Introduction to ONETEP

Chris-Kriton Skylaris

c.skylaris@soton.ac.uk

UNIVERSITY OF Southampton
School of Chemistry

Outline

- Overview of basic concepts
  - OBDMM approaches
  - NGWFS, density kernel
  - Psinc basis set
  - FFT box
  - Linear-scaling examples
  - Parallel scaling
- Functionality available
- Compilation requirements
- Running a simple calculation

Optimal basis density matrix minimization (OBDMM) approaches

(S. Goedecker, *Rev. Mod. Phys.*, **71**, 1085 (1999))

E. Hernandez and M. J. Gillan, *Phys. Rev. B* **51**, 10157 (1995).
J.-L. Fattebert and J. Bernholc, *Phys. Rev. B* **62**, 1713 (2000).
C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, O. Dieguez and M. C. Payne, *Phys. Rev. B* **66**, 035119 (2002).

Efficient linear-scaling:

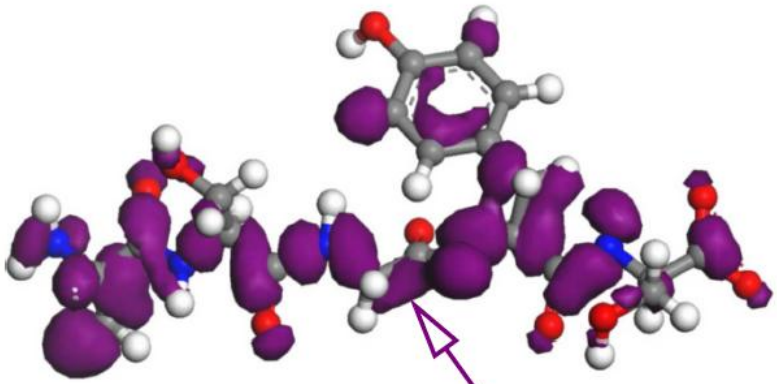
- Use a **small basis set** of localised non-orthogonal functions to express all quantities to obtain sparse matrices (e.g. Hamiltonian matrix, overlap matrix, etc) with modest memory requirements
- Apply your chosen density matrix minimisation algorithm(s), using efficient sparse matrix algebra techniques

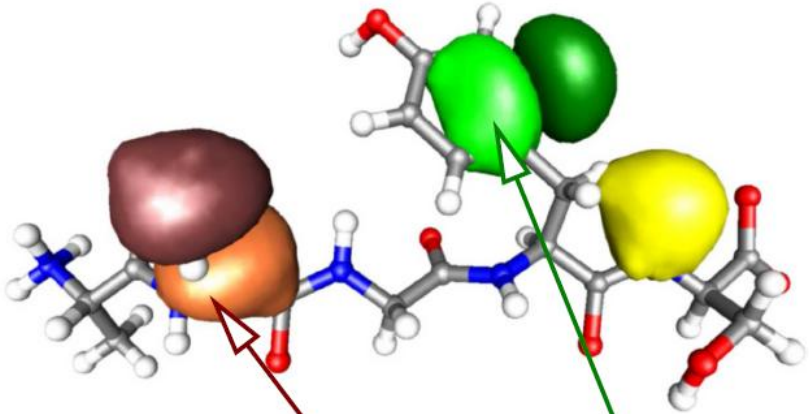
Large basis set accuracy:

- Make sure that the small basis set is "optimal", by determining it variationally, *in situ*
- Therefore the small basis set is not fixed (hence it is not a basis set) but it is expressed in terms of a very detailed, highly accurate **large basis set**

The ONETEP approach

Skylaris, Haynes, Mostofi & Payne, *J. Chem. Phys.* **122**, 084119 (2005)





$$\rho(\mathbf{r},\mathbf{r}') = \sum_n f_n \psi_n(\mathbf{r}) \psi_n^*(\mathbf{r}') = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta^*(\mathbf{r}')$$

Molecular orbitals (MOs)

Non-orthogonal Generalised Wannier Functions (NGWFs)

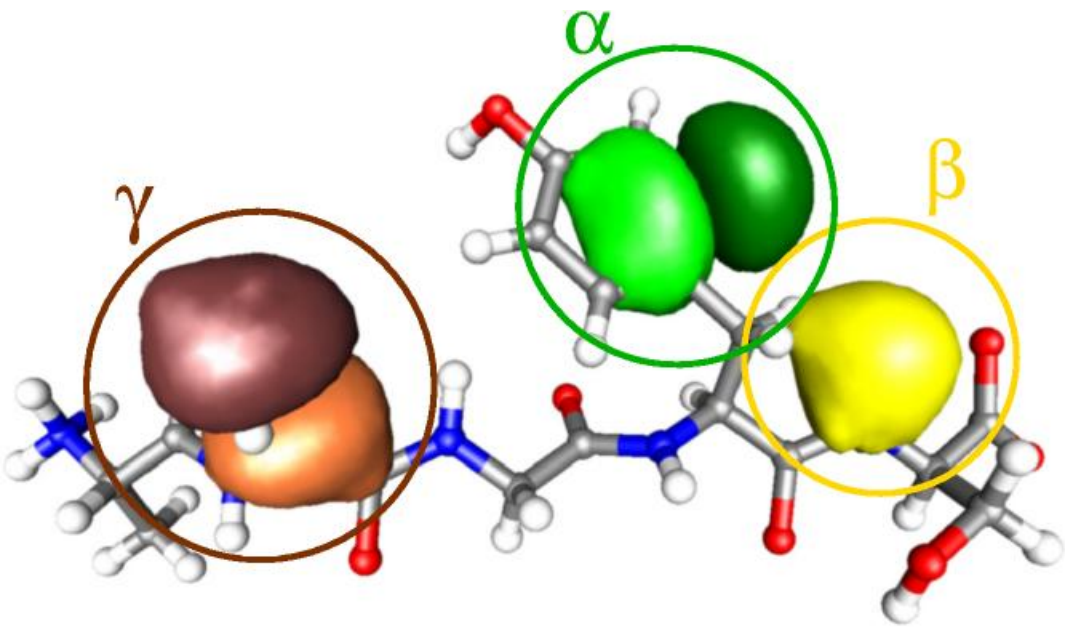
Density kernel

ONETEP is an OBDMM method

- The NGWFs are the “small basis set”
- Each NGWF is expanded in a “large” basis set of psinc functions (this is the actual basis set)

C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, O. Dieguez and M. C. Payne, *Phys. Rev. B* **66**, 035119 (2002).

Density matrix localisation

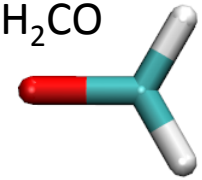


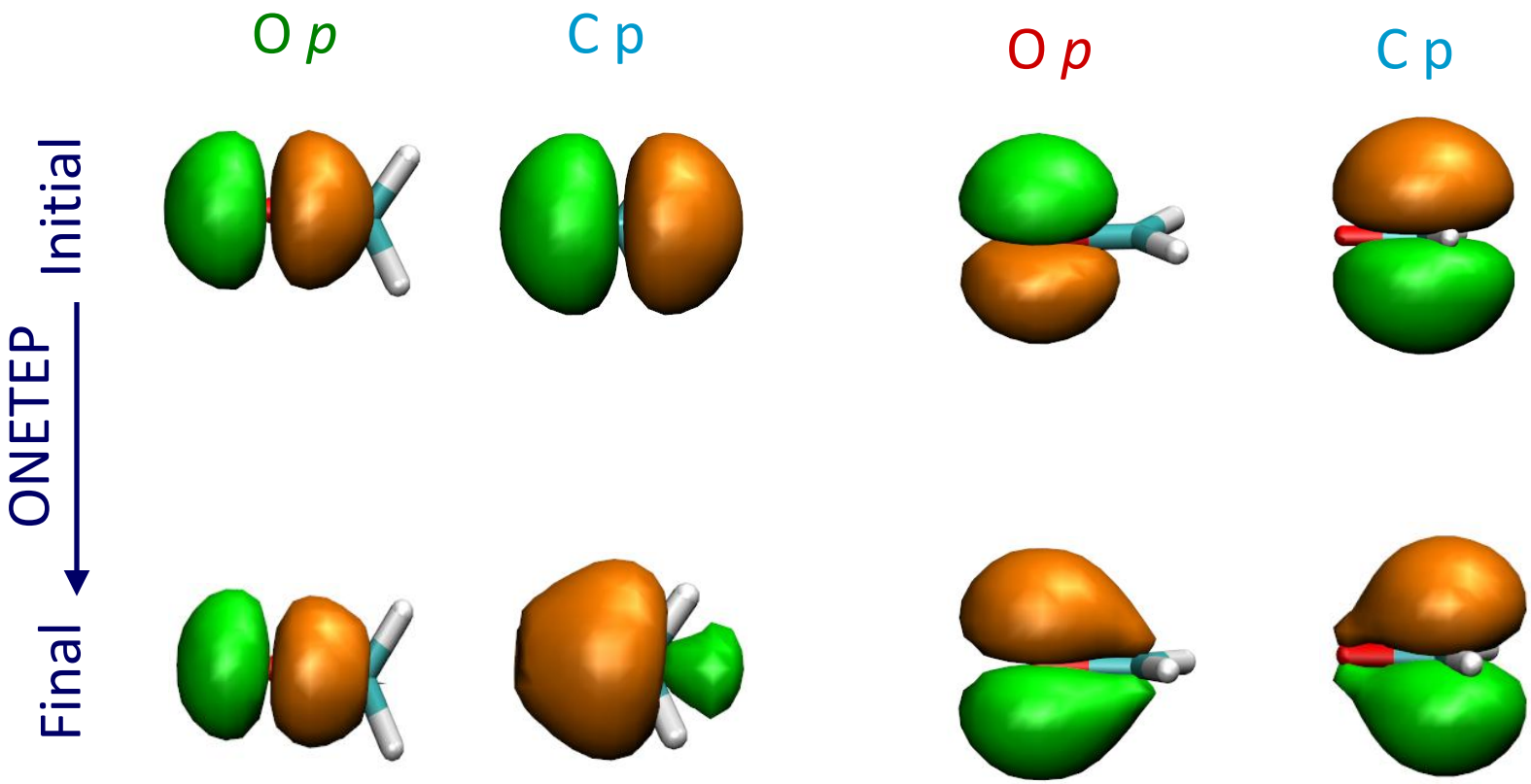
$$\rho(\mathbf{r},\mathbf{r}') = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta(\mathbf{r}')$$

- Impose spatial cut-offs:
  - NGWFs confined to spherical regions
  - Sparse density kernel $\boldsymbol{K}$ by truncation

NGWF optimisation

formaldehyde, H_2CO

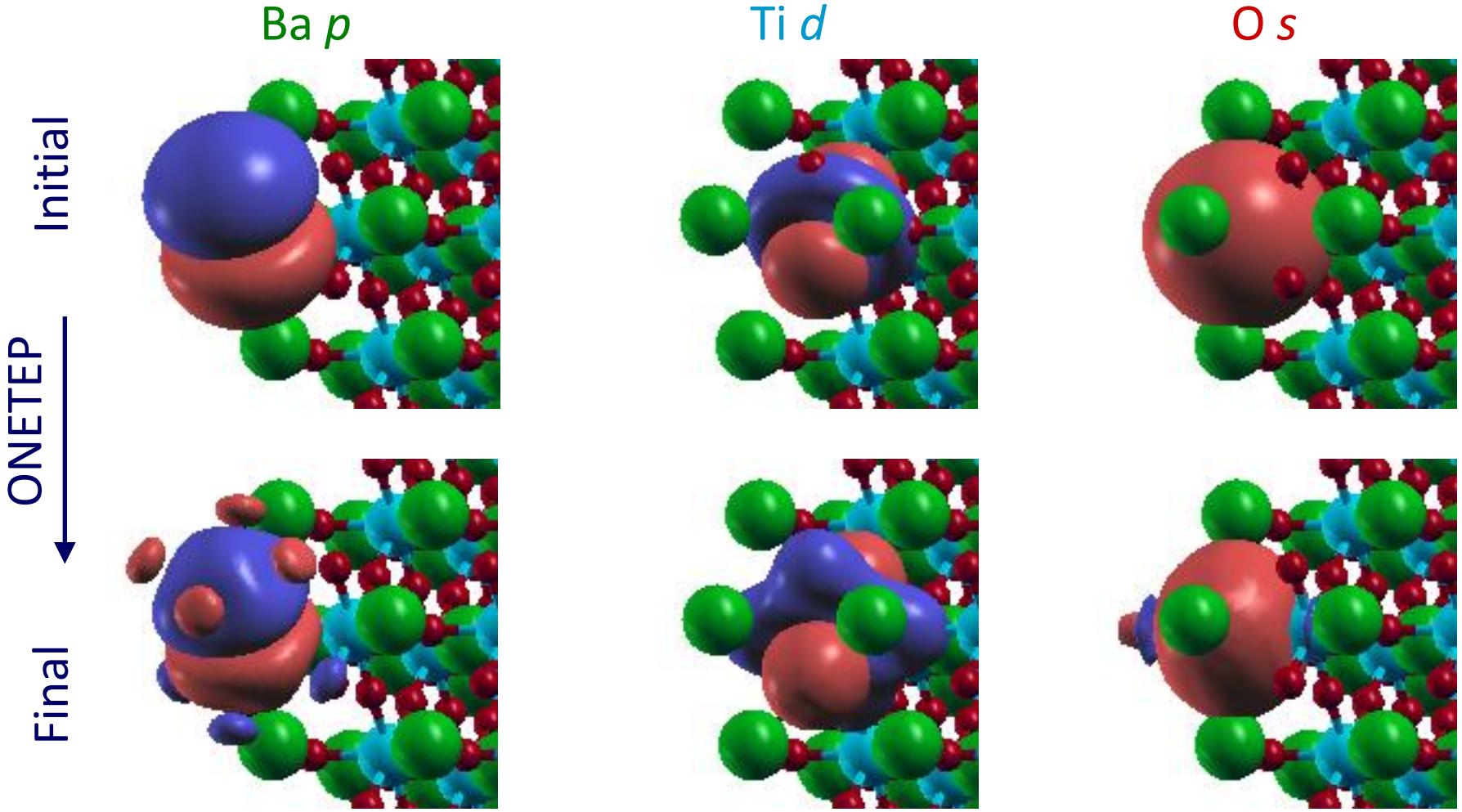




On-site rotation from Foster & Weinhold, *J. Am. Chem. Soc.* **102**, 7211 (1980)

NGWF optimisation

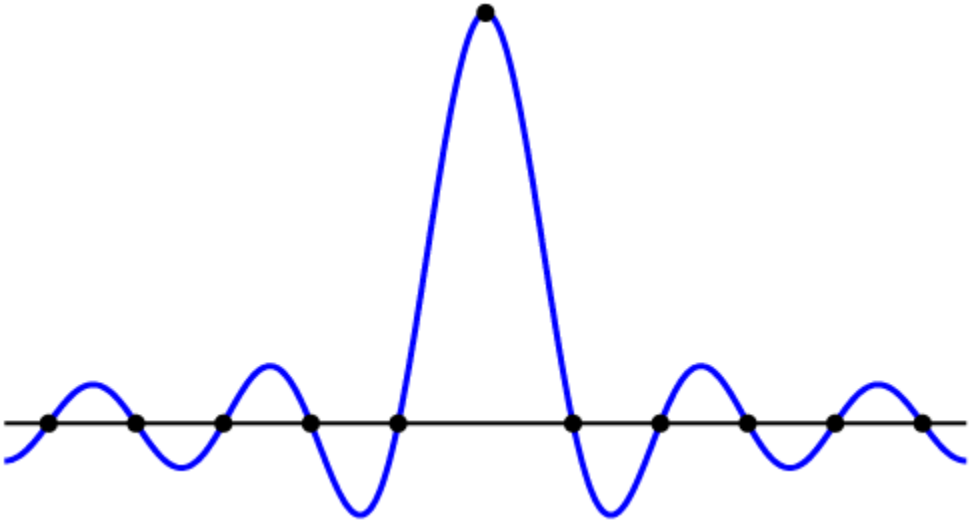
$BaTiO_3$

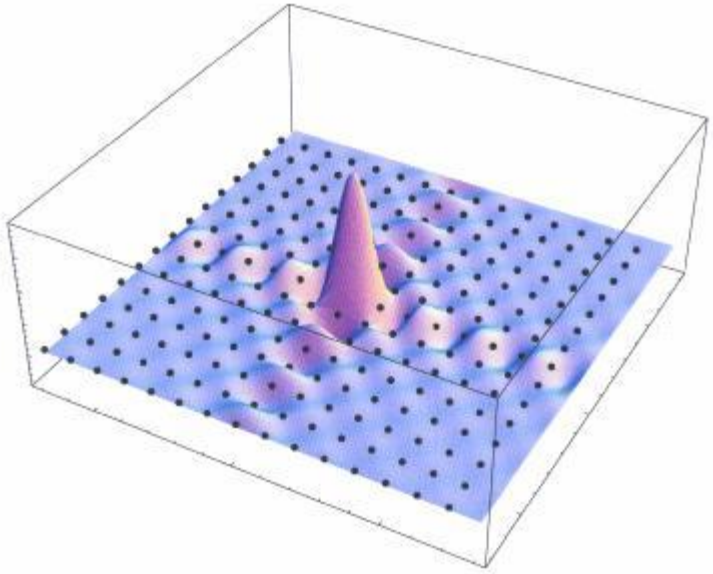


On-site rotation from Foster & Weinhold, *J. Am. Chem. Soc.* **102**, 7211 (1980)

Basis set: psinc functions

- "Periodic Cardinal Sine" or Lagrange-mesh functions:

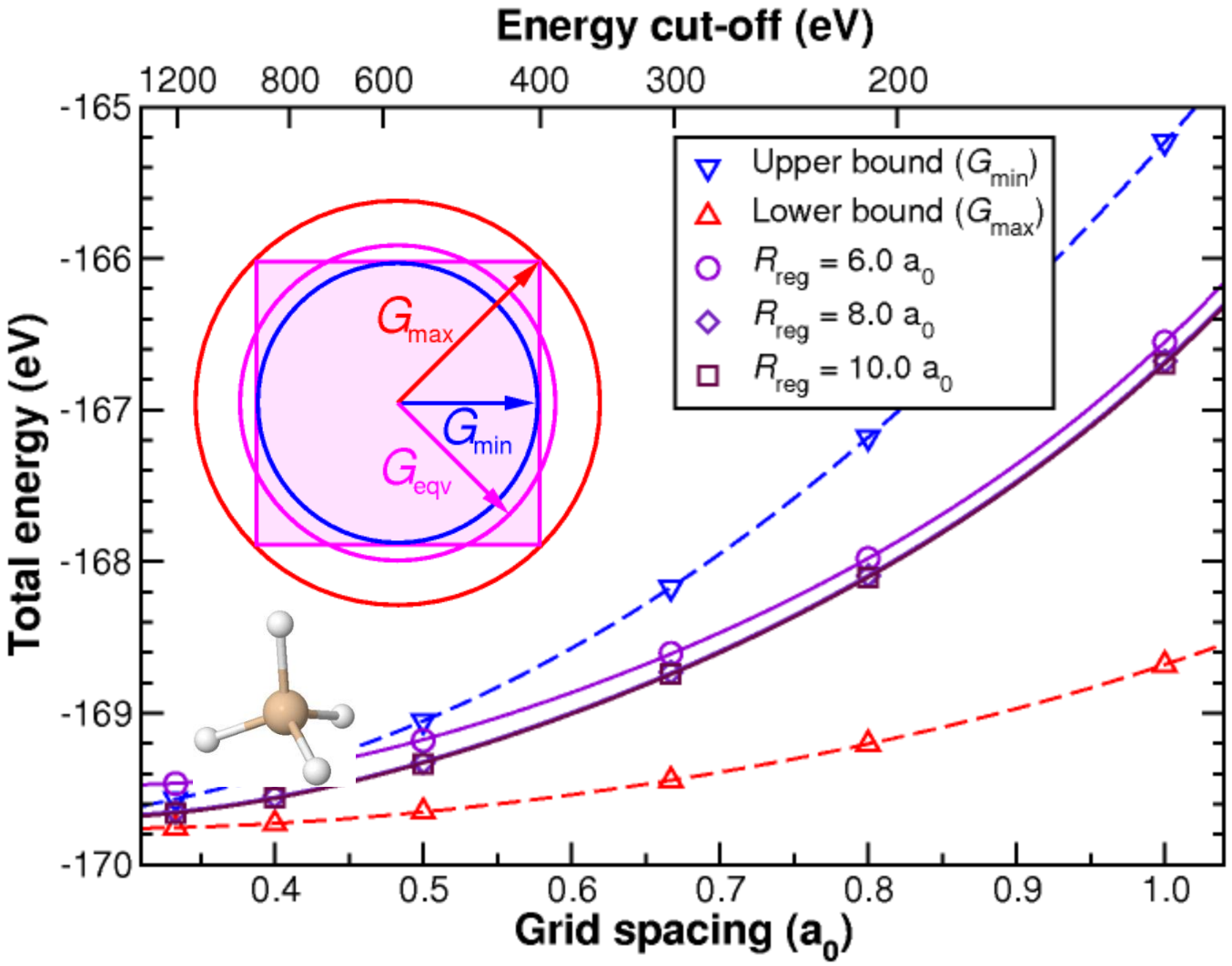




- Real linear combinations of plane waves
- Localised
- Orthogonal

Mostofi, Haynes, Skylaris & Payne, *J. Chem. Phys.* **119**, 8842 (2003)

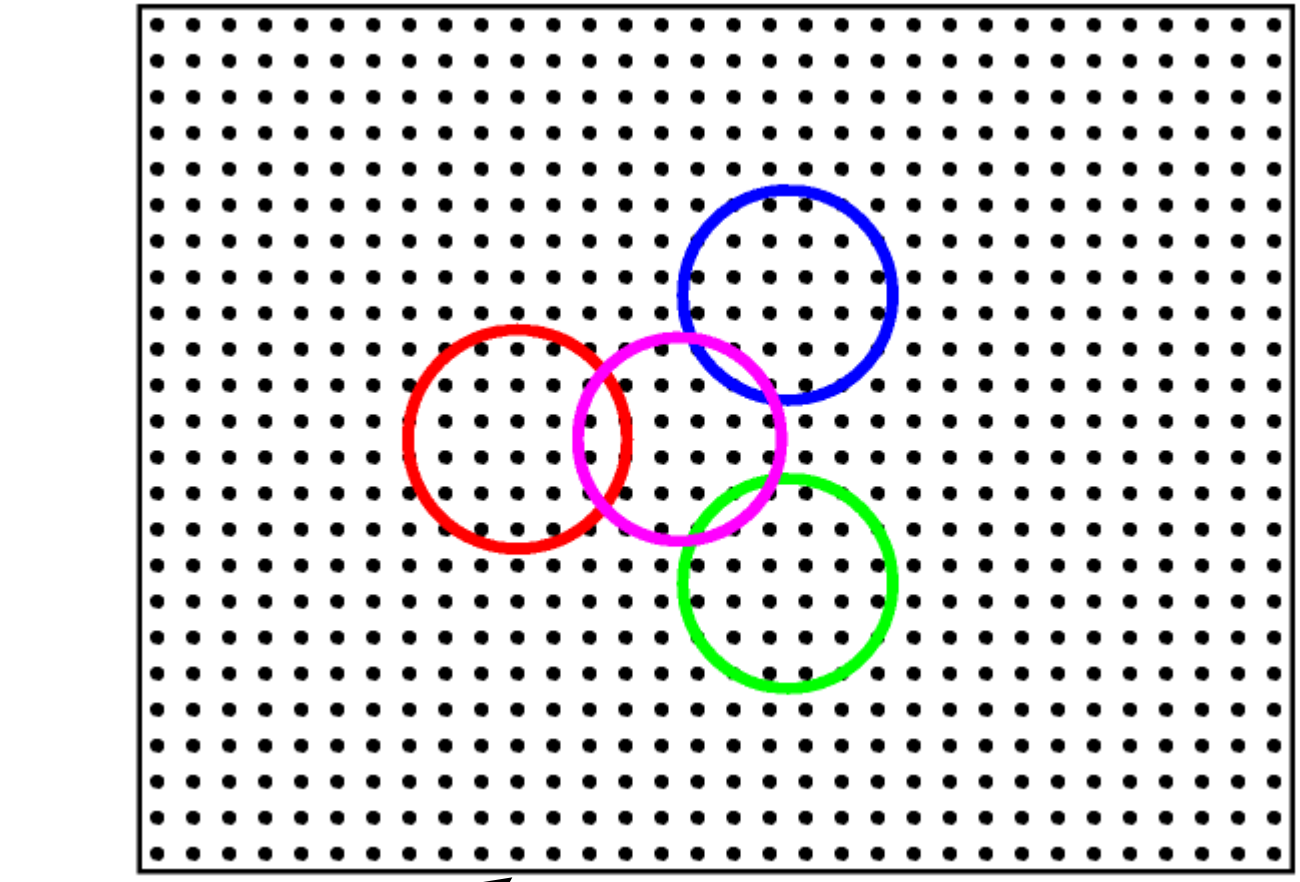
Psinc basis energy cut-off



Basis set variational approaches:

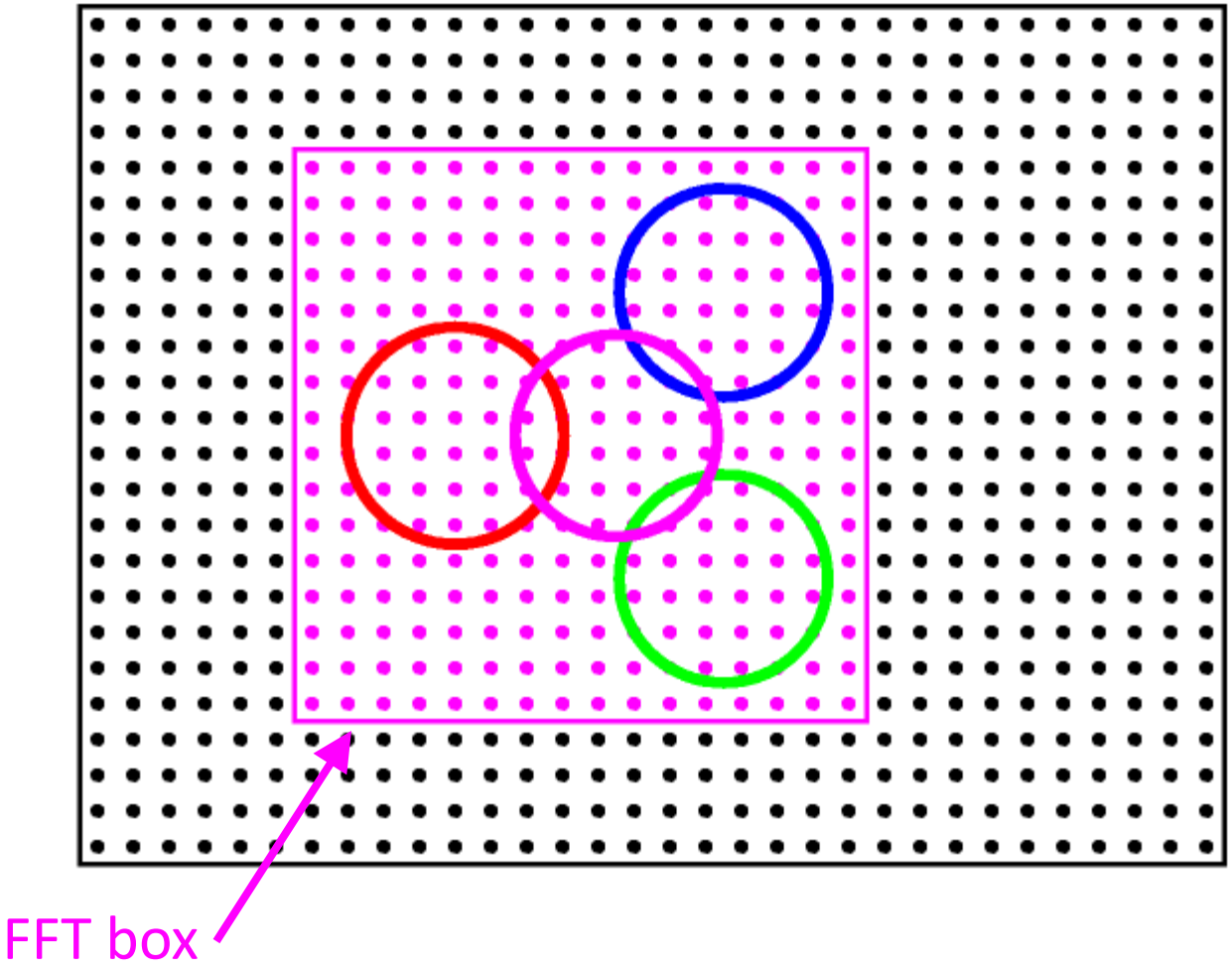
C.-K. Skylaris, O. Dieguez, P. Haynes and M. C. Payne, *Phys. Rev. B* **66**, 073103 (2002).

FFT box technique



simulation cell

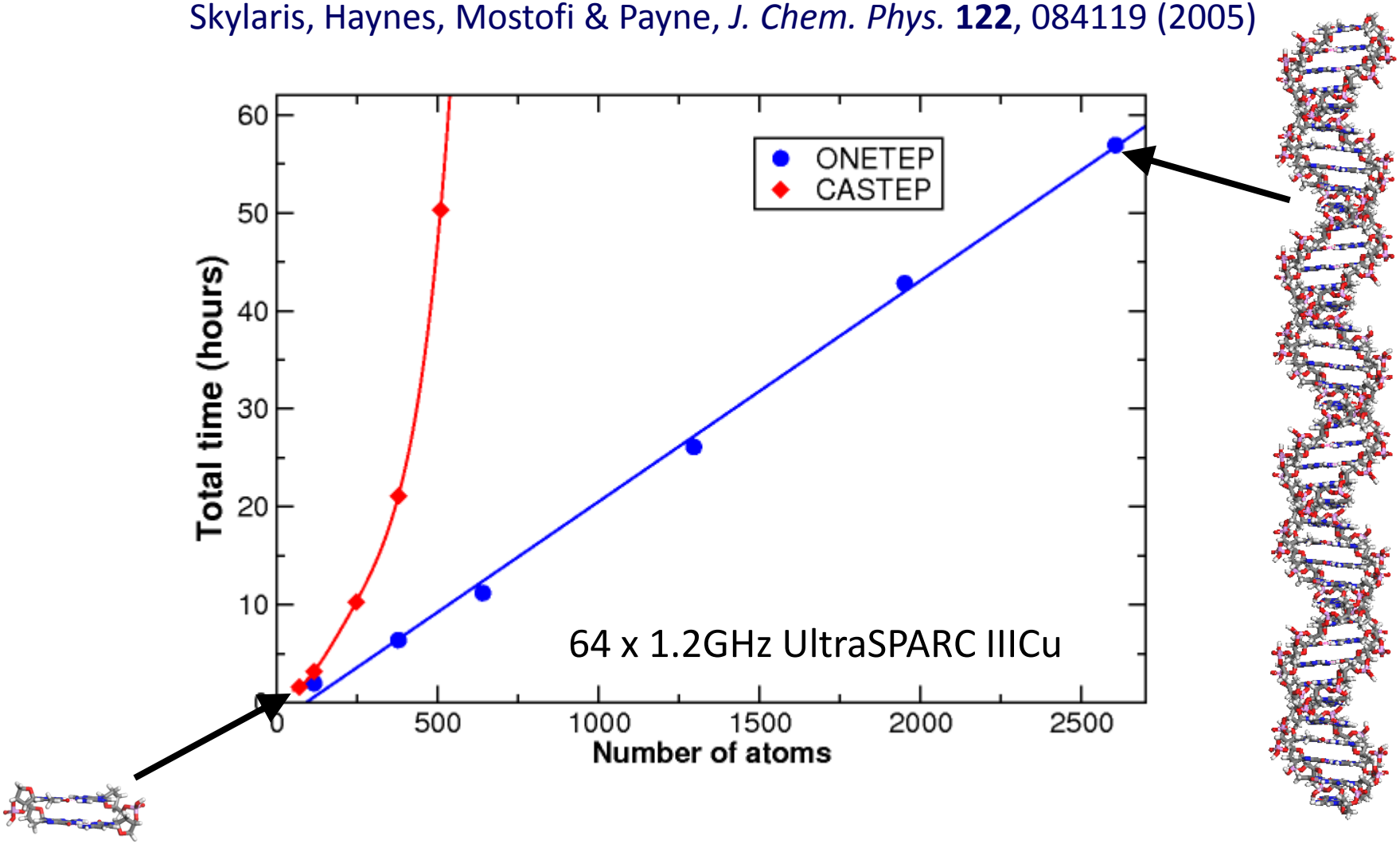
FFT box technique



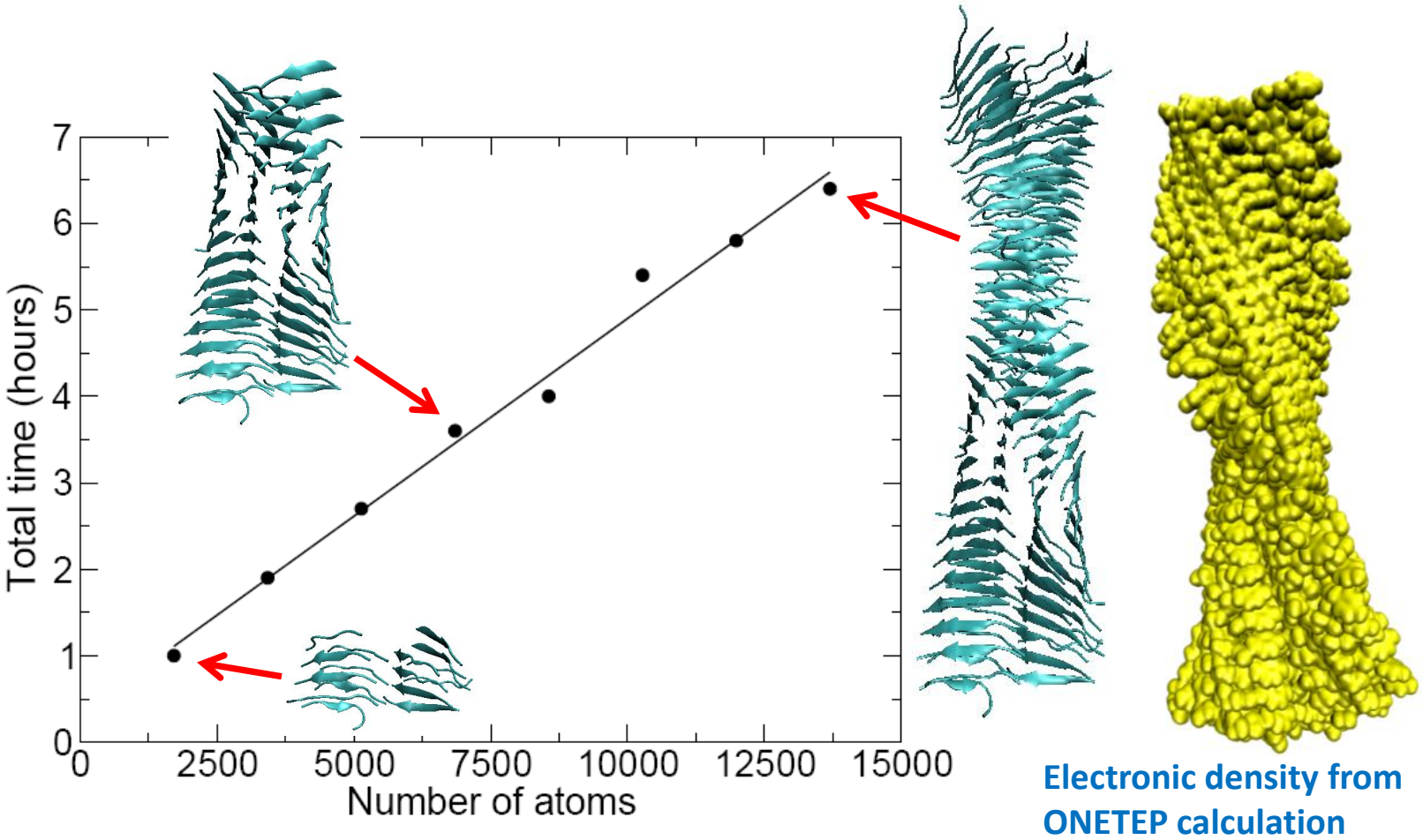
C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, C. J. Pickard & M. C. Payne, *Comp. Phys. Comm.* **140**, 315 (2001)
A. A. Mostofi, C.-K. Skylaris, P. D. Haynes & M. C. Payne, *Comp. Phys. Comm.* **147**, 788 (2002)

Linear scaling: DNA

Skylaris, Haynes, Mostofi & Payne, *J. Chem. Phys.* **122**, 084119 (2005)

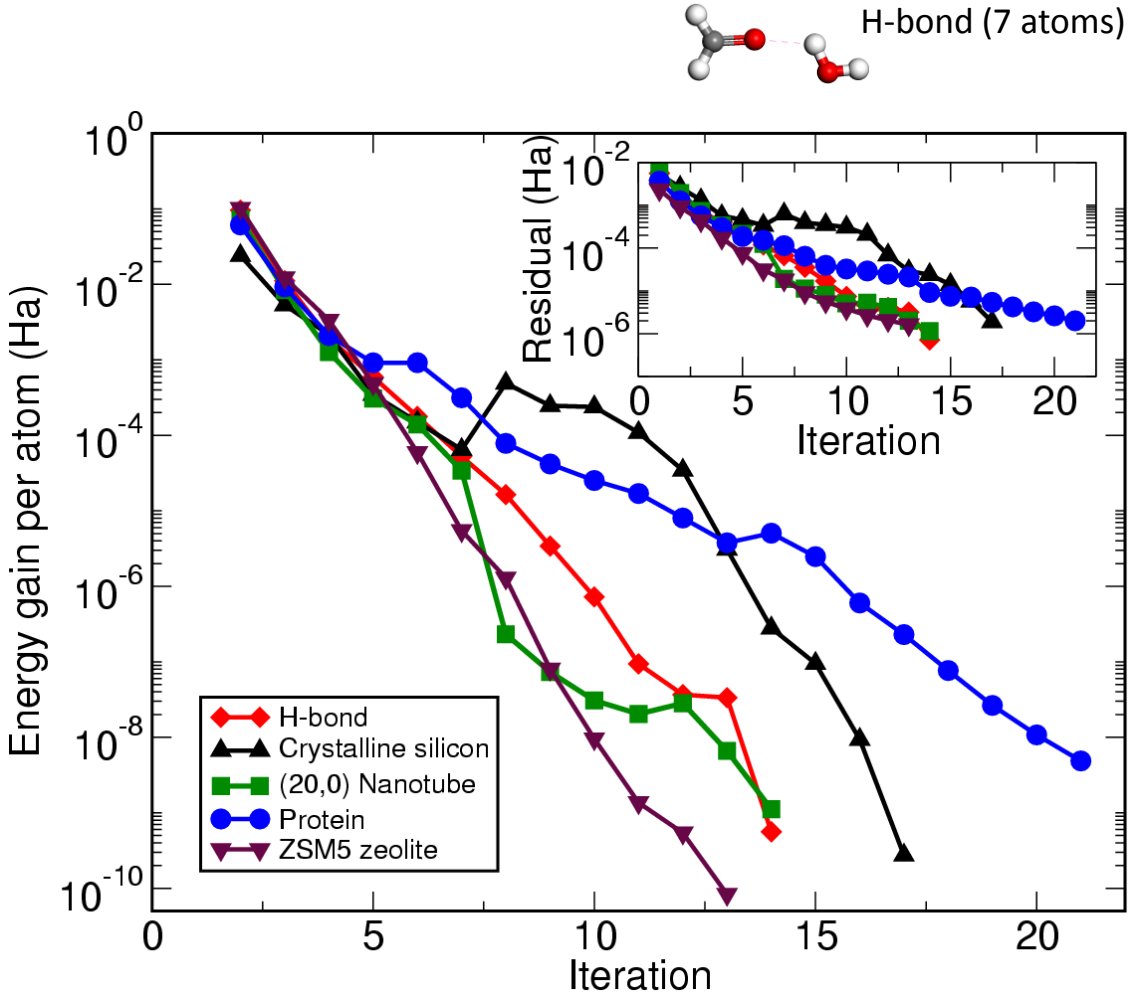


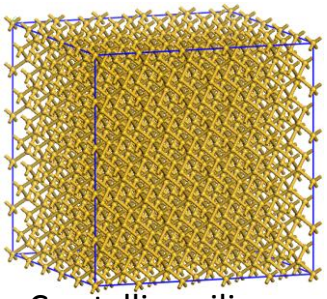
Linear-scaling: Amyloid fibrils



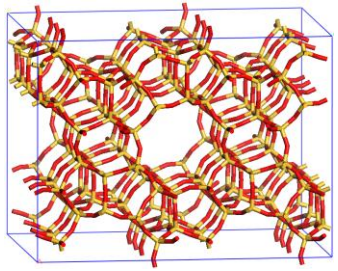
Structures of the amyloid fibril kindly provided by the authors of
J. T. Berryman, S. E. Radford and S. A. Harris, Biophysical Journal, ***97*** *1 (2009)*

True linear scaling

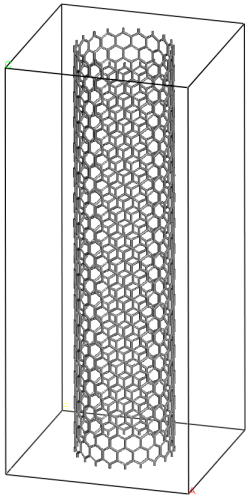




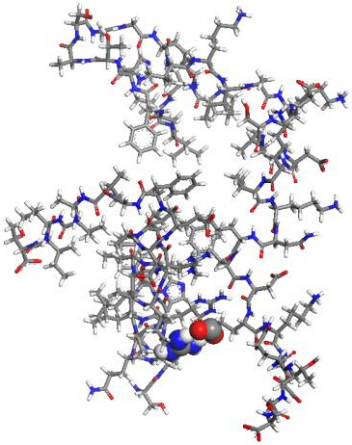
Crystalline silicon (1000 atoms)



ZSM5 zeolite (576 atoms)



(20, 0) Nanotube (1280 atoms)

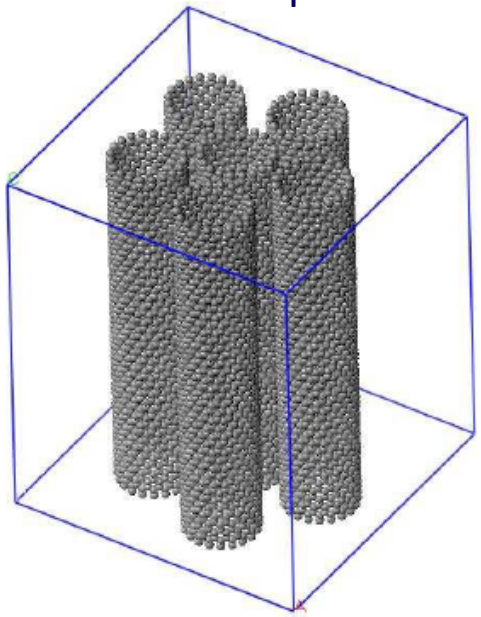


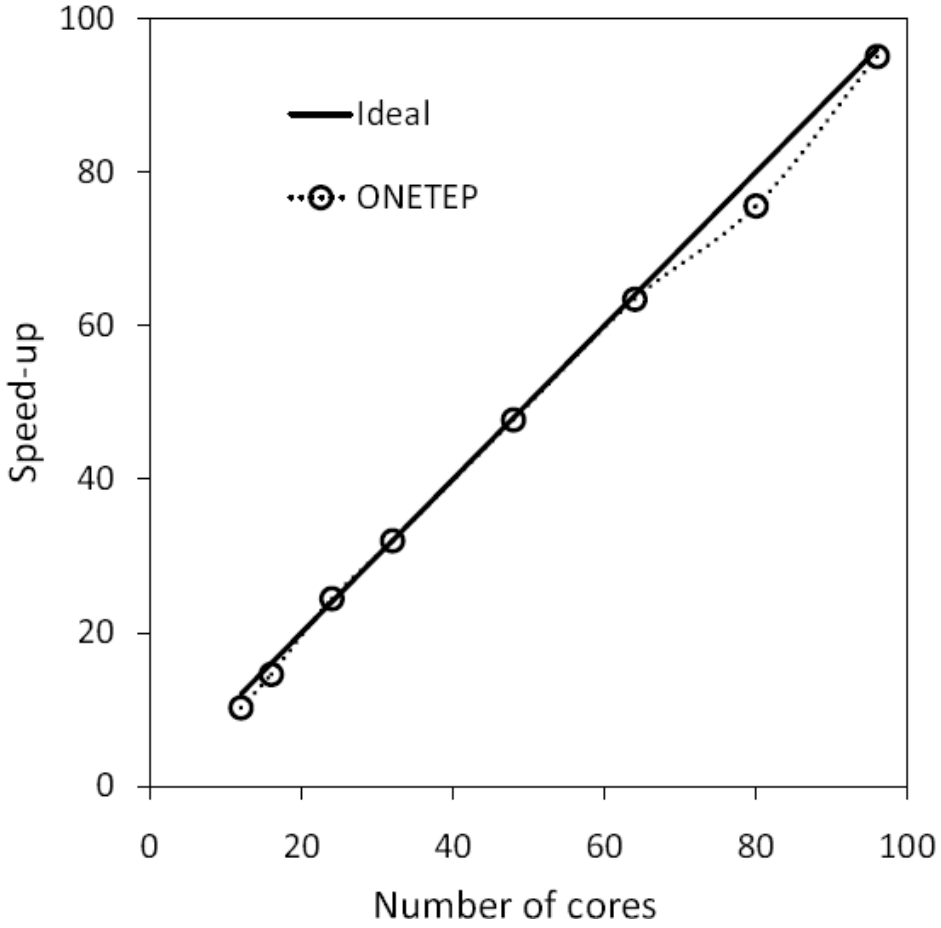
Protein (988 atoms)

Skylaris, Haynes, Mostofi & Payne, *J. Phys.: Condens. Matter* **17**, 5757 (2005)

Parallel scaling: carbon nanotubes

- 9600 atoms
- Commodity cluster of 24 dual socket dual core Opteron nodes





Skylaris, Haynes, Mostofi & Payne, *Phys. Stat. Sol. (b)* **243**, 973 (2006)
Hine, Haynes, Mostofi, Skylaris & Payne, *Comput. Phys. Comm.* **180**, 1041 (2009)

Functionality

- LDA (Ceperley-Alder-Perdew-Zunger, Vosko-Wilk-Nusair) and GGAs (Perdew-Wang ’91, Perdew-Burke-Ernzerhof, revPBE, RPBE, BLYP, XLYP)
- Spin polarisation
- Forces
- Geometry optimisation
- Electronic structure analysis
- Visualisation
- Modified Coulomb interactions (Cut-off Coulomb and Martyna-Tuckerman approaches)

Compiling ONETEP

Simple multi-platform build system, needs:

- Fortran 95 compiler
- BLAS and LAPACK numerical libraries
- FFT library: vendor-supplied or FFTw
  - www.fftw.org
- MPI library for parallel version
  - www.lam-mpi.org

Running ONETEP

- Parallel computer
  - Minimum 2 GB per processor (core)
  - Typically distribute 10-100 atoms per processor
  - Cross-over >100 atoms

- Prepare input file: free format
  - Documentation at www.onetep.org

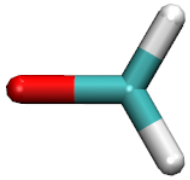
- Supply pseudopotential files (**`.recpot`** format)

Input file

- Keywords of different types:
  - Integer
  - Boolean
  - String
  - Real
  - Physical (real + unit)
  - Block data e.g. atomic positions, delimited by **`%block`** and **`%endblock`**
- Atomic units by default (hartree and bohr)
- Beware older keywords e.g. **`kernel_cutoff`**

Example input file: formaldehyde

```
! Example input file for the ONETEP program
! Formaldehyde molecule

cutoff_energy 600 eV

%block   lattice_cart
  48.00     0.00     0.00
   0.00    48.00     0.00
   0.00     0.00    48.00
%endblock  lattice_cart

%block  positions_abs
O     24.887507   23.896975   22.647313
C     27.731659   23.667449   22.643306
H     28.655157   21.721170   22.637547
H     28.955467   25.440371   22.646039
%endblock  positions_abs

%block species
O    O    8    4    6.5
C    C    6    4    6.5
H    H    1    1    6.5
%endblock species

%block species_pot
O    oxygen.recpot
C    carbon.recpot
H    hydrogen.recpot
%endblock species_pot
```



ONETEP calculation outline

- Initialisation phase:
  - Construct initial NGWFs (STOs or PAOs)
  - Construct initial charge density (atomic superposition) and effective potential
  - Construct initial Hamiltonian
  - Obtain initial (non-self-consistent) density kernel using canonical purification
  - Refine initial density kernel (self-consistently) using penalty functional

ONETEP calculation outline continued

- Main optimisation phase:
  - Combination of nested self-consistent loops
  - Outer loop optimises the NGWFs (density kernel fixed)
  - Inner loop optimises the density kernel (NGWFs fixed) using DMM approaches
- Final analysis phase:
  - Calculate forces
  - Write out potentials, densities, NGWFs for plotting
  - Mulliken population analysis
  - Diagonalisation yields wave functions, DOS etc.

Example output file: formaldehyde

View h2co.out

More information

- www.onetep.org

- Scientific highlight of the month:
  - Ψ_k Newsletter **72**, December 2005
  - http://psi-k.dl.ac.uk/

- *J. Chem. Phys.* **122**, 084119 (2005)

THE ROYAL SOCIETY
CELEBRATING 350 YEARS

EPSRC