Linear-Scaling Quantum Chemistry:

Coupled Cluster

Divide-Expand-Consolidate (DEC)

Scheme

Dmytro Bykov

OLCF, ORNL

May 18, 2018

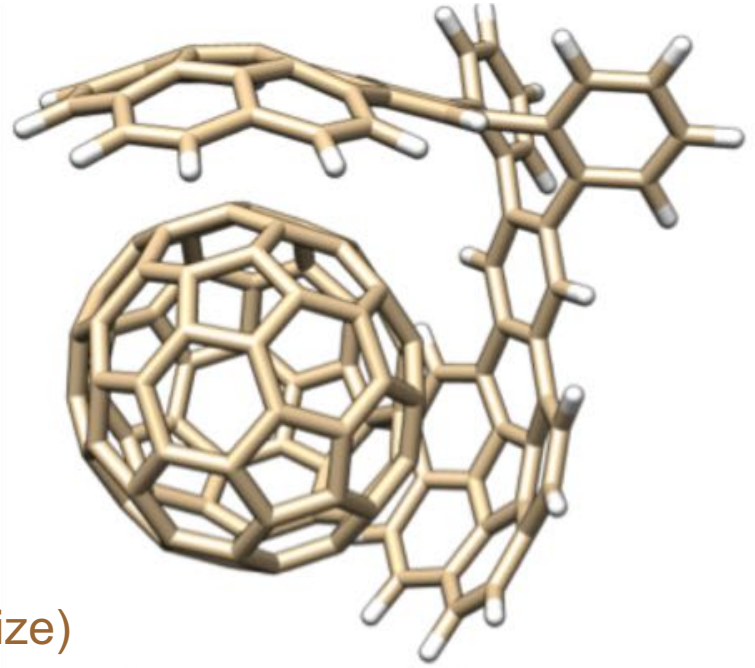
ORNL is managed by UT-Battelle
for the US Department of Energy

OAK RIDGE National Laboratory

Motivation

Support, Explain and Guide the Experiment

a) DFT is a “working horse” in ab initio methods
b) Coupled cluster theory is the method of choice
c) Hierarchies of coupled models with improved accuracy



MP2 : Second order Møller Plesset theory , N^5 scaling
CCSD : Coupled cluster singles and doubles model, N^6 scaling
CCSD(T) : CCSD with perturbative triples corrections , N^7 scaling

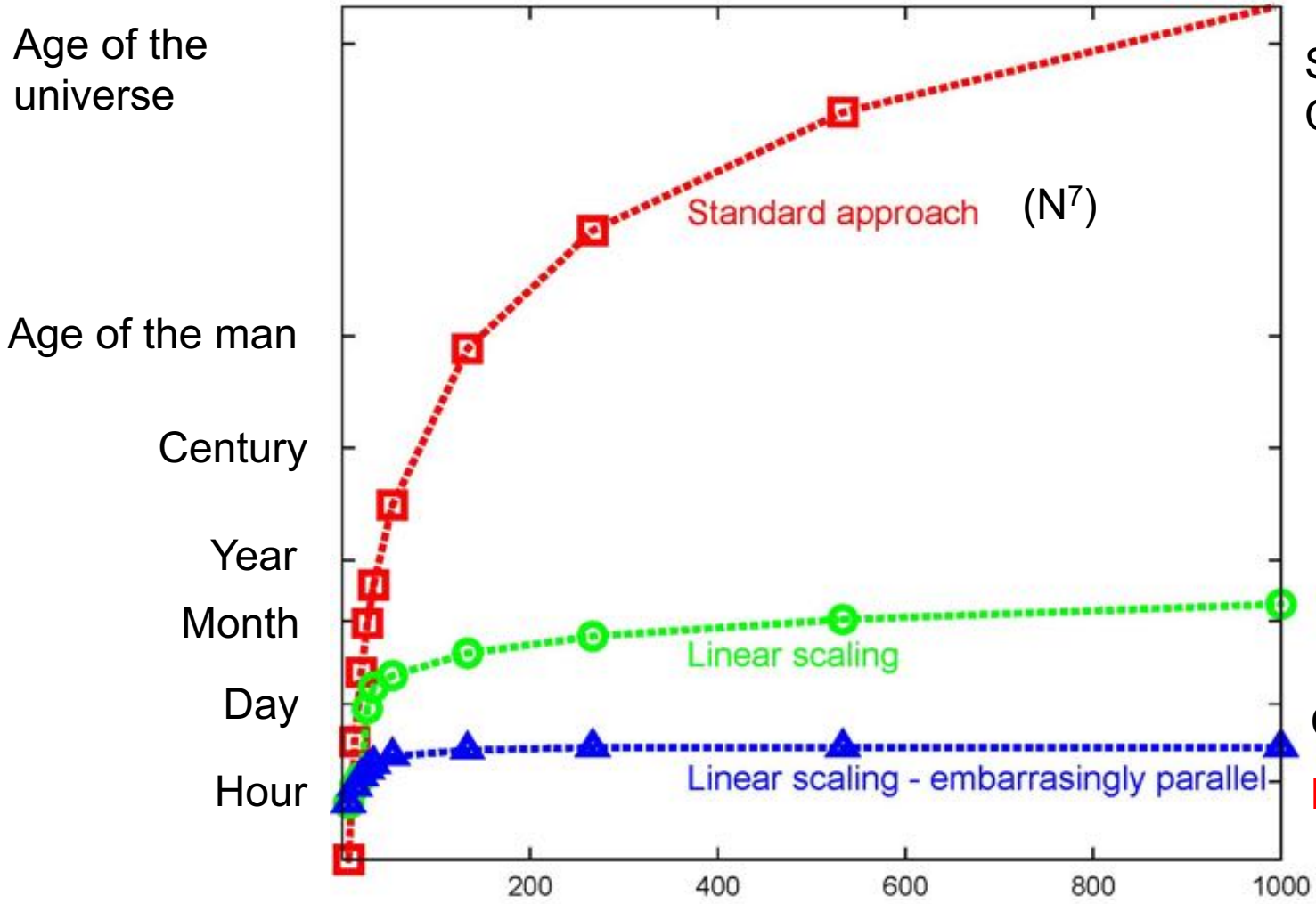
(N denote system size)

CCSD(T) is the gold standard of quantum chemistry

Many molecular properties are described to experimental accuracy or better e.g. equilibrium geometries and reaction enthalpies

Problem: The scaling in standard implementations

Scaling in Coupled Cluster theory CCSD(T)

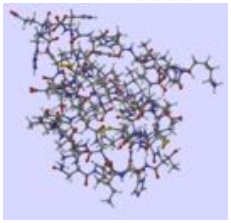


Standard Coupled Cluster

Canonical HF orbitals

DEC Coupled Cluster

Local HF orbitals

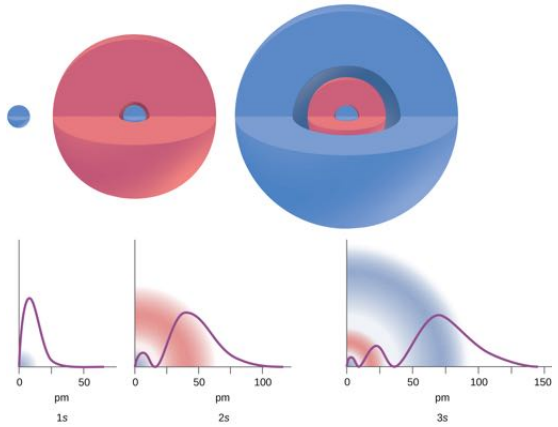


Insuline

Coupled cluster wave function calculation

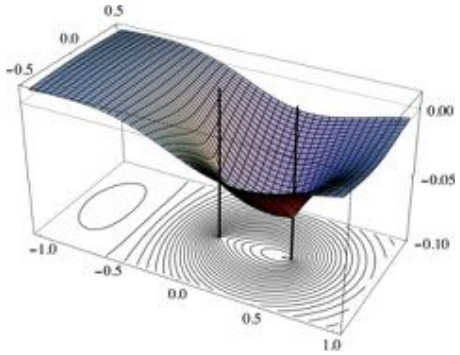
Hartree-Fock calculation determines the reference state

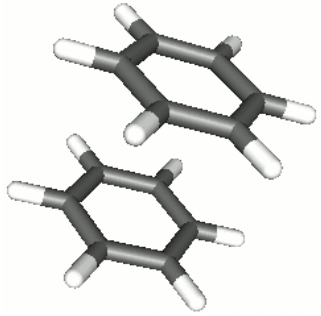
- Electrons move in an averaged field of the others
- Long range potential described to high accuracy
- **Give local orbitals for both occupied and virtual space**



Coupled cluster calculation describes local electron correlation effects

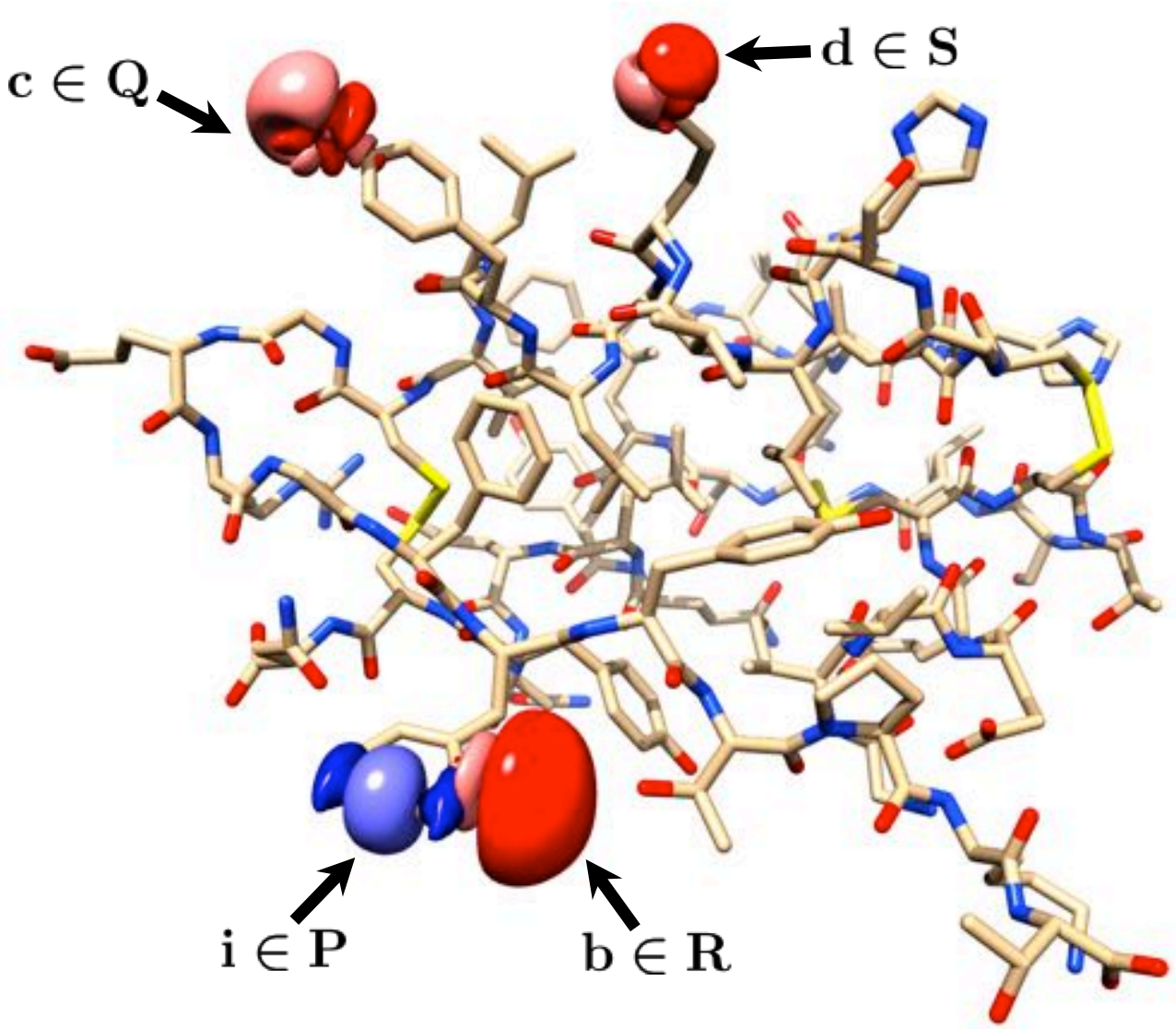
- Coulomb hole, Short range
- Dispersion forces, R^{-6}





Local phenomena described in **local basis** give linear scaling

Determine local Hartree-Fock orbitals



Occupied orbitals:
i,j,k,l

Virtual orbitals:
a,b,c,d

Atomic sites:
P,Q,R,S

Assign orbitals to atomic sites

INSULIN MOLECULE (HYDROGENS OMITTED)

The Divide-Expand-Consolidate DEC coupled cluster (CC) method

Method

- ***Local occupied*** and ***virtual*** HF orbitals must be determined
- Assign ***local occupied*** and ***virtual*** HF orbitals to atomic sites, *P,Q*,..
- Partition correlation energy into atomic fragment E_P and pair ΔE_{PQ} energies
- Evaluate E_P and ΔE_{PQ} from small orbital fragment spaces adjusted to give the energies to a predefined fragment optimization threshold (FOT)

Features

- Adjustable orbital spaces give error control for energies and amplitudes
- Black box, linear scaling, and embarrassingly parallel

Precision of Coupled Cluster calculation

Standard Coupled Cluster

- Residual norm (***R***) of amplitude equation

DEC Coupled Cluster (new strategy which exploit locality efficiently)

- Fragment optimization threshold (***FOT***) for atomic fragment energies

Standard and DEC Coupled Cluster on par with respect to precision

- Single parameter (***R*** or ***FOT***) define the precision of energy, density, ...

Correlation energy MP2 and CCSD

$$E_{corr} = \sum_{ijab} (t_{ij}^{ab} + t_i^a t_j^b)(2g_{iajb} - g_{ibja})$$

Quadratic scaling

$$E_{corr} = \sum_{P} E_P + \sum_{P>Q} \Delta E_{PQ}$$

Coulomb hole Dispersion energy

DEC Strategy

Assign local orbitals to atomic sites P,Q,..

Replace summation over two occupied i,j (virtual a,b) orbitals with summations over sites P and pair sites P,Q and summations over orbitals belonging to these sites

Occupied atomic fragment energy

$$E_P = \sum_{\substack{ij \in P \\ ab}} \left(t_{ij}^{ab} + t_i^a t_j^b\right)\left(2g_{iajb} - g_{ibja}\right)$$

Occupied pair interaction energy

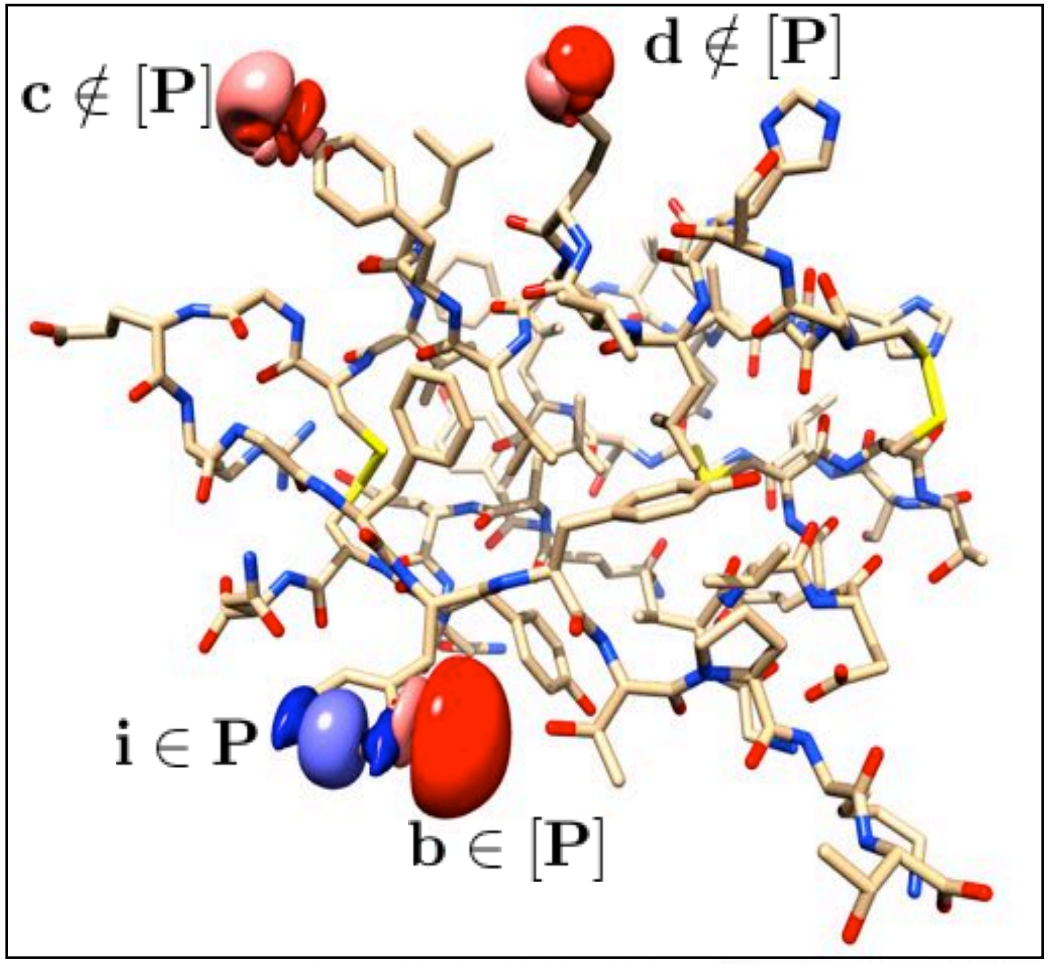
$$\Delta E_{PQ} = \sum_{\substack{i \in P, j \in Q \\ ab}} \left(t_{ij}^{ab} + t_i^a t_j^b\right)\left(2g_{iajb} - g_{ibja}\right) + \quad P \leftrightarrow Q \text{ term}$$

No approximations have so far been made, only a reorganization

Atomic fragment energy E_P

$$E_P = \sum_{ij \in P \; ab \in [P]} \left(t_{ij}^{ab} + t_i^a t_j^b \right) \left(2 g_{iajb} - g_{ibja} \right)$$

Determine [P] in a black box manner such that errors in E_P is smaller than **Fragment Optimization Threshold (*FOT*)**



CCSD and MP2

Charge distributions in integrals determine distance decay from site P

$$g_{iajb} \quad ; \; i,j \in P \quad , \; a,b \in [P]$$

Occupied space partitioning of E_{corr}

Occupied atomic fragment energy

$$E_P = \sum_{\substack{i \in P, j \in P \\ ab \in [P]}} \left(t_{ij}^{ab} + t_i^a t_j^b \right) \left(2 g_{iajb} - g_{ibja} \right)$$

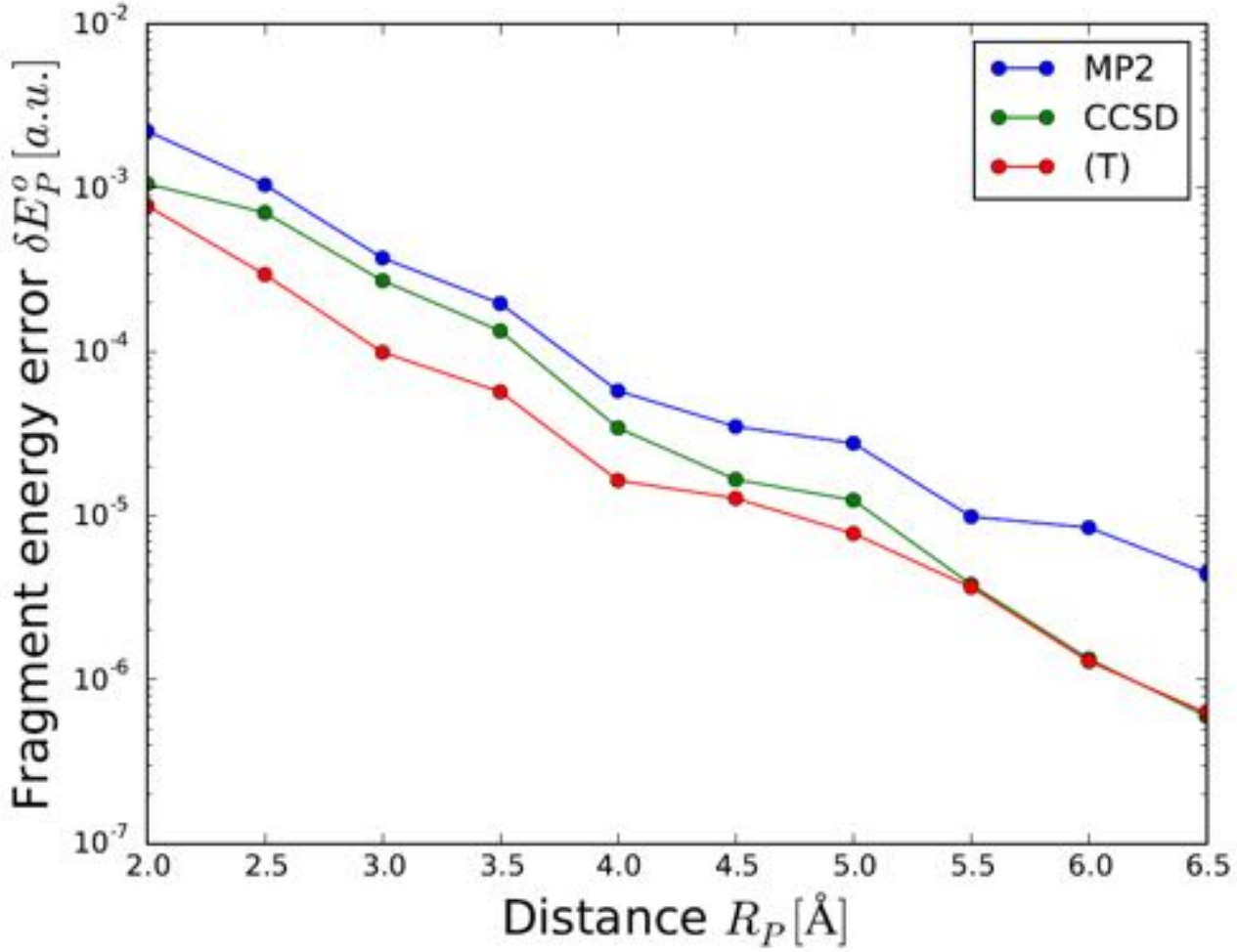
(atomic fragment orbital space)

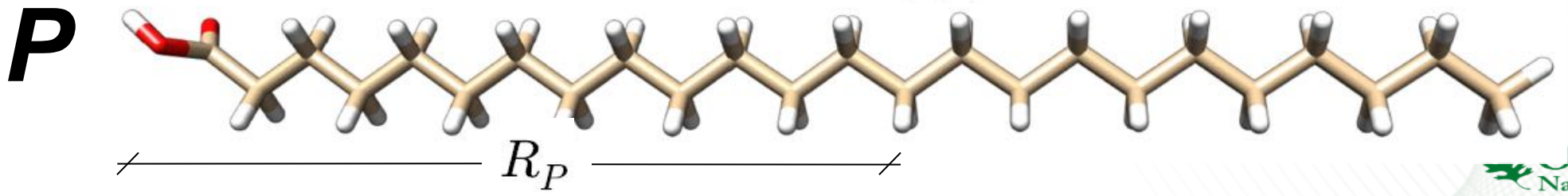
Occupied atomic pair interaction energy

$$\Delta E_{PQ} = \sum_{\substack{i \in P, j \in Q \\ ab \in [P] \cup [Q]}} \left(t_{ij}^{ab} + t_i^a t_j^b \right) \left(2 g_{iajb} - g_{ibja} \right) + \quad P \leftrightarrow Q \ term$$

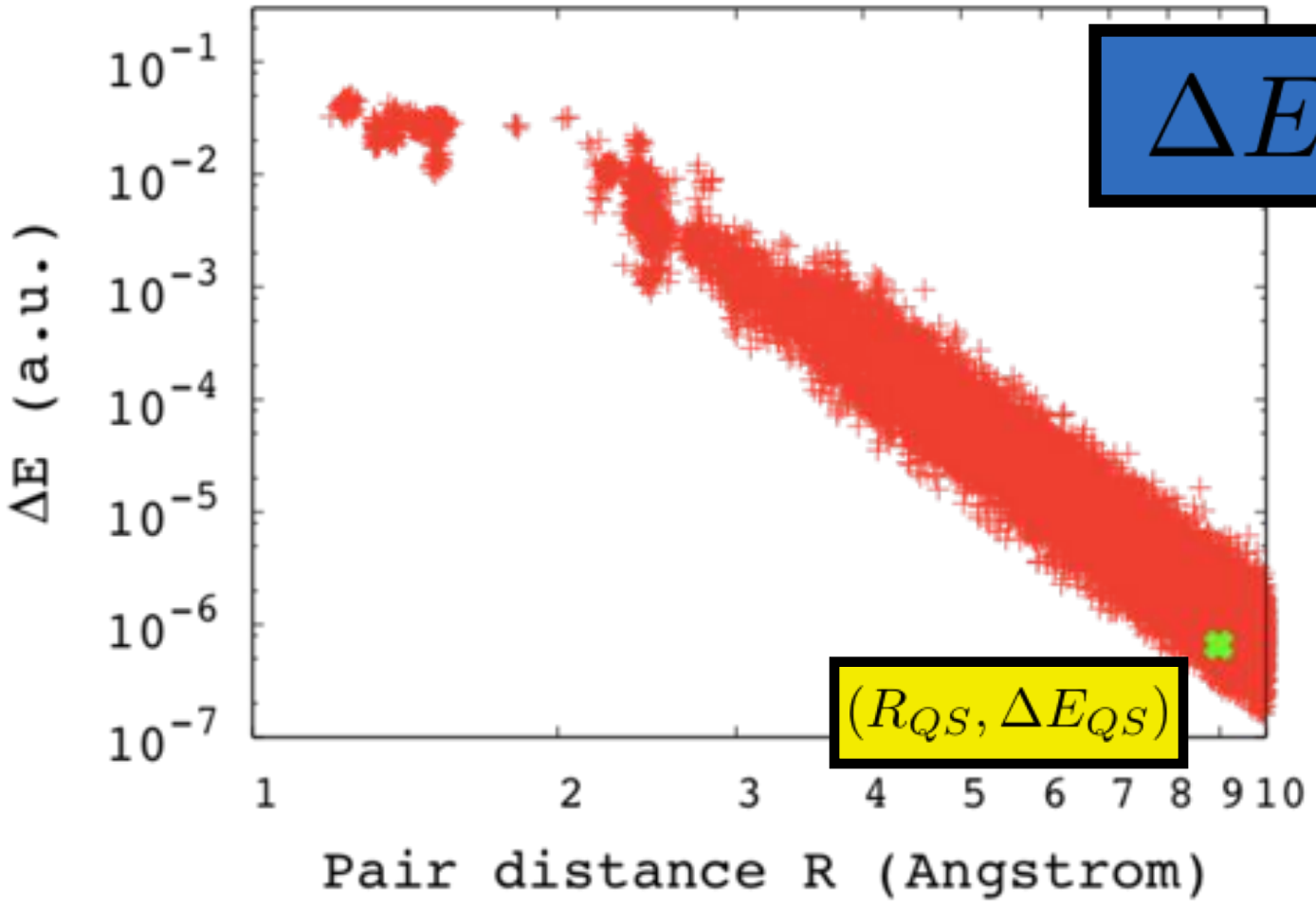
(union of atomic fragment orbital spaces)

Distance decay of E_P

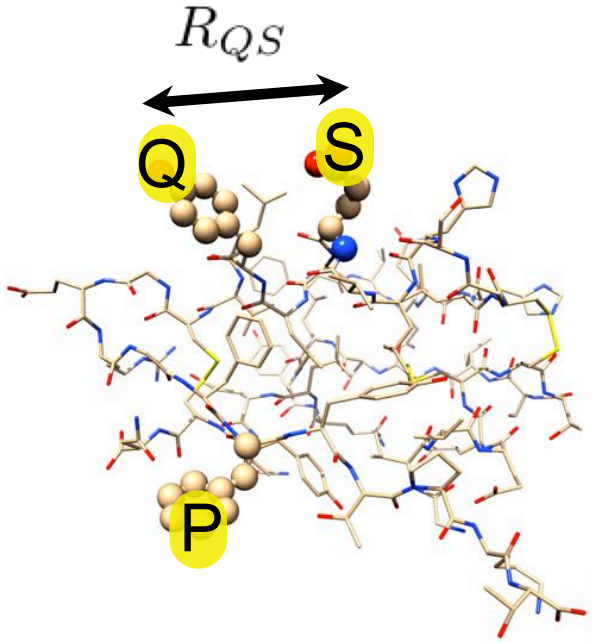




Insulin - pair interaction energies ΔE_{PQ}



$$\Delta E \propto R^{-6}$$



Pairs separated by more than ~10 Å can be neglected without affecting the precision of the calculation

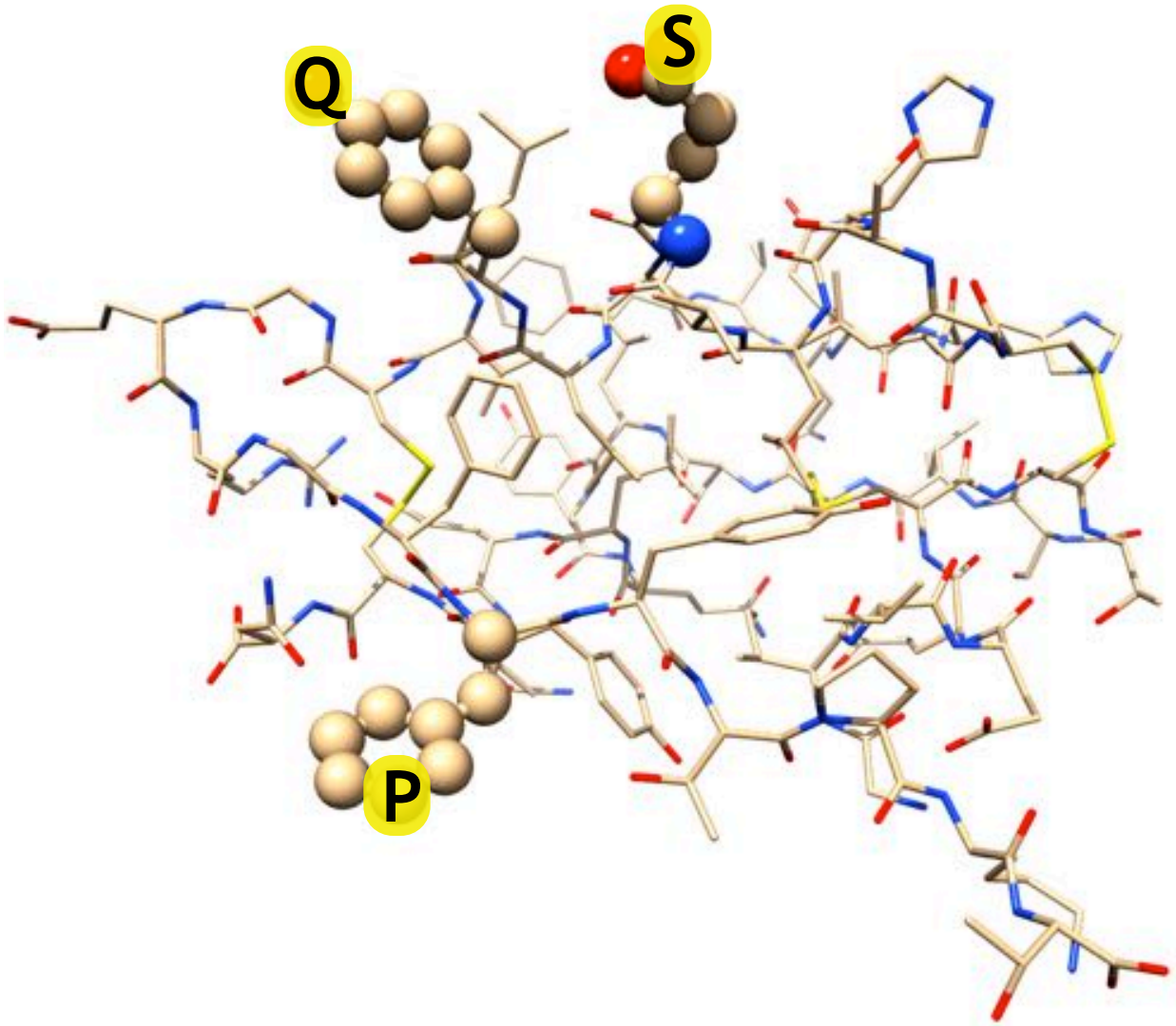
Linear scaling algorithm

OAK RIDGE National Laboratory

Insulin - three examples of atomic fragments

$$E_{\mathrm{corr}} = \sum_{P} E_P + \sum_{P>Q} \Delta E_{PQ}$$

Quadratic scaling

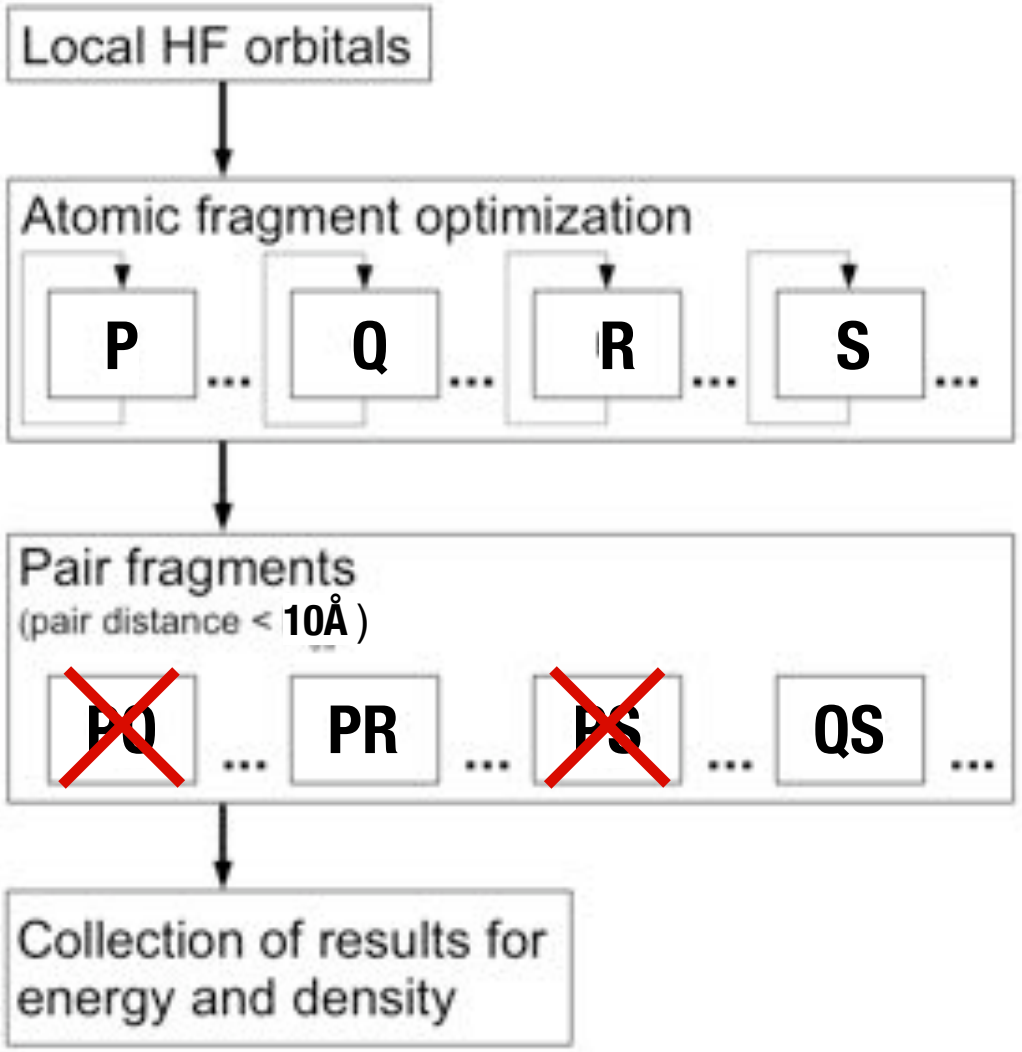


E_P determined to FOT precision

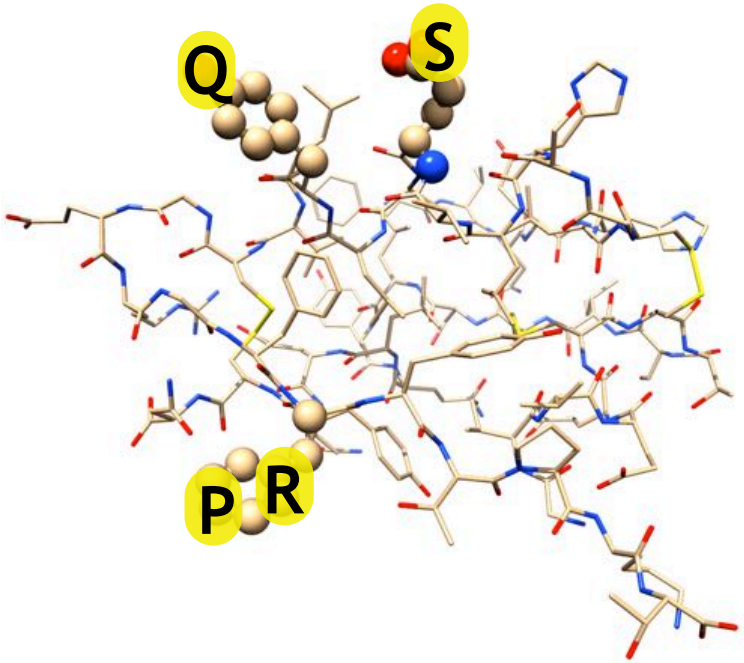
ΔE_{PQ} determined to FOT precision

E_{corr} determined to FOT precision

Summary of the DEC scheme



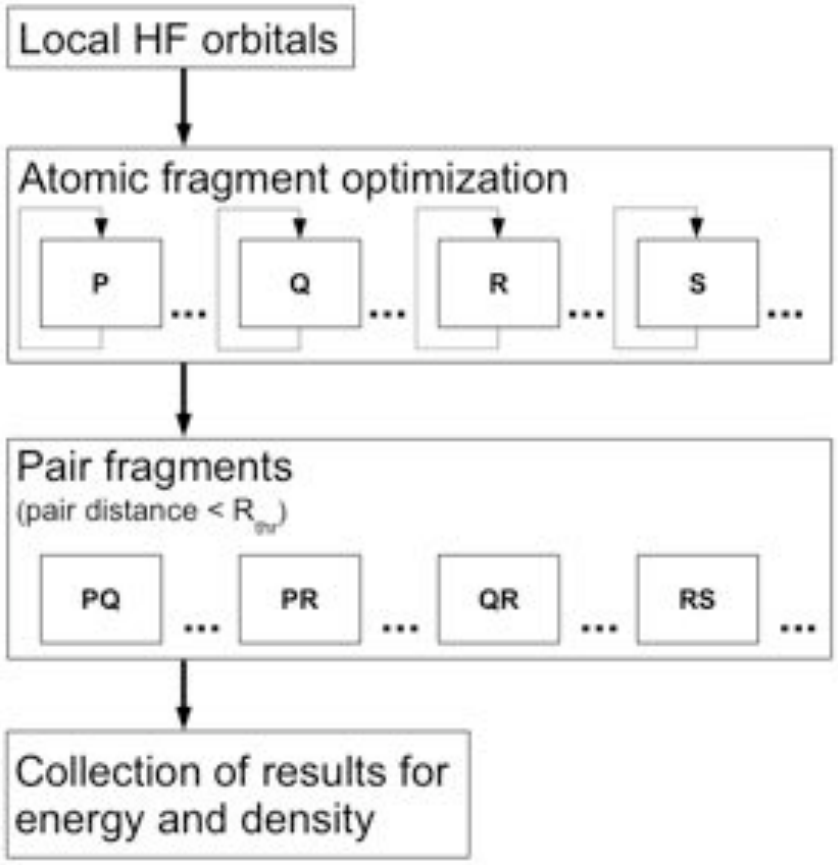
N atomic fragments



Formally: N*(N-1)/2 pair fragments
Use cut-off: const*N pair fragments

DEC is linear scaling and embarrassingly parallel

Parallelism in DEC calculation



Parallelism at three levels

Coarse grained parallelism

All fragment calculations E_P and ΔE_{PQ} carried out independently

Medium and fine grained parallelism

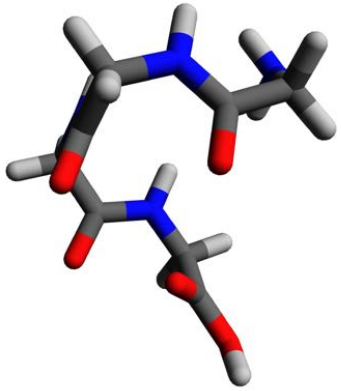
Individual fragment calculations parallelized at two levels (MPI and OpenMP)

(as in conventional implementations)

DEC Error vs. FOT

4 glycine residue in α-helix structure

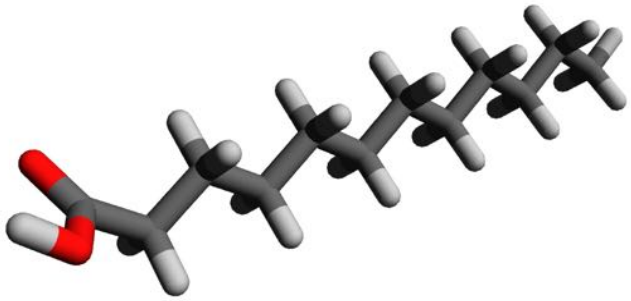
cc-pVDZ



FOT	10^{-3}	10^{-4}	10^{-5}
Δ MP2	$3.61\ 10^{-2}$	$5.36\ 10^{-3}$	$3.99\ 10^{-4}$
Δ CCSD	$9.33\ 10^{-3}$	$1.16\ 10^{-3}$	$-6.03\ 10^{-4}$
ΔCCSD(T)	$2.06\ 10^{-2}$	$3.06\ 10^{-3}$	$-2.81\ 10^{-4}$

Dodecanoic acid

cc-pVTZ

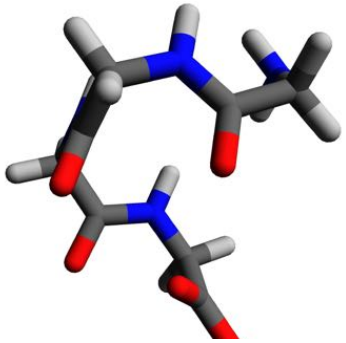


FOT	10^{-3}	10^{-4}	10^{-5}
Δ MP2	$4.68\ 10^{-2}$	$4.54\ 10^{-3}$	$5.50\ 10^{-4}$
ΔCCSD	$2.77\ 10^{-3}$	$6.93\ 10^{-4}$	$-1.37\ 10^{-5}$
Δ CCSD(T)	$3.78\ 10^{-2}$	$2.59\ 10^{-3}$	$3.49\ 10^{-4}$

Pair interaction energies ΔE_{PQ}

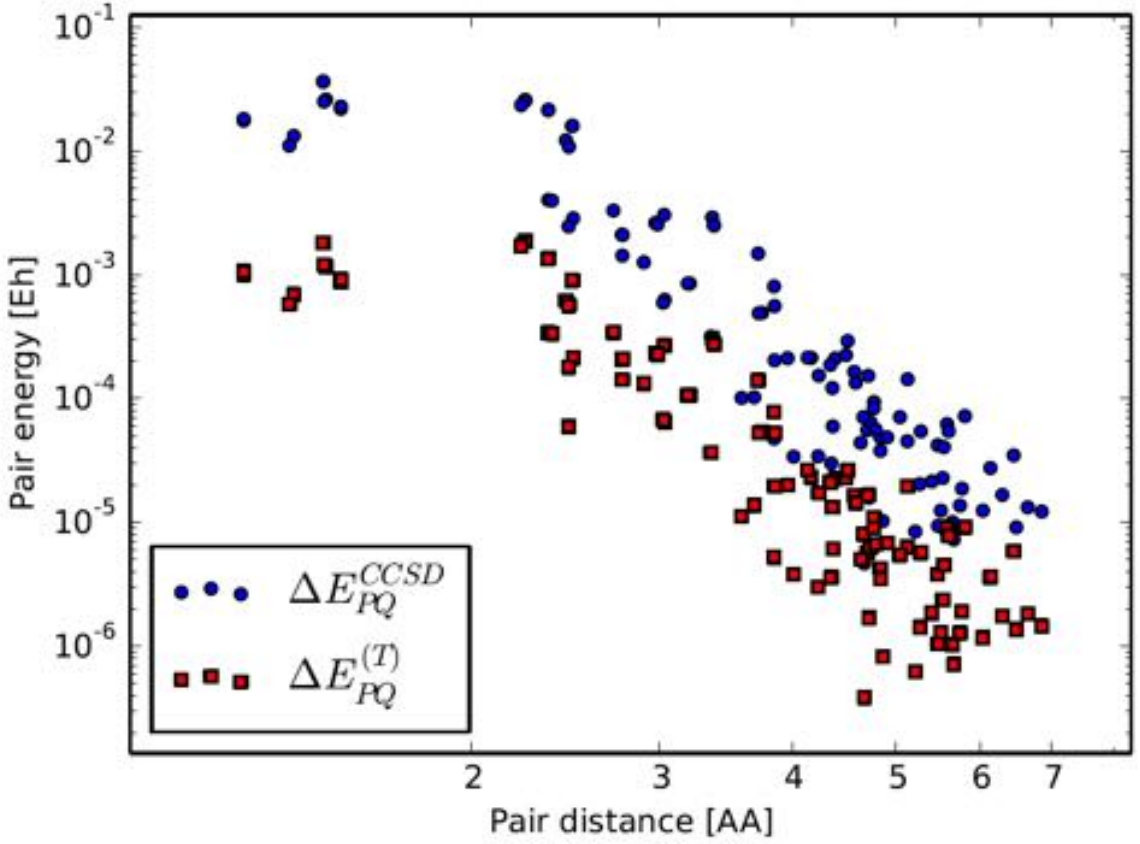
alpha-helix with
4 glycine
residues

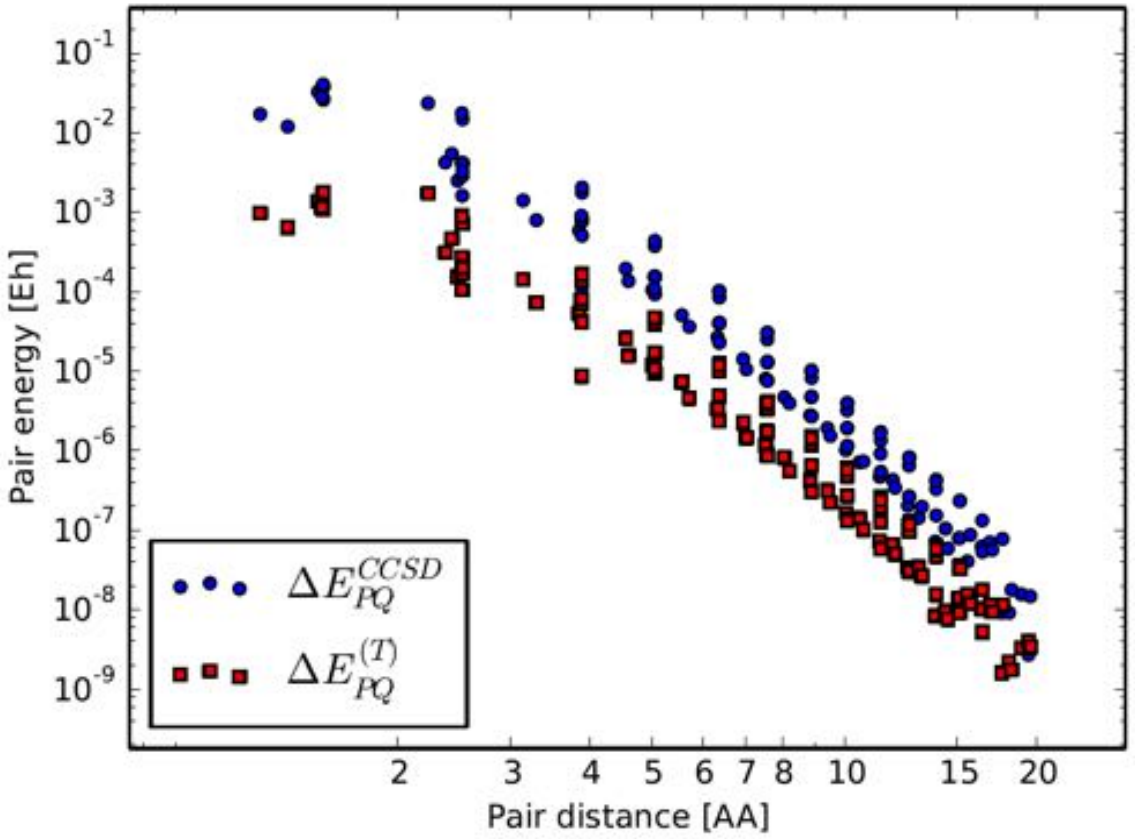
cc-pVDZ



Hexadecanoic acid

cc-pVDZ





DEC Correlation energy(au) vs. FOT

FOT	10^{-3}	10^{-4}
MP2	-11.41503	-11.59505
CCSD	-12.18642	-12.25120
CCSD(T)	-12.63602	-12.74882

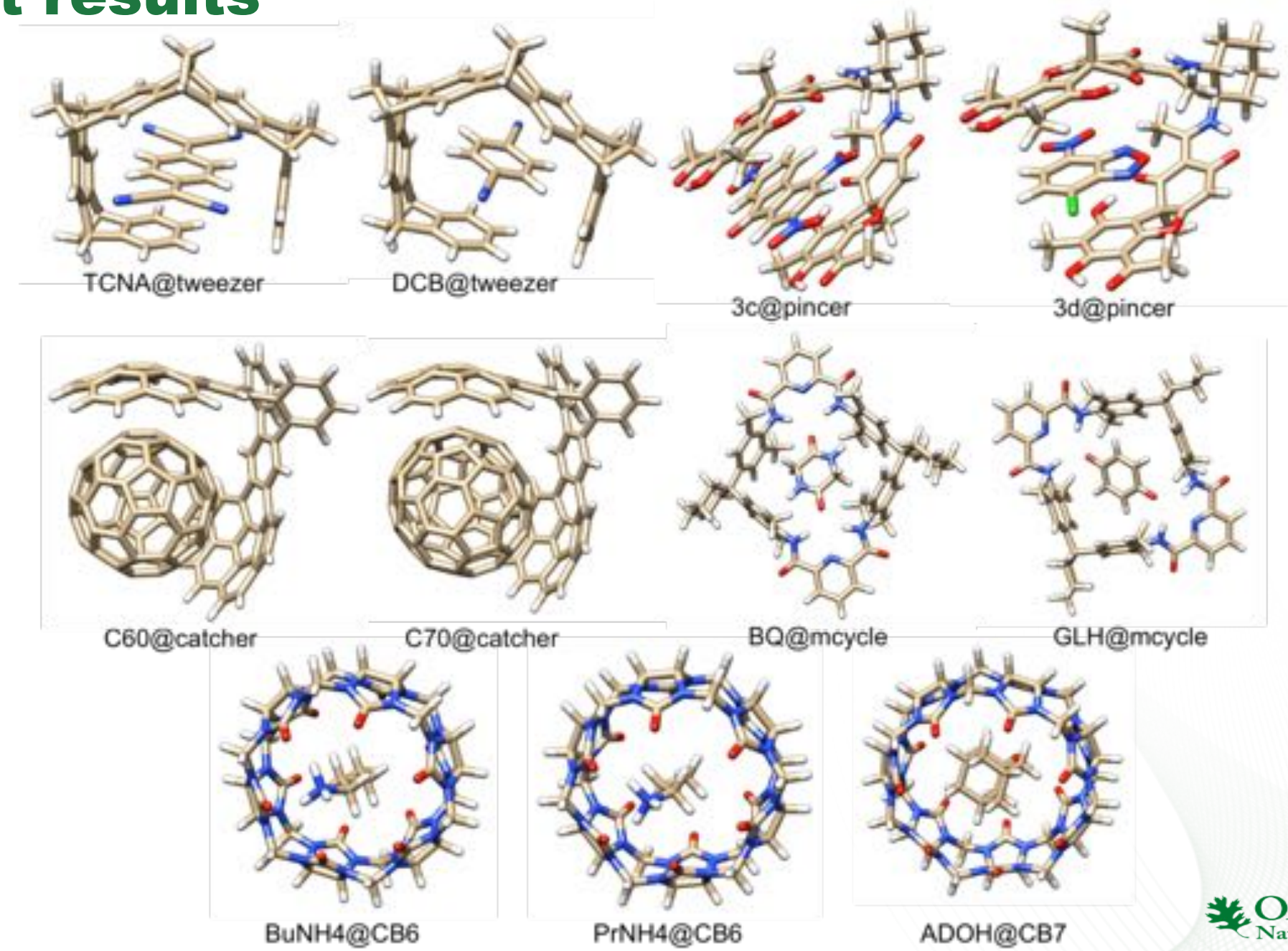
$C_{64}O_2H_{128}$

cc-pVTZ

3772 AOs

- 10^{-4} DEC-CCSD(T) is possible with current computer facilities for large systems

Test set results

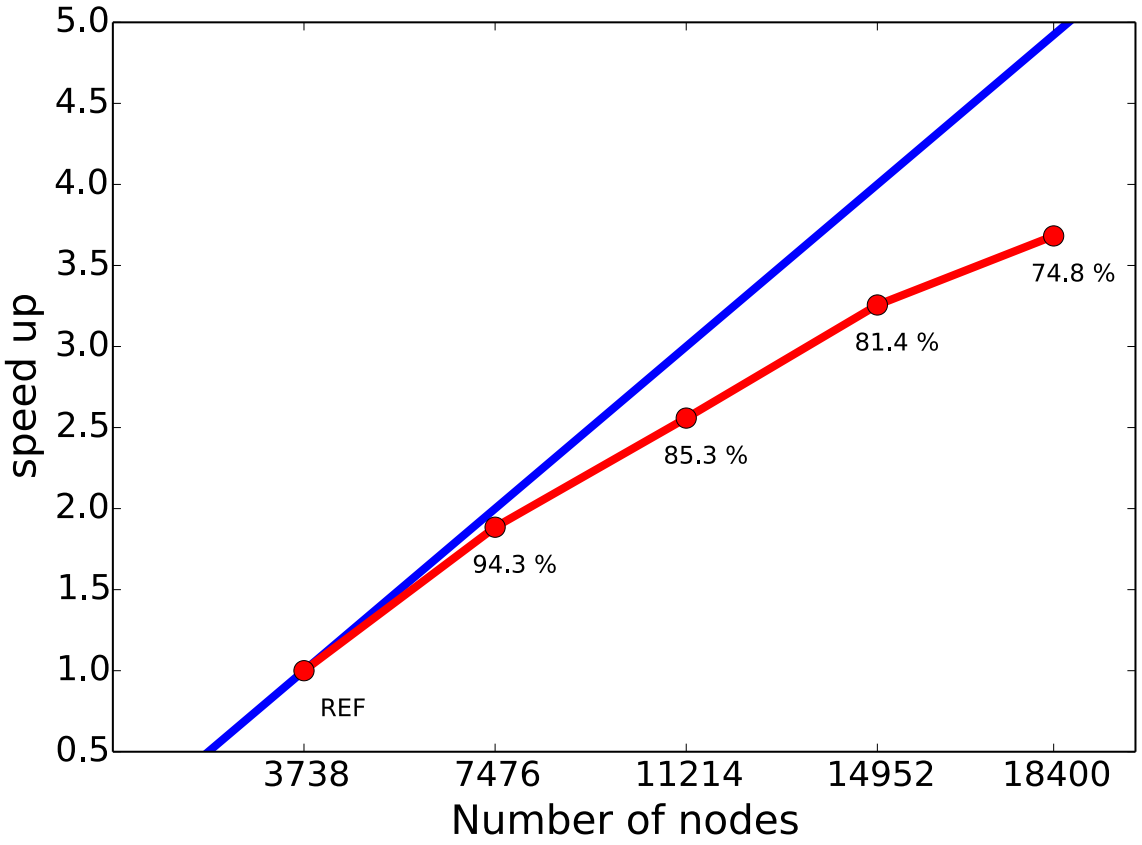


DEC examples

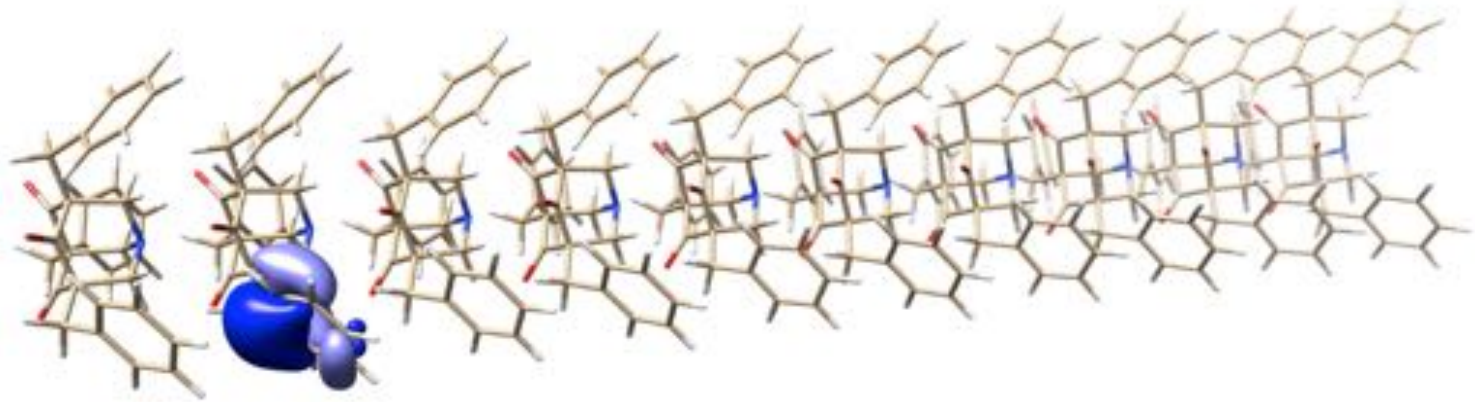
- Calculate MP2 correlation energy and density using cc-pVDZ basis.
- Nanospresso (system **1**): 528 atoms (4278 BF.)
- Nanospresso Doppio (system **2**): 1056 atoms (8556 BF.)
- **Parallel (strong) scaling** (ideal: 0.5)

TTS(11780) / TTS(5890) = 0.56 (system 1)

TTS(11780) / TTS(5890) = 0.53 (system 2)



Nanospresso



Conclusion for DEC

- New strategy for CC energy and properties which explore locality efficiently
- Full system in terms of CC calculations on small fragments of total orbital space
- Linear scaling and massive parallel algorithm
- Full control of errors in energy and cluster amplitudes
- Black box method

DEC perspective

MP2

- Energy, density, molecular gradient and geometry optimizer
- F12 energies, Nuclear shifts, ...

CCSD

- Energy implemented
- Future work, density, molecular gradient, F12 energy ...

CCSD(T)

- Energy implemented
- Future work, density, molecular gradient, ...

Thanks!

ORNL:
D. Liakh
A. Barnes
D. Bykov

Aarhus:
P. Ettenhuber,
J. Erikssen,
K. Kristensen,
P. Baudin,
T. Kjaergaard,
Y. M. Wang,
F. Pawlowski

P. Jørgensen