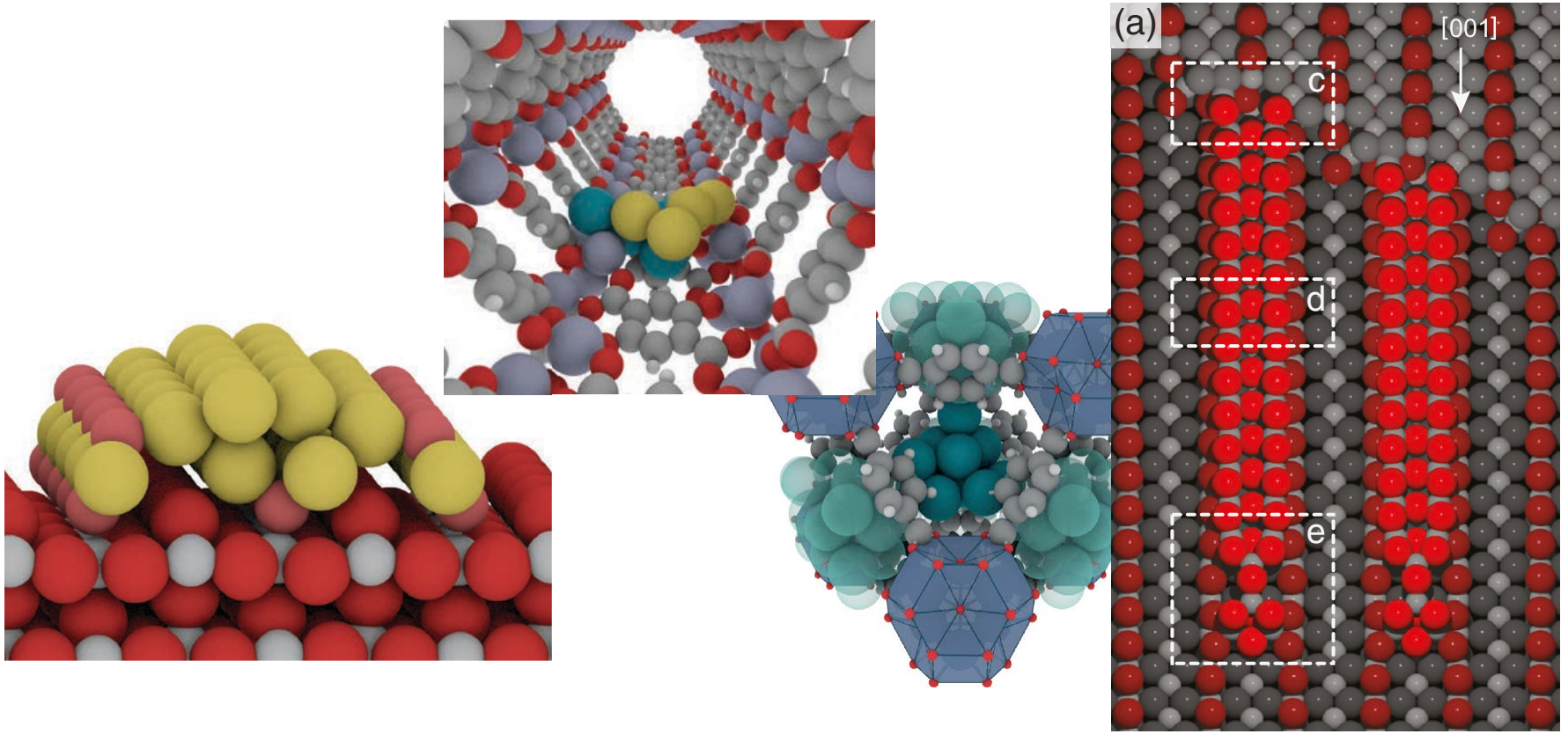
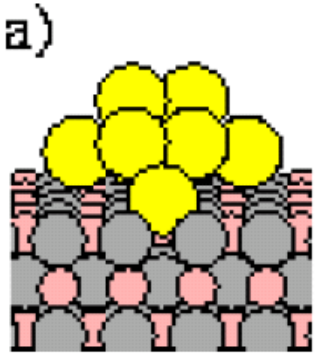
Automated Two Step Structure Prediction within GPAW



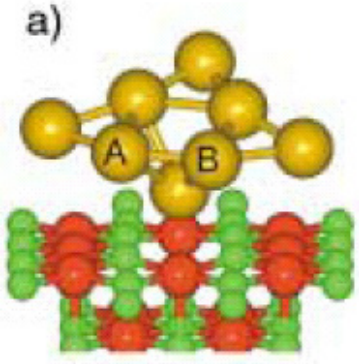
Lasse B. Vilhelmsen
and Bjørk Hammer

Why We Need Automated Methods



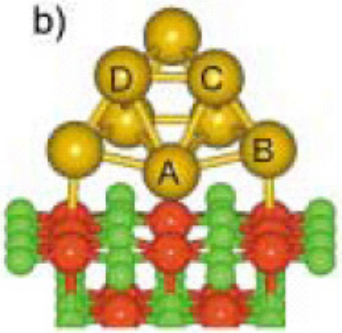
2007

$\Delta E = -4.2$ eV



2003

$\Delta E = -3.8$ eV



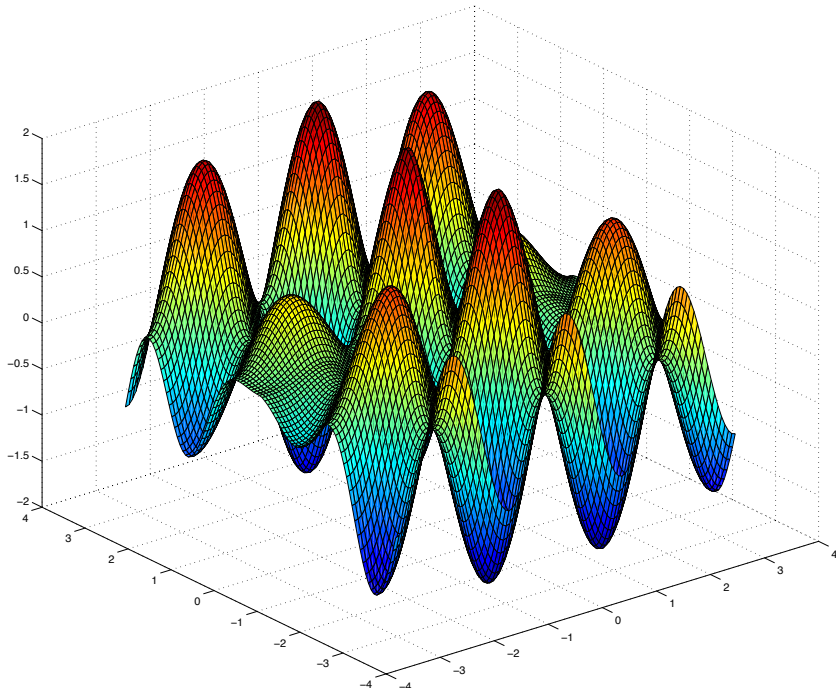
1999

$\Delta E = -3.6$ eV

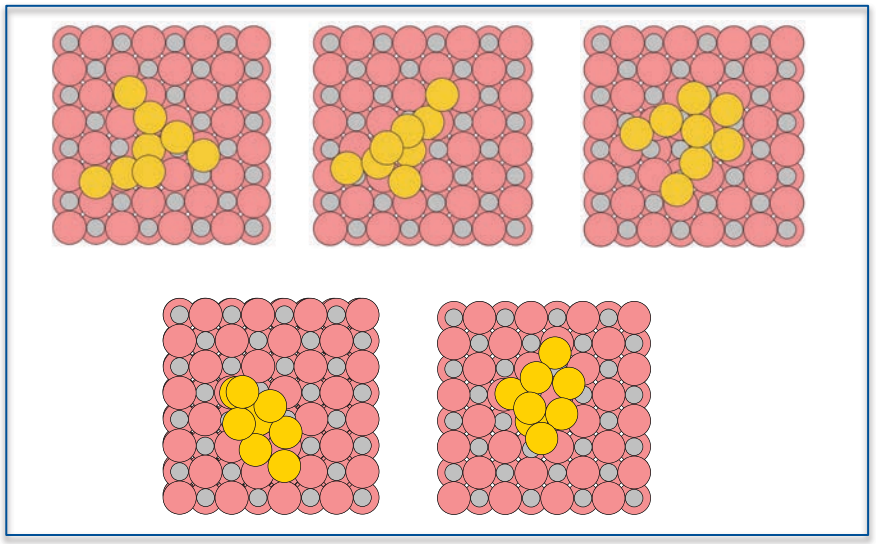
J. Wang and B. Hammer, Top. Catal. 44, 49 (2007)
H. Häkkinen, S. Abbet, A. Sanchez, U. Heiz, and U. Landman, Angew. Chem. 42, 1297 (2003).
A. Sanchez et al., J. Phys. Chem. A 103, 9573 (1999).

Genetic Algorithms

- A method for finding global minima
- Follows a Darwinian evolution scheme
- Based on physical intuition and no formal convergence criteria

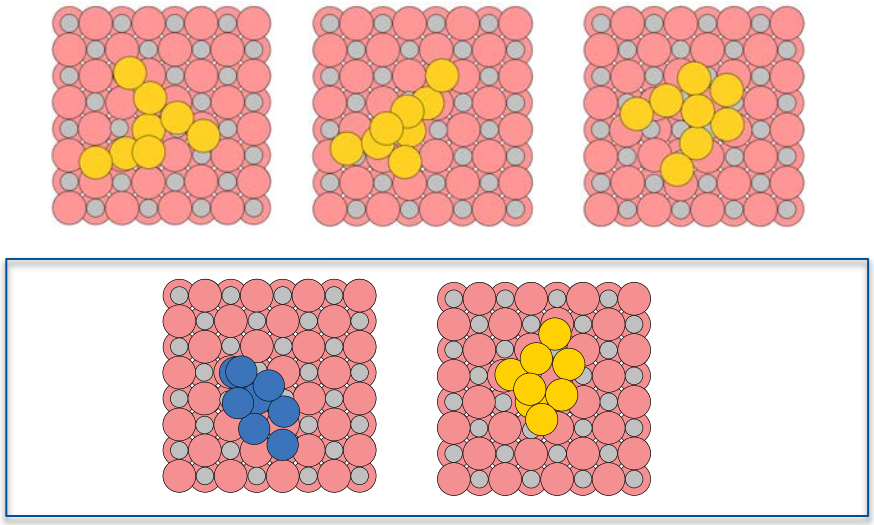


Genetic Algorithms



Generate random start population

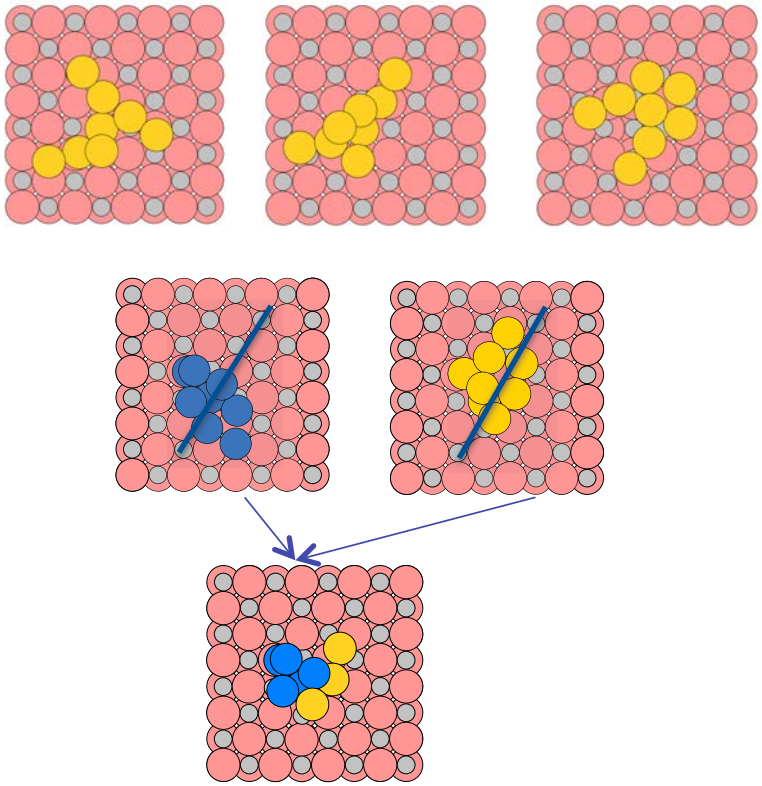
Genetic Algorithms



Generate random start population

Select two parents for mating

Genetic Algorithms

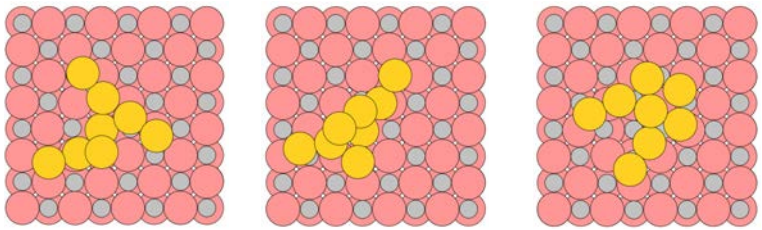


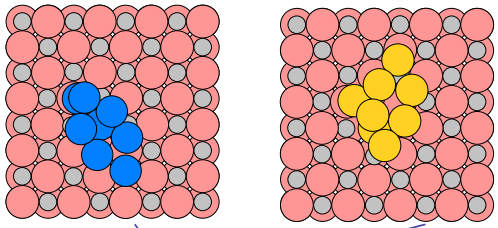
Generate random start population

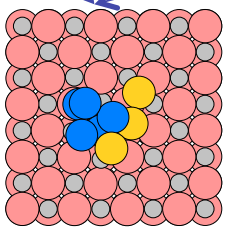
Select two parents for mating

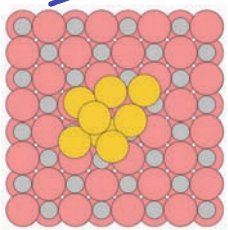
Mate the parents and perhaps mutate

Genetic Algorithms









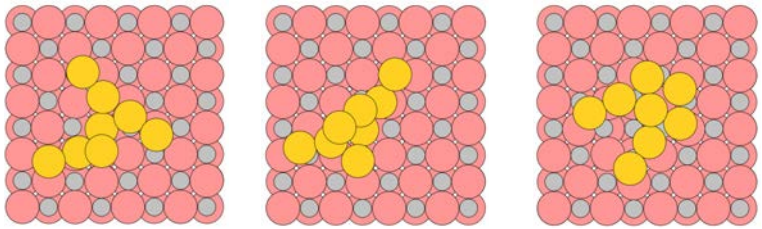
Generate random start population

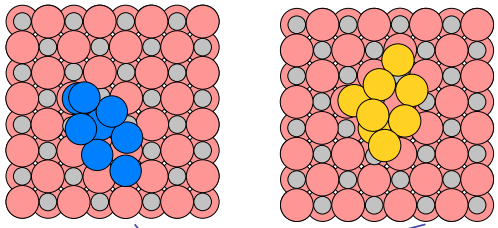
Select two parents for mating

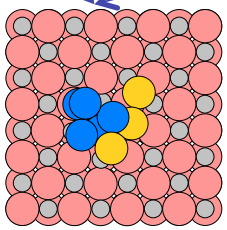
Mate the parents and perhaps mutate

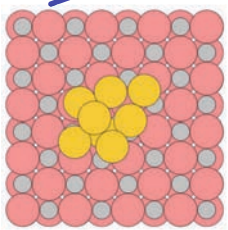
Relax to local minimum

Genetic Algorithms









Generate random start population

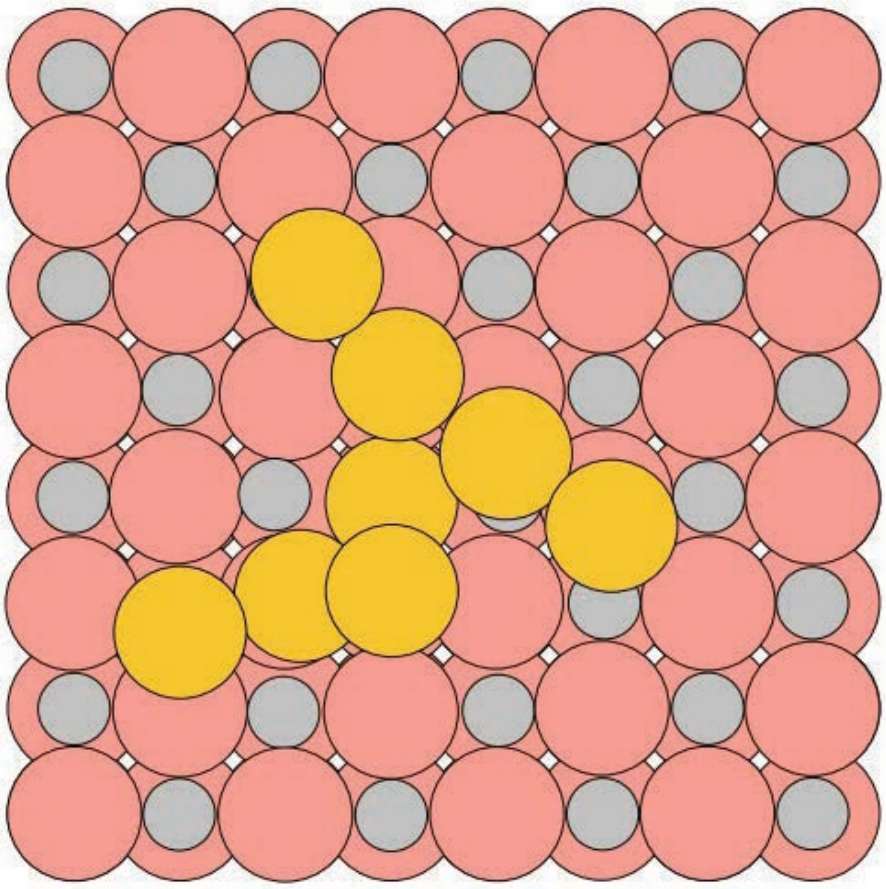
Select two parents for mating

Mate the parents and perhaps mutate

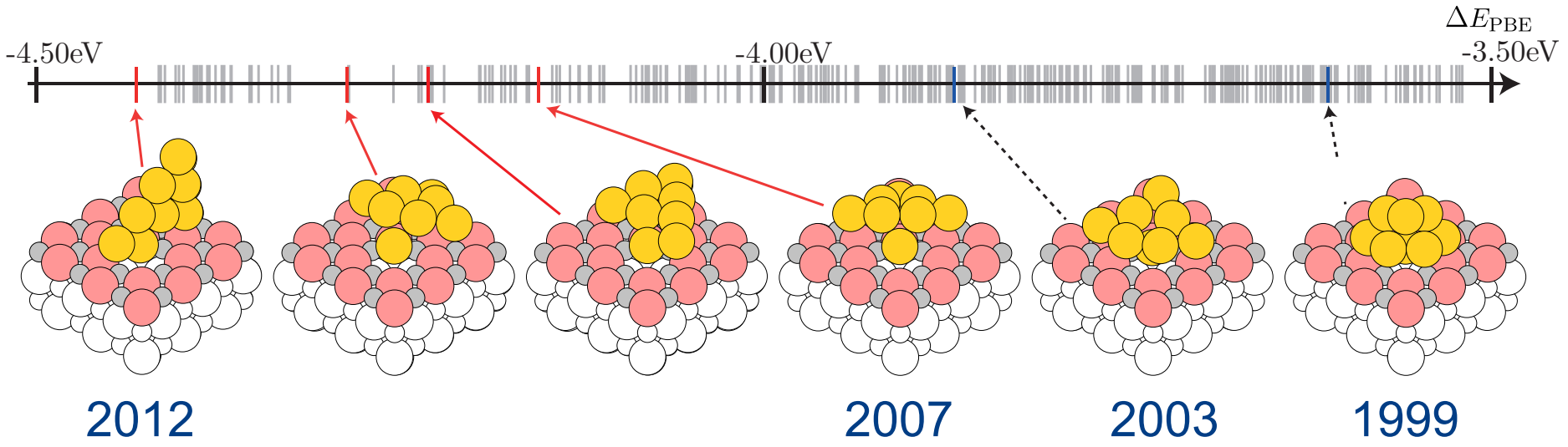
Relax to local minimum

Adopt the structure in to the population

Encountered Au_8 structures

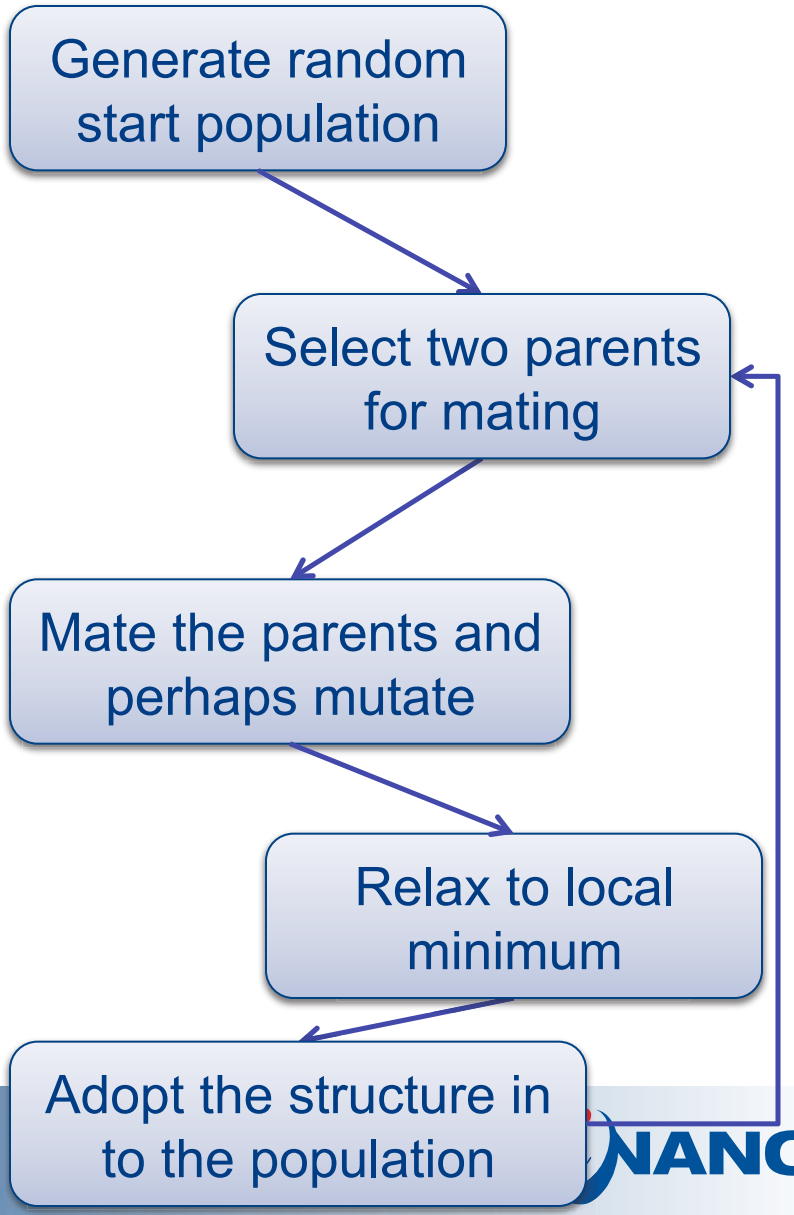


Au_8 Structures

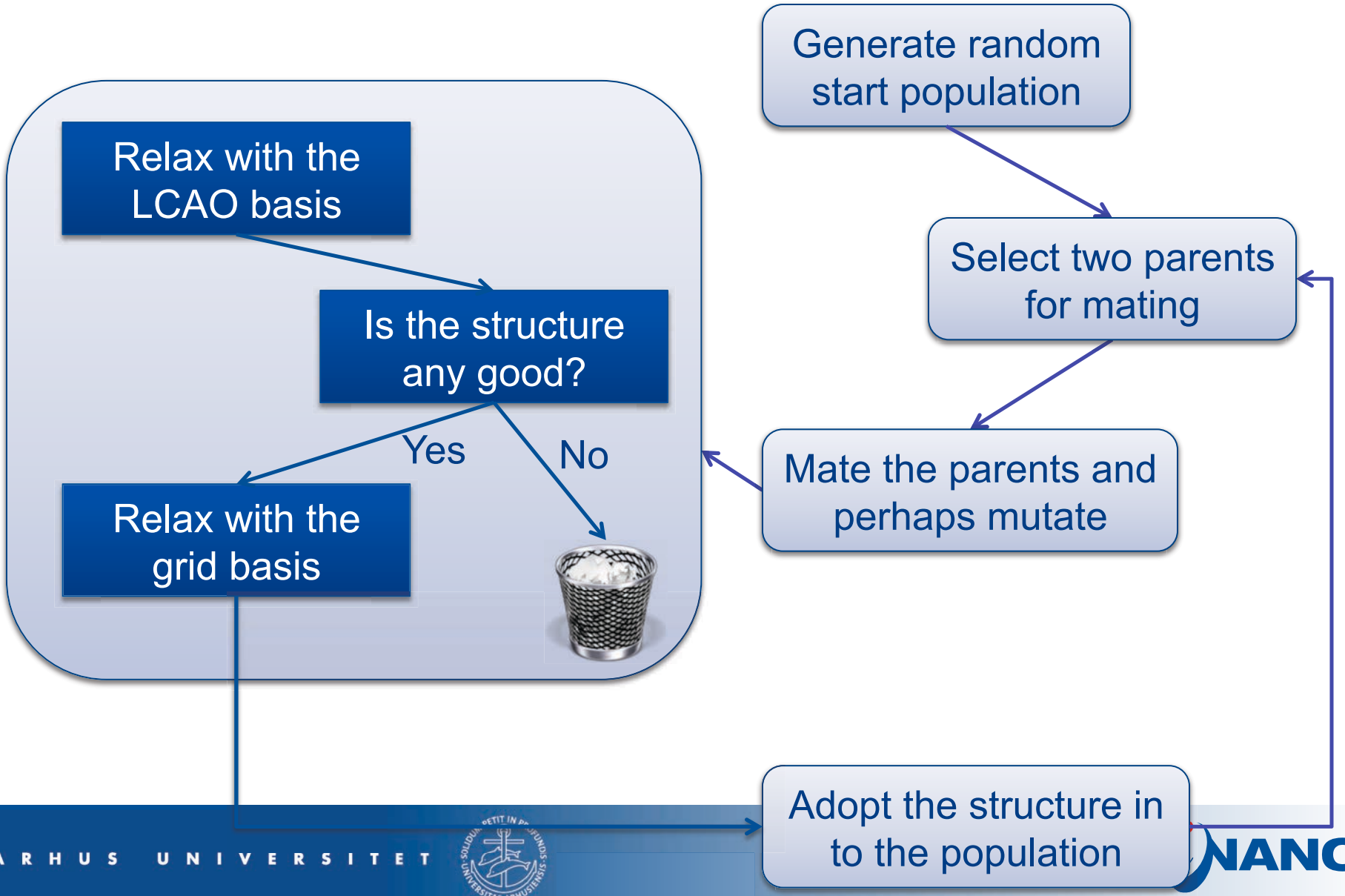


LB Vilhelmsen and Bjørk Hammer, PRL 108, 126101 (2012)

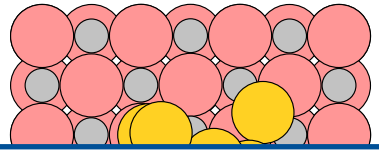
Two Step Optimization Technique



Two Step Optimization Technique

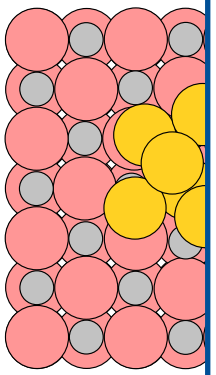


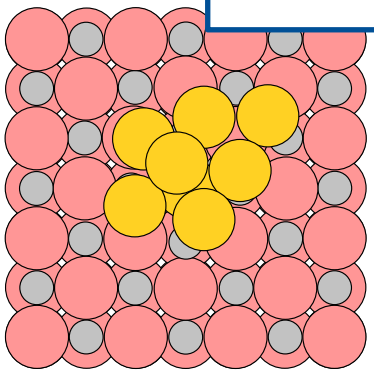
CPU Time Used



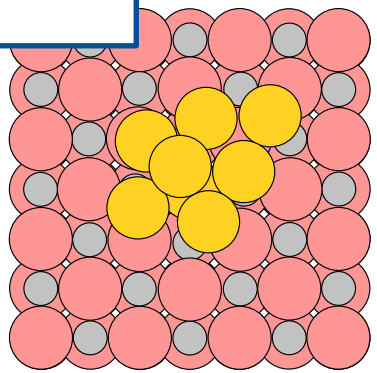
A structural comparison removes 75% of the structures before FD relaxation.

Net speedup of 7

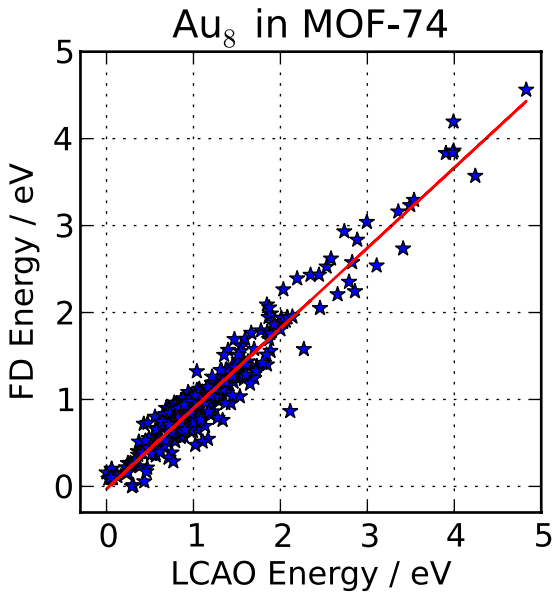


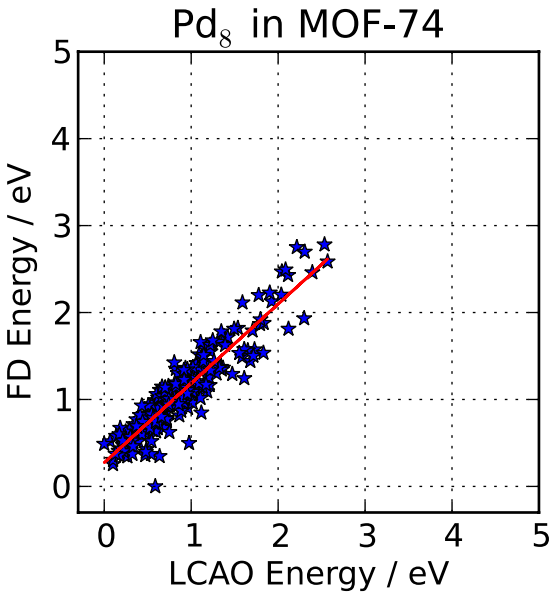


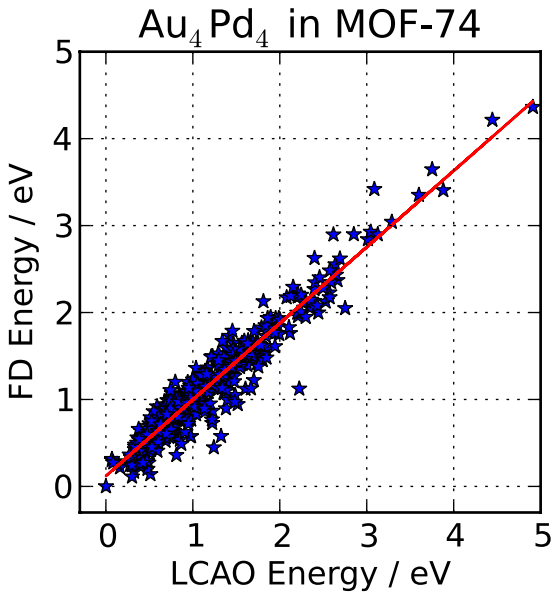
17 steps
148 CPU hours

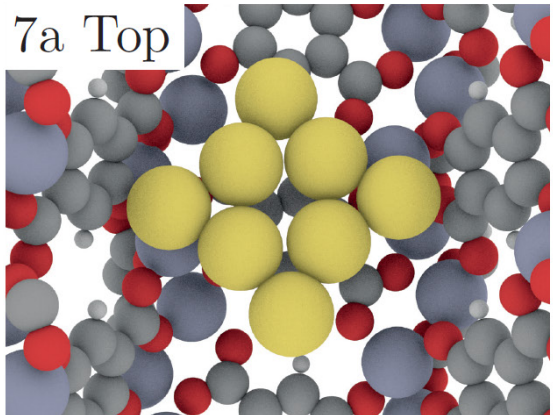


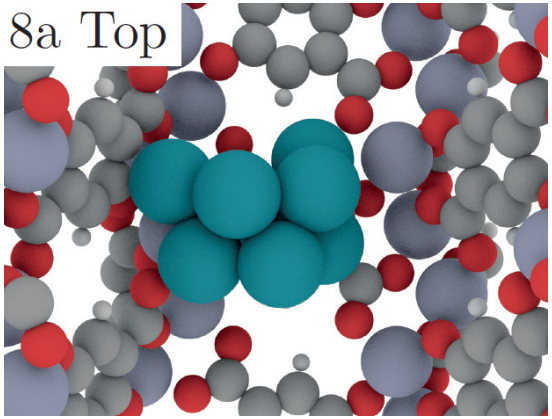
Au, Pd and Au/Pd in MOF-74

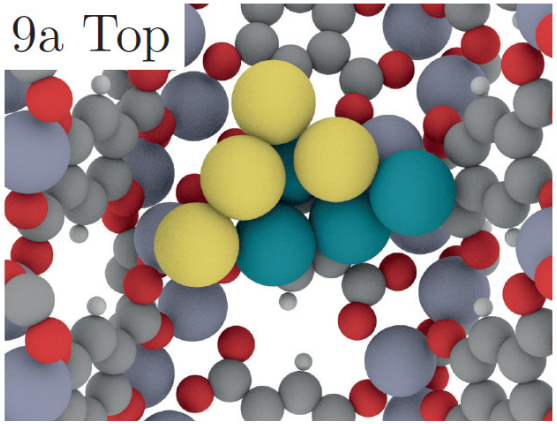




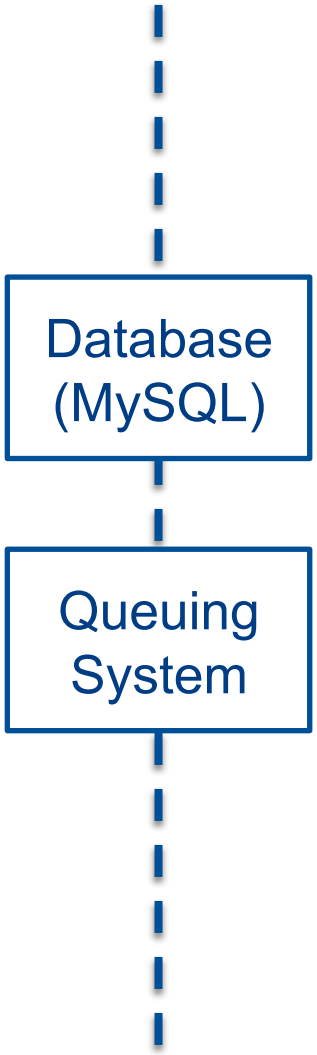






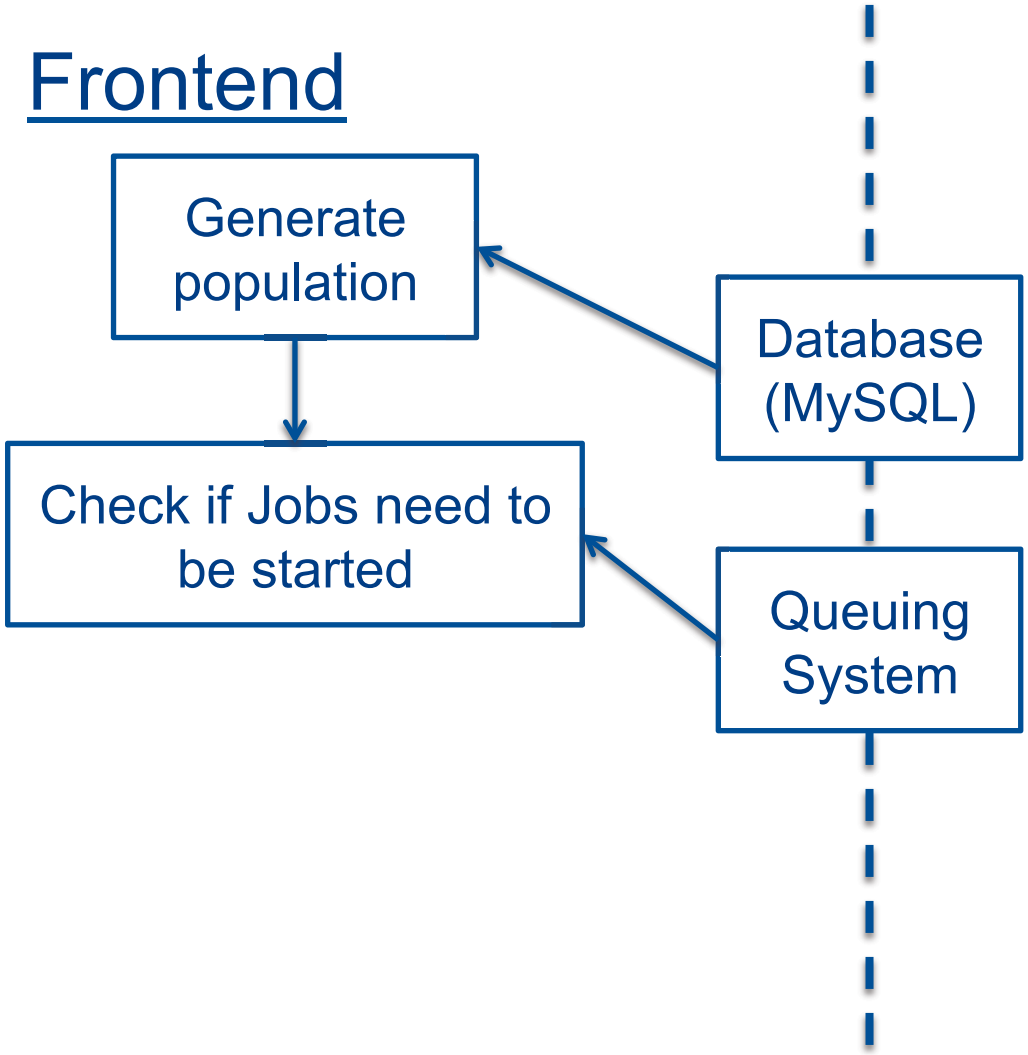


Parallelization



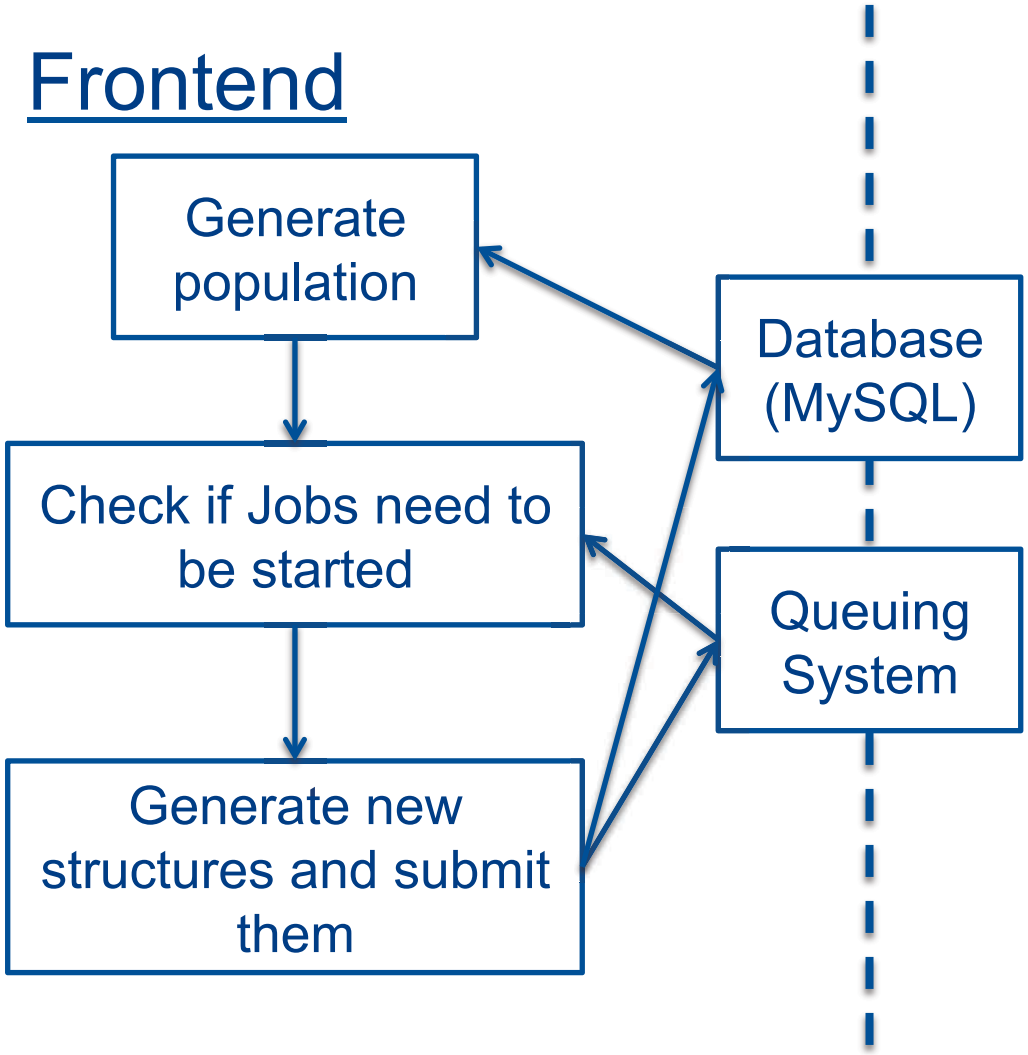
Parallelization

Frontend

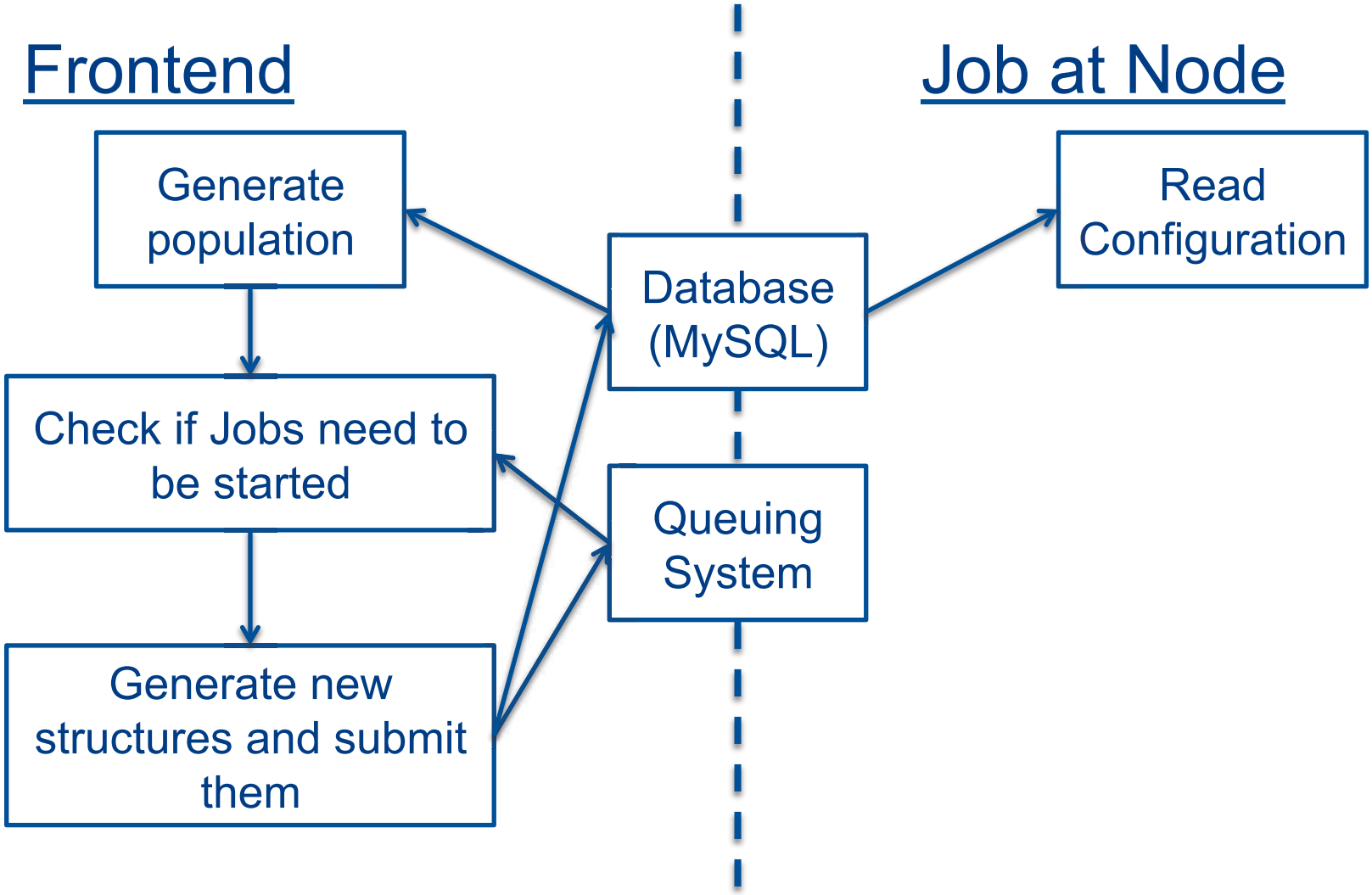


Parallelization

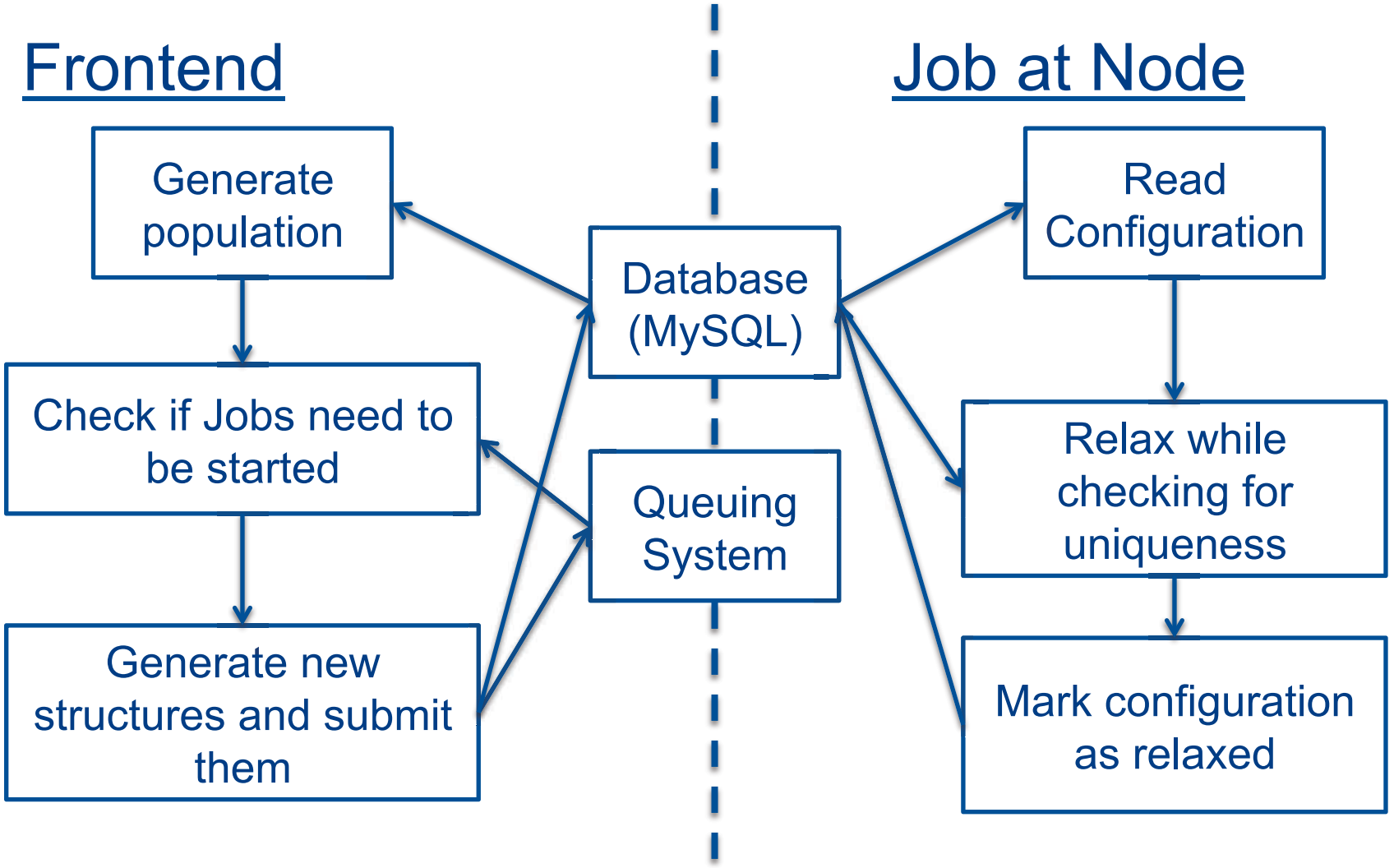
Frontend



Parallelization



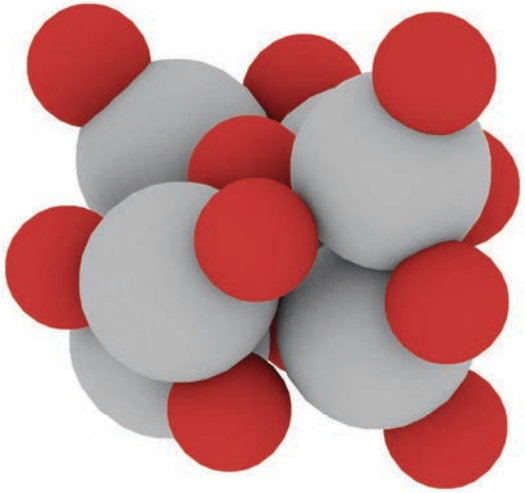
Parallelization



The Challenge of Testing

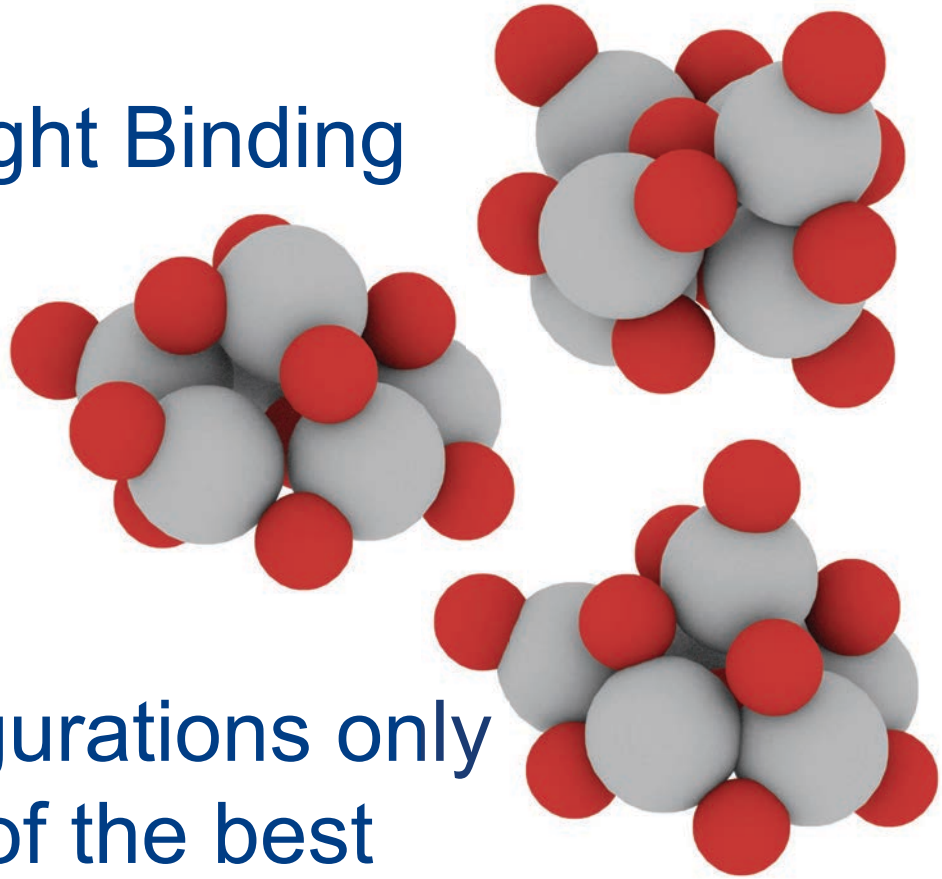
A GA run includes many random factors

Only multiple runs can test the performance



The System To Test With

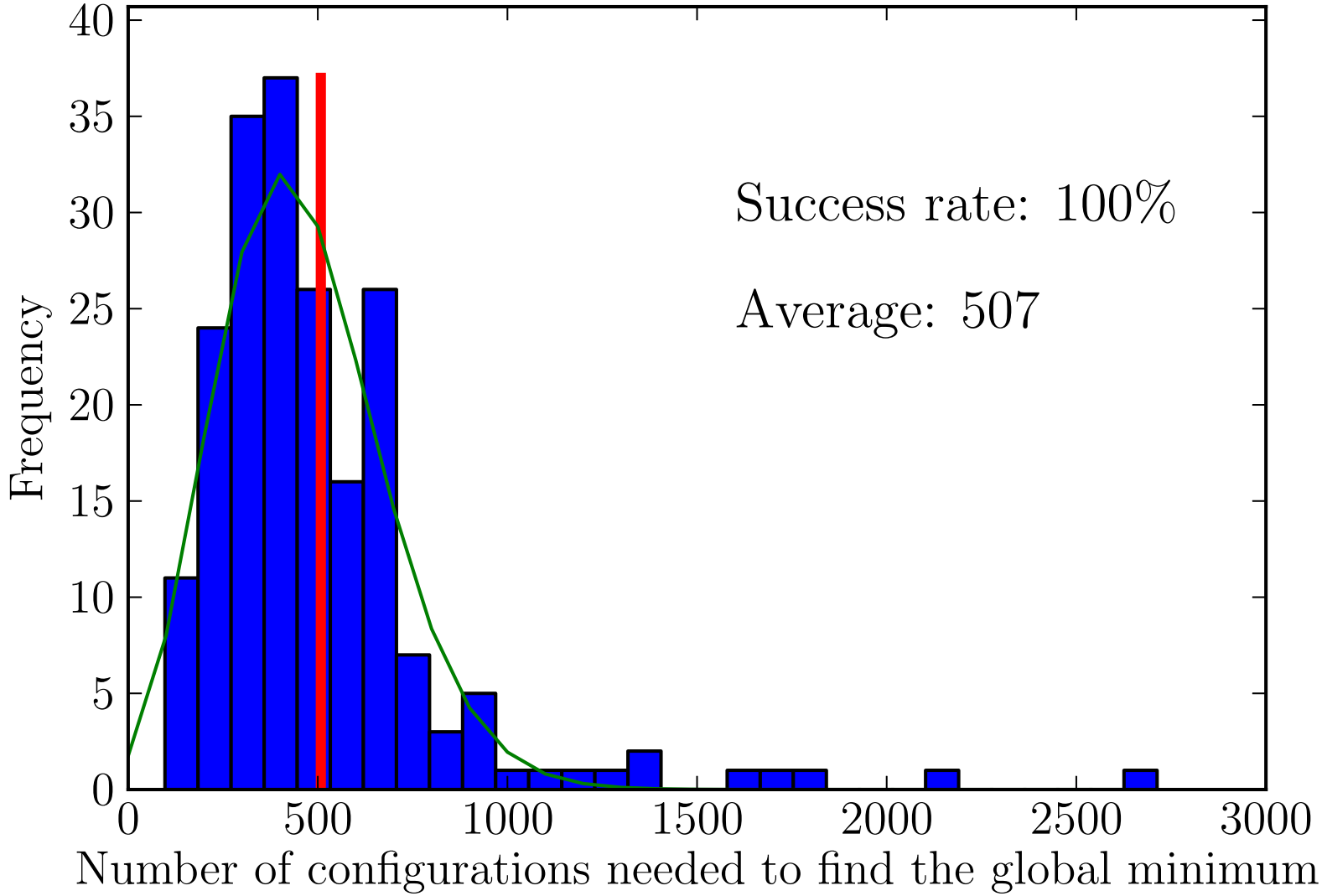
Density Functional Tight Binding calculations (DFTB)



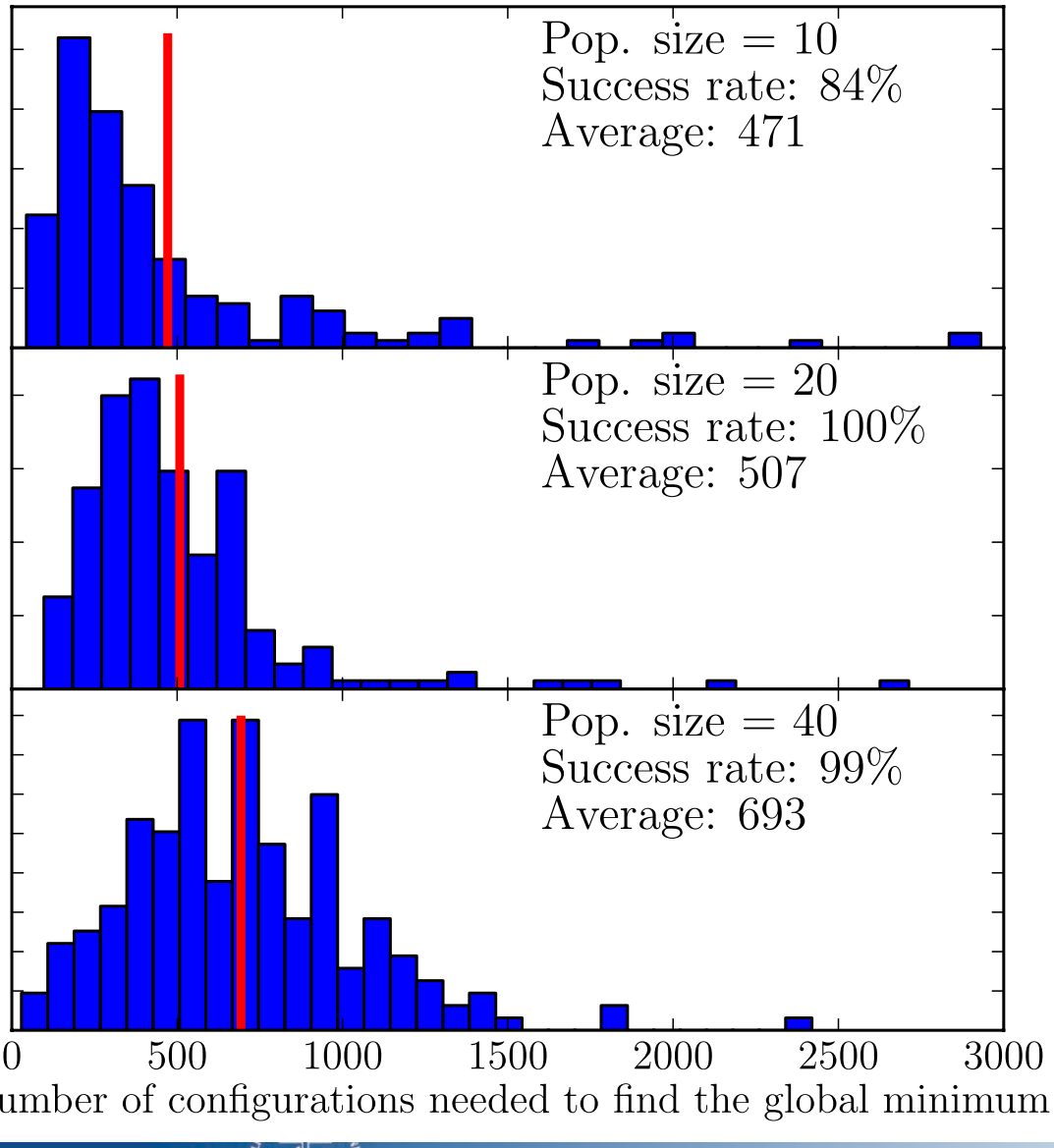
Ti_6O_{12} cluster

10.000 random configurations only come within 0.66 eV of the best configuration

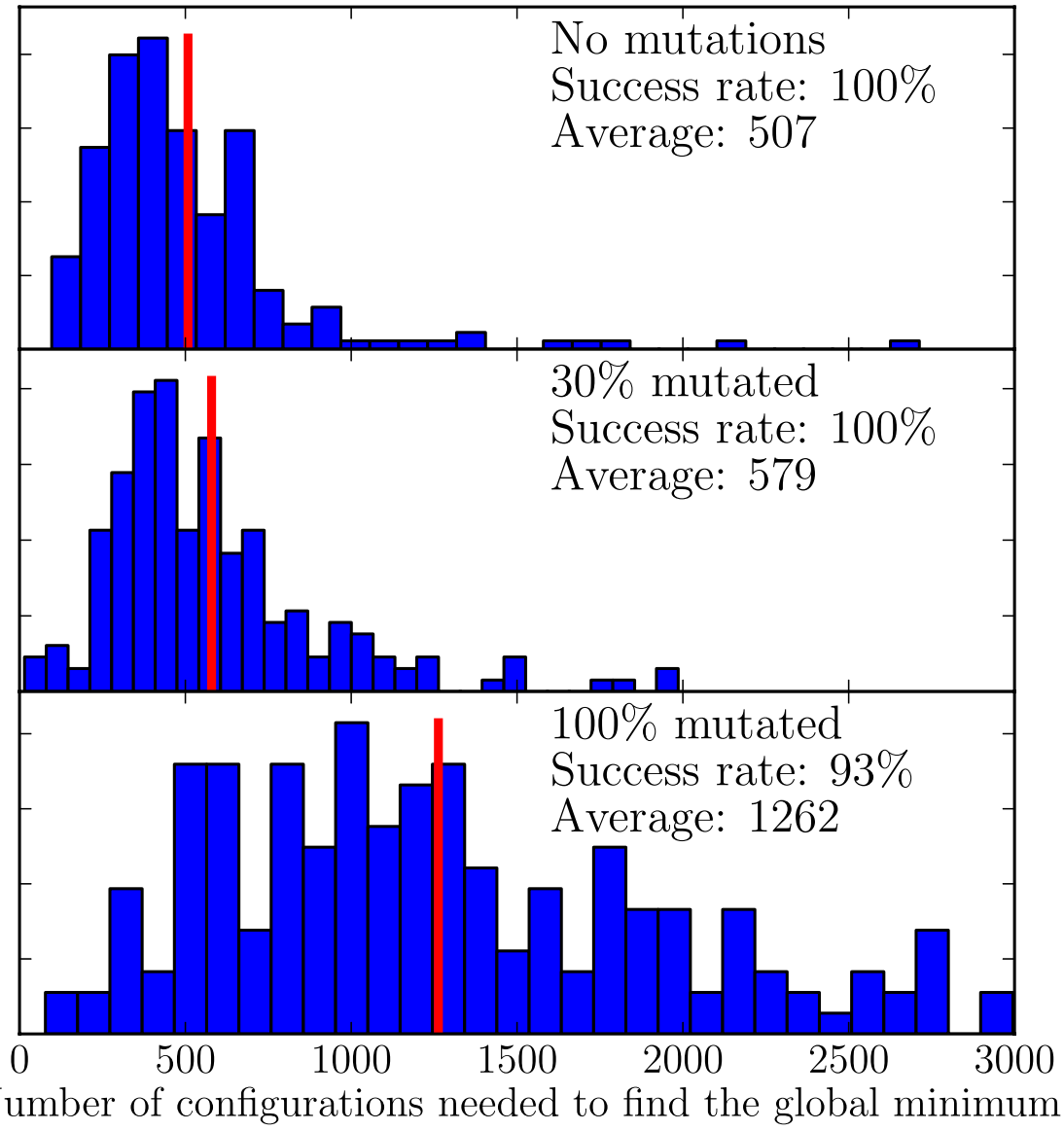
Distribution of Attempts



Importance of Population Size



Should We Use Mutations?

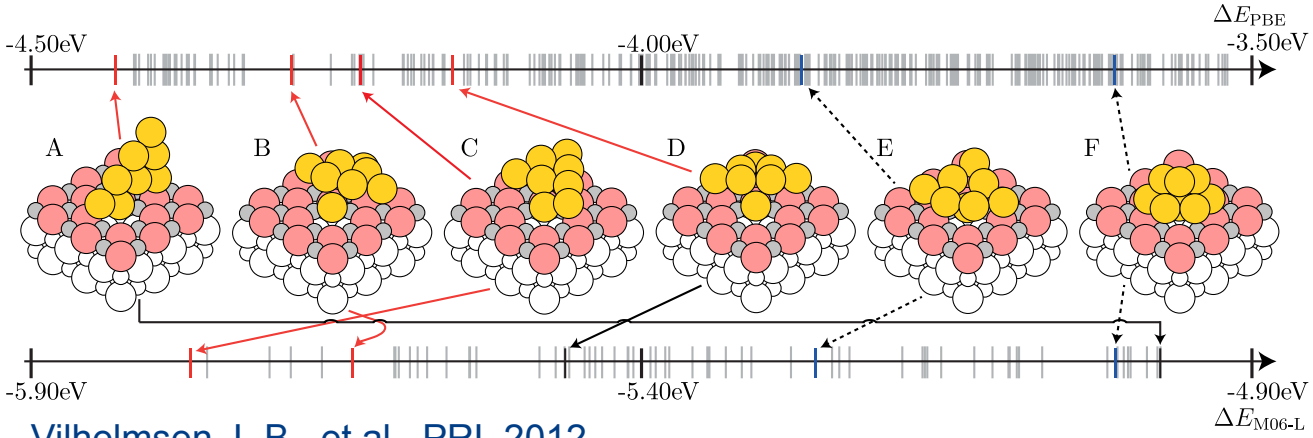


Conclucions

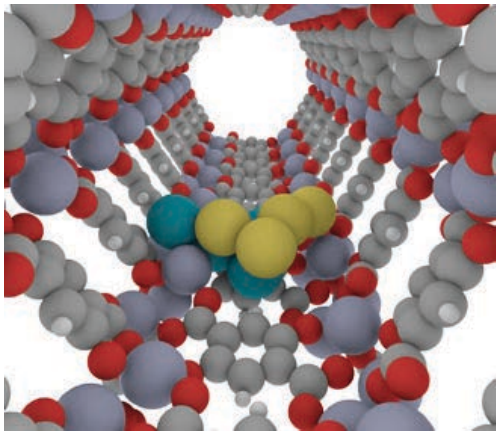
- Energies and forces correlate well between the LCAO and FD bases
- The GA is highly successful in predicting structures across many different system types
- Using fast methods one can investigate the GA performance

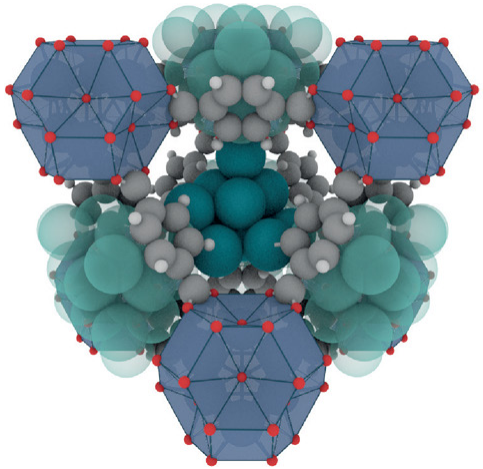
Acknowledgements

Successes

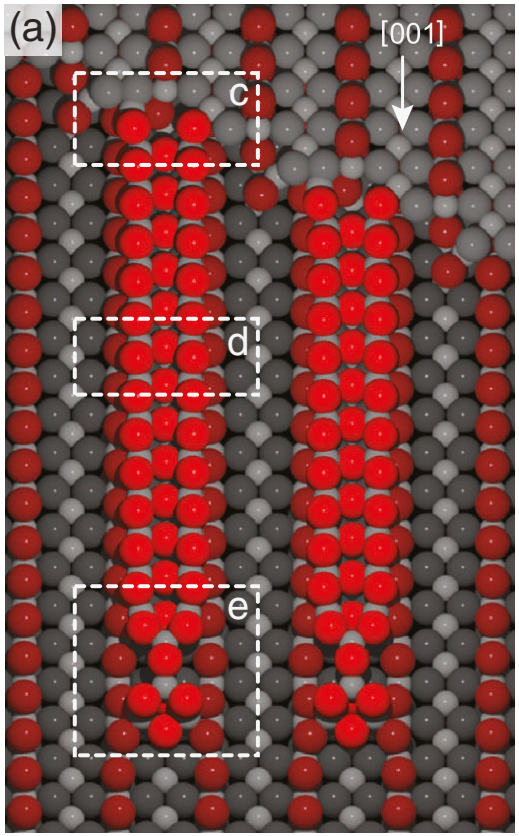


Vilhelmsen, L.B., et al., PRL 2012



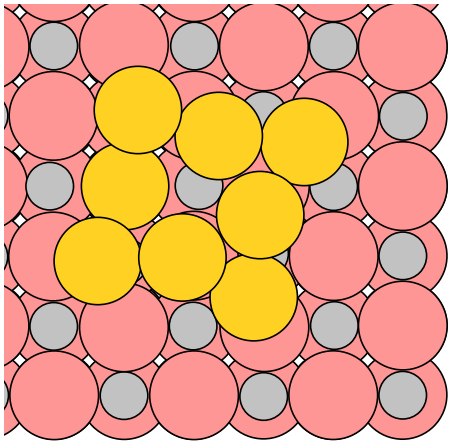


Vilhelmsen, L.B. et al. JACS, 2012
Vilhelmsen, L.B. et al. JPCL, 2012

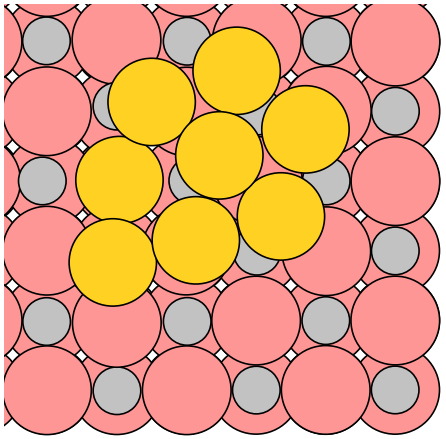


Bechstein, R. et al., PRL 2012
Martinez, U., et al., PRB 2011

Similiarity Criteria



$$\Delta E_i = 0.7\text{meV}$$
$$d_{\text{max,i}} = 1.026\,\text{Å}$$
$$\text{d}_{\text{rel},i} = 0.002$$



Energy criteria:

$$d_i = \left|E_j - E_i\right| > \Delta E \text{ for all } E_j$$

Structural criteria:

$$d_{rel,\text{i}} = \frac{\sum_k \left|D_i(k) - D_j(k)\right|}{\sum_k D_i(k)} < d_{\text{rel}} \text{ and}$$

$$d_{\text{max,i}} = \max\left(\left|D_i - D_j\right|\right) < d_{\text{max}}$$