Science & Technology Facilities Council

DL_Software Modelling Capability

C W Yong, A. Brukhno
Computational Chemistry Group,
Scientific Computing Department,
STFC,
Daresbury Laboratory, Sci-Tech Daresbury,
Warrington WA4 4AD

CCP5

EPSRC

Computer simulation of condensed phase materials
(liquids, polymers, solids, surfaces, macromolecules)

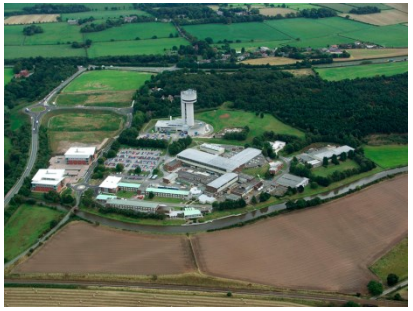
Science & Technology Facilities Council

DL_Software

A collective term for a number of scientific software developed at the Computational Chemistry Group at the Scientific Computing Department, Daresbury Laboratory, spanning across multiple time and length scales.

Computational Chemistry Group

- Develops and sustains applications and associated expertise to facilitate UK's Computational Science over chemistry-dominated time/length scales
- Provides stewardship for UK academia (CoSeC), for stakeholders.
- Facilitates Knowledge Transfer between academia and industry (via Hartree Centre)
- Public outreach and awareness in STEM



Science & Technology Facilities Council

A selection of DL_Software

Many of the software projects are funded by EPSRC via CCP5 and HEC consortia (MCC, UKCOMES).

DL_POLY - The molecular dynamics (MD) program suite.

DL_FIELD - Force-field model setup tool for DL_POLYs.

DL_ANALYSER - General analysis tool for DL_POLY.

DL_CGMAP - Coarse-graining utility for DL_POLY

DL_MONTE - The monte-carlo simulation program.

DL_MESO - Mesoscopic simulation including lattice Boltzmann and dissipative particle dynamics.

CHEMSHELL - Program interface that handles data and communications of various computational chemistry packages.

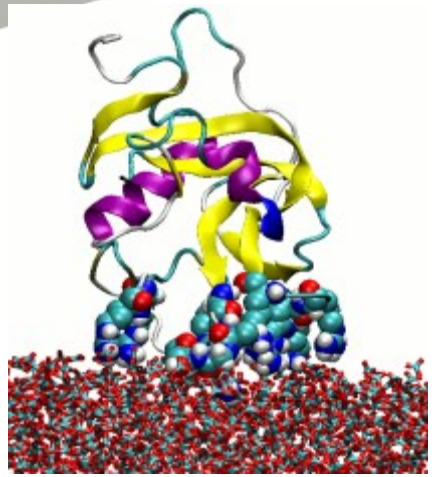
Details, registrations: https://www.ccp5.ac.uk/software

DL_POLY

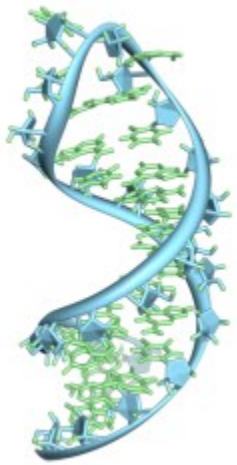
MD is the solution of the classical equations of motion for atoms and molecules to obtain the time evolution of a system.

- General purpose molecular dynamics (MD) simulations. Domain decomposition MD (DL_POLY_4). Up to ~2 billion particles.
- Written in modularised Fortran90 (Nagware & FORCHECK compliant) with MPI2 (MPI1+MPI-I/O) & fully self-contained.
- On average 1500 licences pa world-wide.
- DL_POLY_C. Open source BSD license. Replicated data. Up to ~30000 atoms

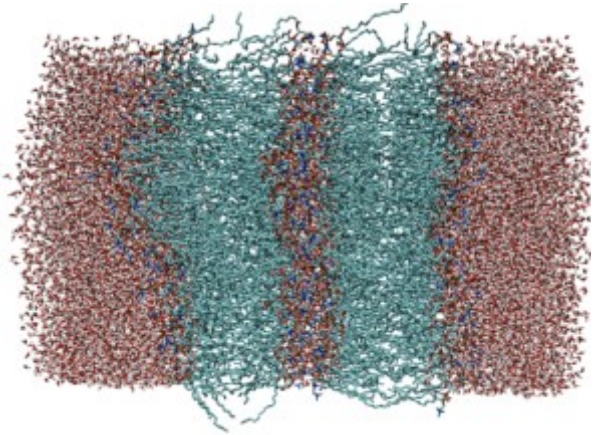
DL_POLY Capabilities
Examples of Model Systems



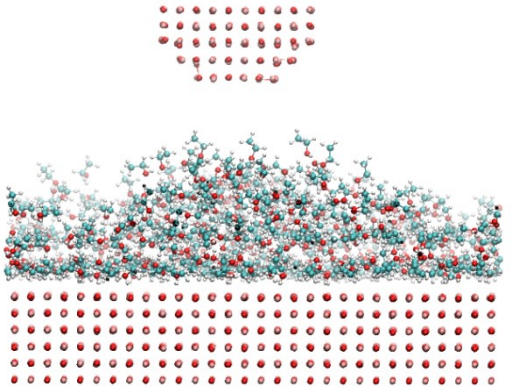
Proteins
solvation & binding



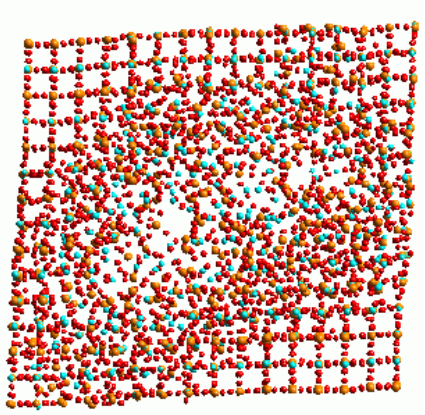
DNA strands
dynamics



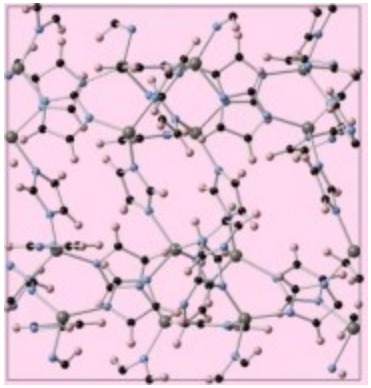
Membranes' processes

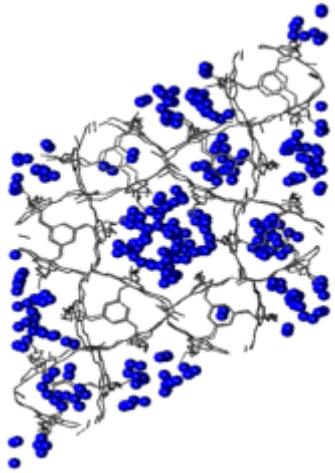


Bio-inorganic models

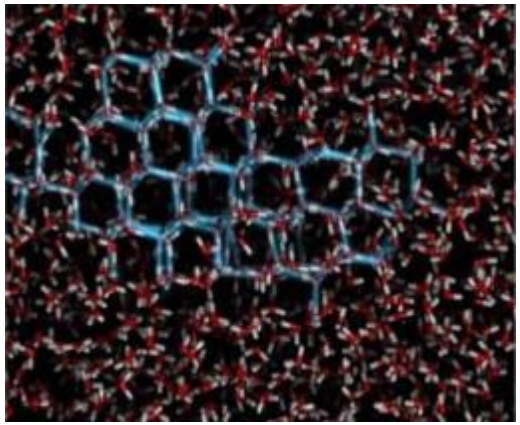


Crystalline & Amorphous
Solids – damage and recovery





Dynamic processes in
Metal-Organic & Organic Frameworks

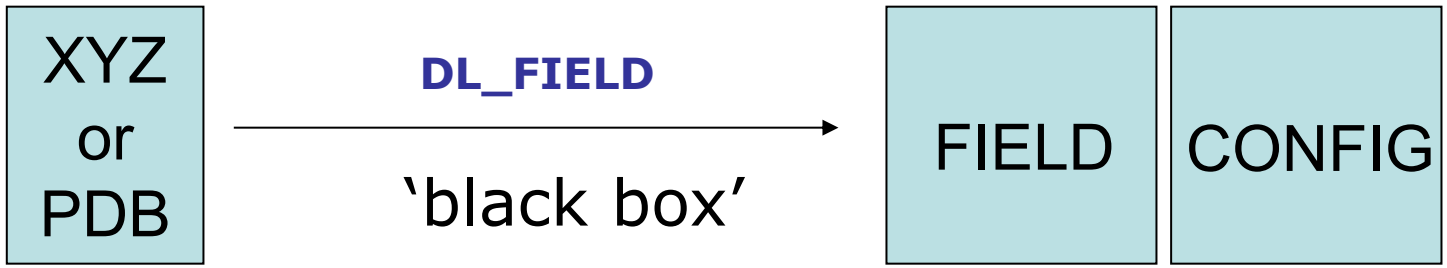


Dynamics at Interfaces &
of Phase Transformations

DL_FIELD

Force field model: a set of equations with fitted parameter values that describe the behaviour of a molecular systems.

- Utility software for setting up force field models.
- Full automatic atom typing an identification of chemical nature of atoms.
- A computer program package written in C that primarily serves as a support application software tool for DL_POLY molecular dynamics (MD) simulation package.
- Important application tool to enhance the usability of DL_POLY MD simulation package and to facilitate the use of a wide range of advance features included in the DL_POLY program.



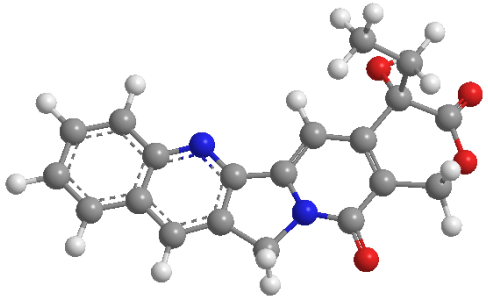
Science & Technology Facilities Council

DL_FIELD Capabilities

Available force field: CHARMM, AMBER, OPLS-AA/2005, DREIDING, PCFF, CVFF, G54A7, Inorganic (clay, oxides, glass). All in a single-unified format.

System models: Proteins, sugars, drugs, polymers, networked structures, random structures, inorganic, organic/inorganic systems.

Implements **DL_F Notation***, a universal standardised atom typesetting expression for molecular simulation.



Camptothecin – a quinoline alkaloid

PCFF

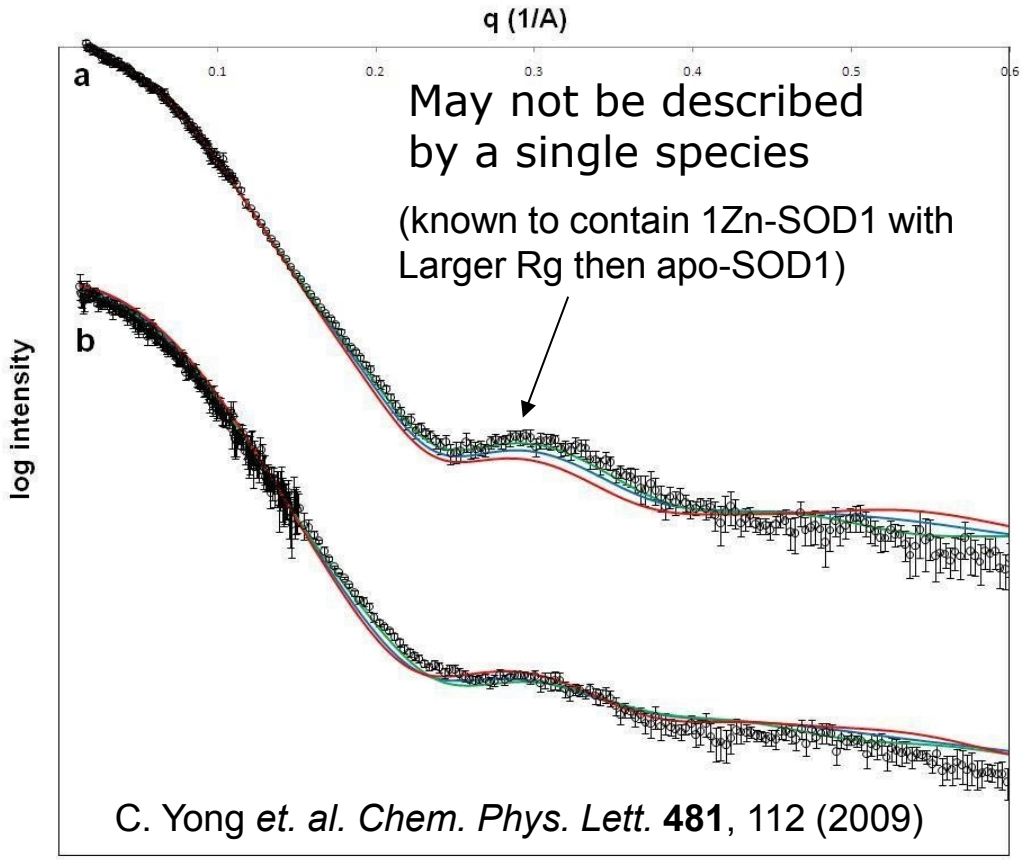
```
Generated by DL_FIELD v3.50
Units kcal/mol
Molecular types 1
Molecule name not_define
nummols 1
atoms 42
cp      12.01150   -0.12680   1   0
cp      12.01150   -0.12680   1   0
cp      12.01150   -0.12680   1   0
cp      12.01150    0.24050   1   0
cp      12.01150    0.00000   1   0
cp      12.01150   -0.12680   1   0
np      14.00670   -0.48100   1   0
...
```

1 - C6_quinoline
2 - C7_quinoline
3 - C8_quinoline
4 - C8a_quinoline
5 - C4a_quinoline
6 - C5_quinoline
7 - N1_quinoline
8 - C2_quinoline
9 - C3_quinoline
10 - C4_quinoline
11 - CE_alkene
12 - Nt_amide
13 - Cs_alkane
14 - CL_alkene
15 - CL_alkene
16 - CE_alkene
17 - C_amide
18 - Cq_alkane
19 - C_ester
20 - OL_ester
21 - Cs_alkane
22 - O_amide
23 - OE_ester
24 - O_alcohol
25 - Cs_alkane
26 - Cp_alkane
27 - HC_aromatic
28 - HC_aromatic
29 - HC_aromatic
30 - HC_aromatic
31 - HC_aromatic
32 - HC_alkane
33 - HC_alkane
34 - HC_alkene
35 - HC_alkane
36 - HC_alkane
37 - HO_alcohol
38 - HC_alkane
39 - HC_alkane
40 - HC_alkane
41 - HC_alkane
42 - HC_alkane

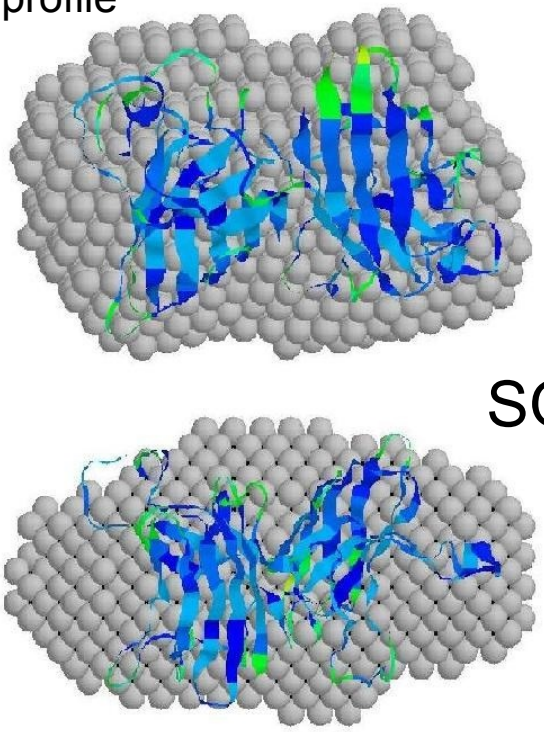
*C. W. Yong, J. Chem. Inf. Model. vol.56, pp 1405-1409 (2016)

Science & Technology Facilities Council

Long time-scale MD to provide global structural information, making comparison with SAXS



MD structures superimpose On 3D reconstructed scattering profile



SOD1

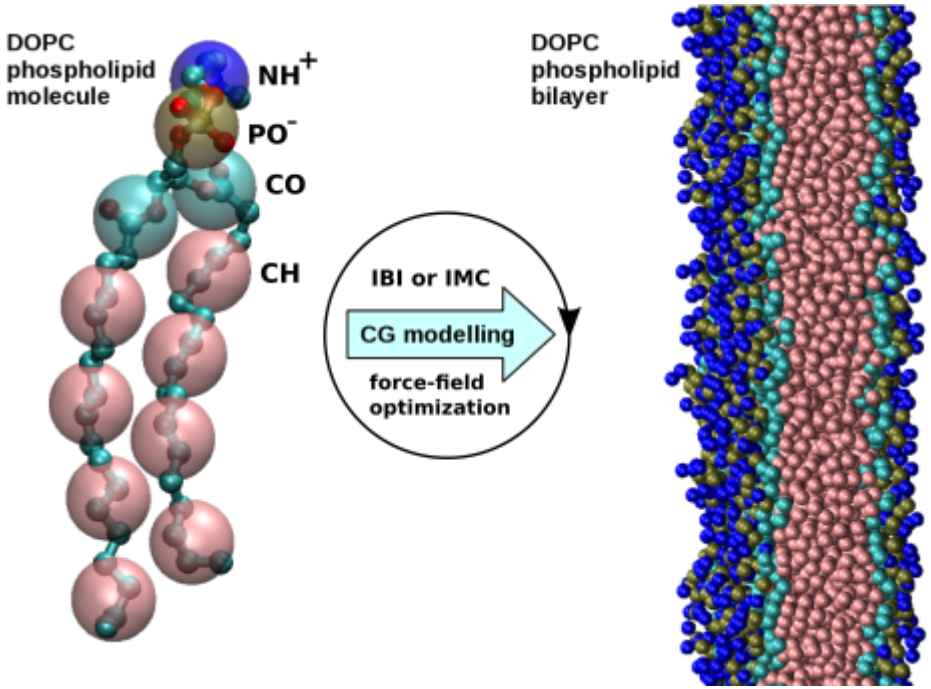
Experimental scattering curves (black with error bars) collected from holo-SOD1 (a) and apo-SOD1 (b) in solution and simulated scattering curves from MD structures:

Get a better fit, if generated a theoretical scattering curve using a combination of the average profiles both from apo-SOD1 and 1Zn-SOD1 structures derived from our MD calculations.

Use MD data as a reference basis to determine the states of protein molecules in solution?

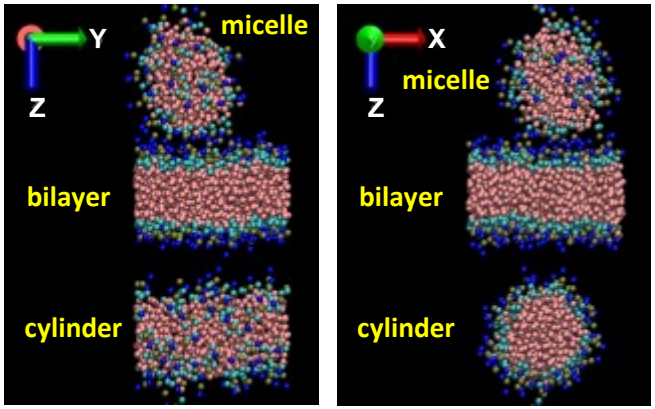
DL_CGMAP

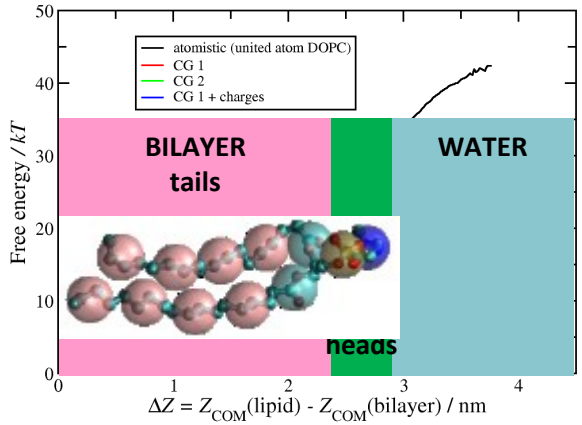
- Interfaces DL_POLY with VOTCA coarse-graining toolkit
  - Developed at Daresbury in collaboration with Universities of Bath and Manchester
  - International collaboration with VOTCA team (Germany, USA, UK)
  - Can determine coarse-grained potentials based on all-atom MD trajectories from DL_POLY
  - Can either use standalone utility or run DL_POLY_4 within VOTCA environment:
  - Both DL_POLY_4 and VOTCA are well-suited to HPC calculations
  - Exploiting SCARF & ARCHER



Developed by A.V. Brukhno

DL_CGMAP





Work by A.V. Brukhno paper is under preparation

- Coarse-graining of DOPC lipid systems

✓ **Molecular Dynamics for united atom model**

– OPLS + Gromos G9643a1 force-field (Gromacs & DL_POLY)

– **2 DOPC lipids in SPCE/TIP4P water, ~4000 molecs**

– *NpT* equilibrated box, $L \approx 50$ Å, $T = 300$-310 K, $p =$ 1 atm

– **2 microsecond MD to collect sufficient stats for RDFs**

✓ **Iterative Boltzmann Inversion & Inverse MC**

– reproducing RDFs for two DOPC lipids

– reproducing bilayer and micellar structures

– **no tweaking needed for pressure correction etc**

✓ **Passing crucial tests:**

– area per lipid is in the correct range: 65 – 70 Å^2

– bilayer thickness is also correct: 2.5 – 3.0 nm

– density distributions across the bilayer are OK

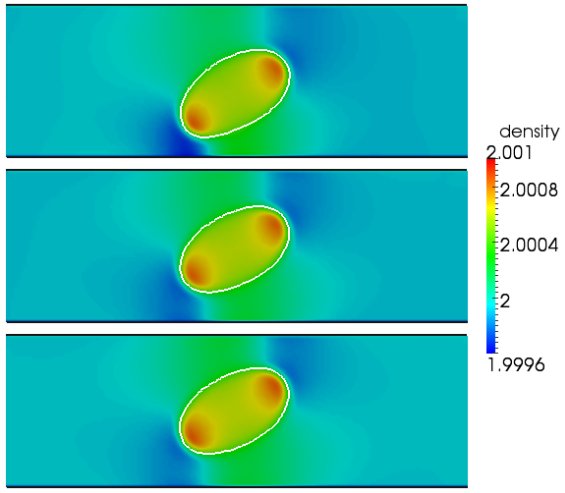
– **strikingly, partitioning FE profile is virtually perfect!**

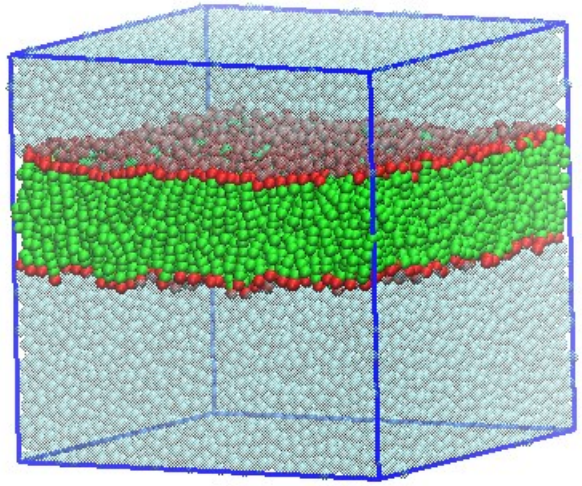
(i.e. pulling a lipid molecule in/out a bilayer)

Science & Technology Facilities Council

DL_MESO

- Mesoscale modelling
  - Between atomistic and fluid dynamics length/time scales
  - Atom-like and fluid-like behaviours
- DL_MESO: mesoscopic modelling software suite
  - Created for CCP5
  - Highly parallelisable codes for Lattice Boltzmann Equation (LBE) and Dissipative Particle Dynamics (DPD)
  - Both methods can model interactions between species with correct hydrodynamic behaviour





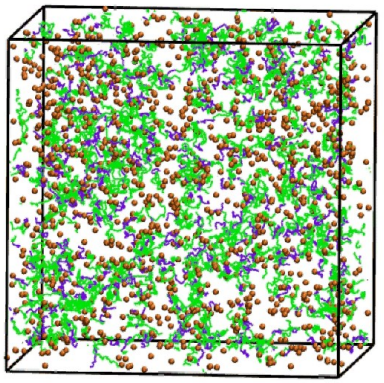
Source: Seaton *et al.*, *Mol Sim*, **39** (10), 796-821 (2013)

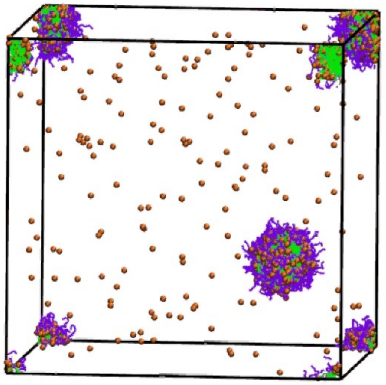
www.ccp5.ac.uk/DL_MESO

Science & Technology Facilities Council

DL_MESO

- Dissipative Particle Dynamics (~μs)
  - Similar to classical Molecular Dynamics: soft repulsive particles
    - Larger timesteps available than atomistic MD, parameterise for e.g. hydrophobicity
  - Pairwise thermostat ensures Galilean invariance: correct hydrodynamics
  - Can include bonded interactions, density-dependent potentials, (smeared) charges
  - Example: loading and release of drug in co polymer vesicles (related to solution pH)

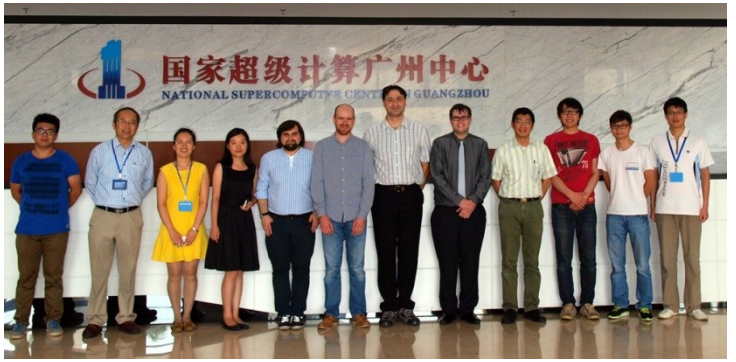




Source: Luo and Jiang, *J Control Release*, **162**, 185-193 (2012)

Supporting software ('after-registration' services)

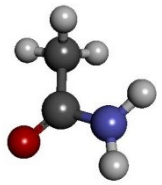
- Documentation
  - User manuals, websites
- Mailing lists
  - Report bugs and new version releases
- Emails to developers
  - Under CoSeC, priority given to UK academics
- DL Software Workshops
  - Held three times per year at different locations in UK and abroad
  - Give current/new users opportunities to try out software and modelling techniques
  - 'Hack days' show how DL software can be modified for new science



Where do we go from here….

Science & Technology Facilities Council

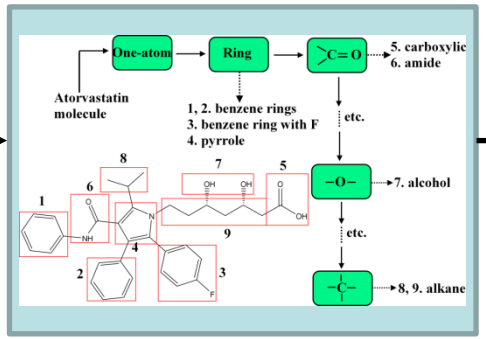
Molecular simulation engine infrastructure



User's structure

```
O  2.129000  0.063000  2.421000
C  2.004000  0.182000  1.026000
H  1.513000 -1.209000 -0.602000
C  1.611000 -1.202000  0.507000
C  0.907000  1.179000  0.657000
H  0.657000  1.067000 -0.426000
C  1.146000  2.672000  0.936000
H -0.012000  0.872000  1.214000
C  2.015000  2.959000  2.178000
H  0.148000  3.161000  1.037000
```

DL_FIELD



FF model

DL_POLY

Force field model,
Configuration file

Dynamic computations

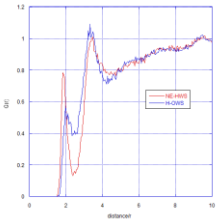
Molecular trajectory files

DL_ANALYSER

Filter module

Formatted atomic structure

Data analysis



Structure extraction

[Template file]

Future development

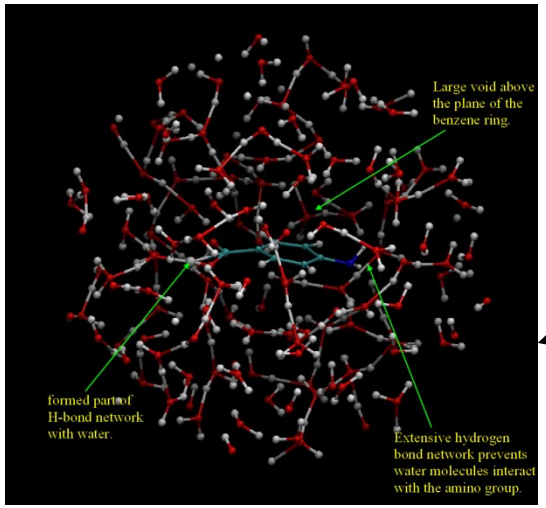
DL_MONTE (monte carlo)

ChemShell (QM/MM)

Force field model

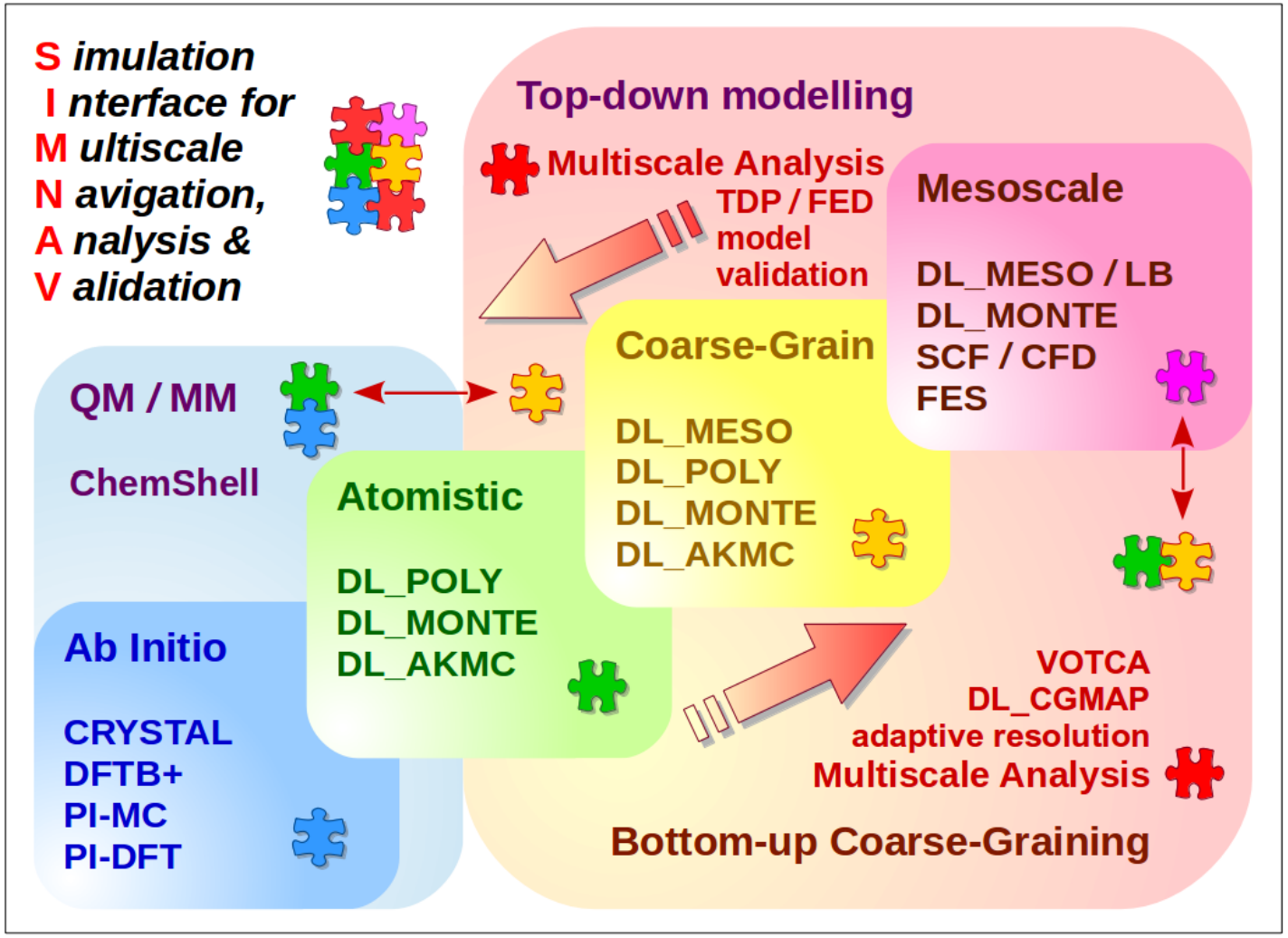
Configuration file

Control file.



https://www.ccp5.ac.uk/software

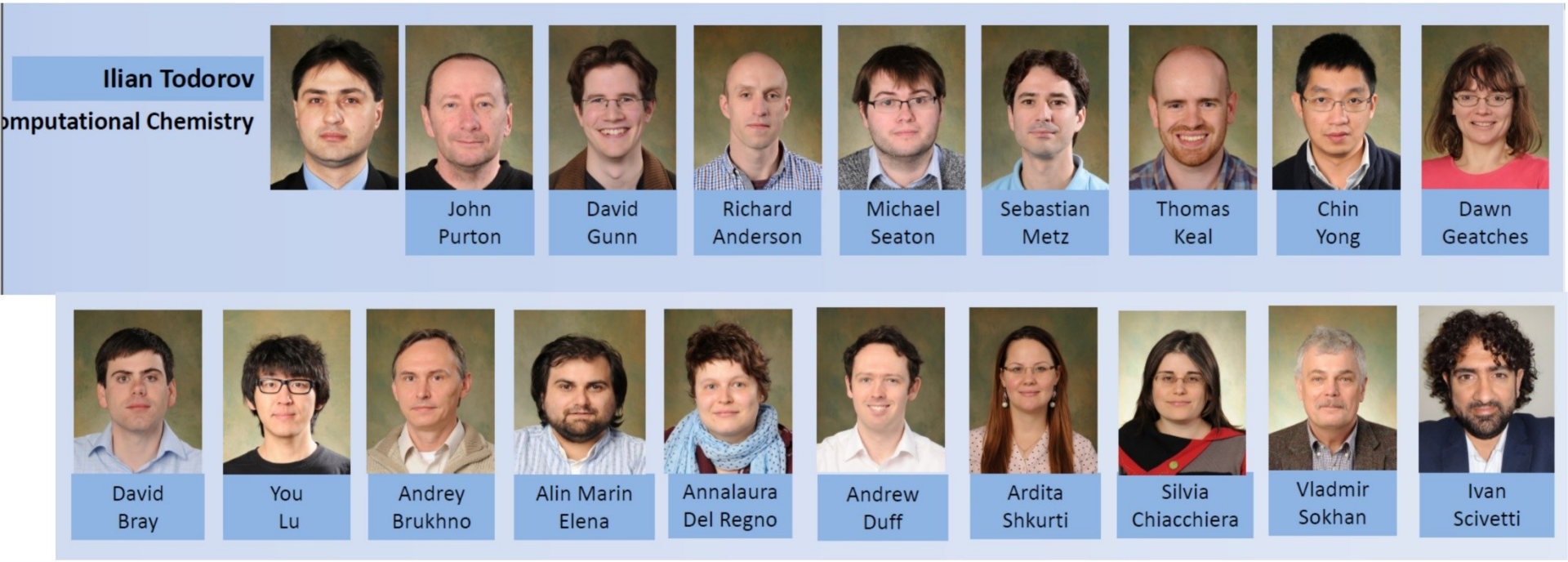
A bigger picture....



RSE Proposal by Andrey Brukhno

Acknowledgements

The CCG Team



CCPs, HECs, Hartree Centre.