive Resolution Simulation for Molecular Liquids. *Physical Review Letters*, 110(10):108301, March 2013.

[82] Guojie Zhang, Livia A. Moreira, Torsten Stühn, Kostas Ch. Daoulas, and Kurt Kremer. Equilibration of High Molecular Weight Polymer Melts: A Hierarchical Strategy. *ACS Macro Letters*, 3(2):198–203, February 2014.

[83] https://www.montblanc-project.eu/.

[84] http://www.gaspi.de/gaspi/.

[85] http://cp2k.berlios.de/.

[86] http://www.ks.uiuc.edu/Research/namd/.

[87] http://openmm.org/.

[88] http://www.votca.org/.

[89] Victor Rühle, Christoph Junghans, Alexander Lukyanov, Kurt Kremer, and Denis Andrienko. Versatile Object-Oriented Toolkit for Coarse-Graining Applications. *Journal of Chemical Theory and Computation*, 5(12):3211–3223, December 2009.

[90] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh ewald method. *Journal Chemical Physics.*, 103:8577, 1995.

[91] R.D. Groot. Electrostatic interactions in dissipative particle dynamics—simulation of polyelectrolytes and anionic surfactants. *Journal Chemical Physics.*, 118:11265, 2003.

[92] H.C. Andersen. Rattle: a “velocity” version of the shake algorithm for molecular dynamics calculations. *Journal Computational Physics.*, 52:24, 1983.