

# Rapid Parameterisation of Small Molecules

Dr Benjamin Roberts  
[\(ben.roberts@nesi.org.nz\)](mailto:ben.roberts@nesi.org.nz)



# Outline

## ① The Background

Introduction to Molecular Mechanics  
Force Fields and Parameterisation

## ② The Challenge

## ③ How We're Tackling It

Overview  
HPC-powered QM Refinement

## ④ Wrap-Up

# Molecular Mechanics

## The what and the why

- An approximation of a molecule's potential energy using a Newtonian model
- Add a time component to get molecular dynamics (MD)
- Useful for large systems (thousands of particles or more)
- Commonly used in biomolecular simulation, drug design, materials science, etc.

# Molecular Mechanics

## The how

- The system is represented in three parts: Coordinates, topology and force field
- Coordinates: Where all the particles are (relative to each other)
- Topology: Which particles are connected to which other particles
- Force fields: What's the deal?

# Force fields



A force field is an approximate mathematical model of the potential energy surface of the system. It can be expressed as an equation in  $3N - 6$  variables, where  $N$  is the number of particles.

# Force fields

## Functional form

- The equations used to model interactions between particles
- All interactions of the same class will use the same basic model (e.g., bonds commonly use a quadratic equation in the manner of Hooke's law of spring elasticity)

## Parameters

- The constant terms used in the equations
- Within any given class, each specific type of interaction will have its own parameters
- Individual interaction types are often quite fine-grained

# The Case of the Missing Parameters

## The Problem

- The equation for the potential energy is a vast sum of terms
- Each term makes use of at least one parameter (usually more than one)
- If any parameter in any term is missing, the equation can't be solved!
- Parameters don't exist for most molecules.

## The Solution

- Parameterise on the fly...
- ...right?

# The Case of the Missing Parameters



# The Case of the Missing Parameters



# A Four-Step Approach

Our approach to parameter generation consists of four steps:

- ① Generate: Automated Topology Builder (ATB)  
<http://compbio.biosci.uq.edu/atb/>
- ② Refine: GAMESS-US quantum mechanical calculations
- ③ Assemble: MAGIC
- ④ Test: MD simulations

# Refinement

- This is the slow part
- Up to 32 QM calculations per torsional degree of freedom
- The speed of each calculation depends on method and basis set
- Best if the calculations can be done in parallel

# Workflow Development

## Parallelisation

- Within each calculation: Run GAMESS in parallel (MPI parallelisation)
- Between calculations: Run calculations for different torsional angles simultaneously

## Hardware selection

- Based on fitness for purpose (from benchmarking work)
- Goal: Fast, straightforward workflow
- AIX on POWER6 versus Red Hat on Intel: Tests run on Fitzroy (NIWA HPCF) and Pan (University of Auckland)

# Results So Far

- GAMESS works faster on Fitzroy than on Pan
- However, the GROMOS software (required by another part of the workflow) only compiles on Pan
- Preliminary results suggest Fitzroy is the best facility for this work, with small amounts of time needed on Pan at the end
- Workflow is still being developed but should facilitate rapid parameterisation of novel molecules
- NeSI CST members are available for consultation regarding machine selection for your workflow

# Acknowledgements

- Jane Allison
- Ivan Welsh
- François Bissey
- Fabrice Cantos

# Questions & Answers

