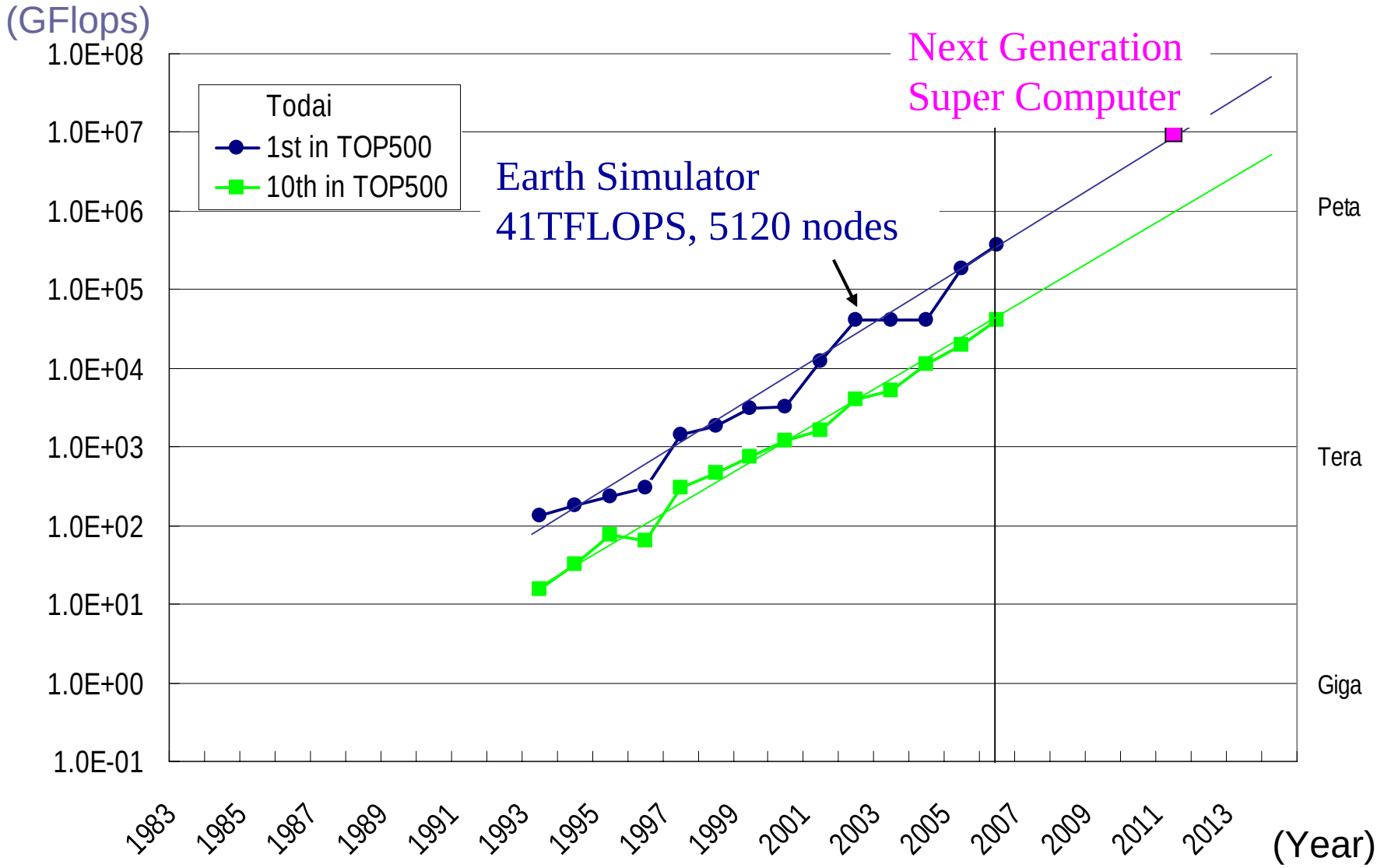
NWChem Meeting on Science Driven Petascale Computing and Capability Development at EMSL
January 25-26, 2007 W.R. Wiley Environmental Molecular Sciences Laboratory
Richland, WA

Toward A Practical DFT For Large Systems

T.Tsuneda, T. Nakajima, and K. Hirao
The University of Tokyo

東京大学
THE UNIVERSITY OF TOKYO

Next Generation Supercomputer Project in Japan



RIKEN Next Generation Supercomputer Project

APRIL 2006: Japan's "3rd Science and Technology Basic Plan" (FY2006-FY2010) launched, placing Supercomputer as "Key Technologies of National Importance" (National Infrastructure)

Total budget $1 billion (2006 - 2012)
10Peta (10^{16}) FLOPS Machine
(250 times faster than Earth Simulator)
Memory 2.5PetaB
Electricity less than 30MW
Floor Space 3,200m^2

Number names in Japanese

一、	十、	百、	千、	万、	億、	兆、	京、
10^{0}	10^{1}	10^{2}	10^{3}	10^{4}	10^{8}	10^{12}	10^{16}

垓、	杼、	穰、	溝、	澗、	正、	載、	極、
10^{20}	10^{24}	10^{28}	10^{32}	10^{36}	10^{40}	10^{44}	10^{48}

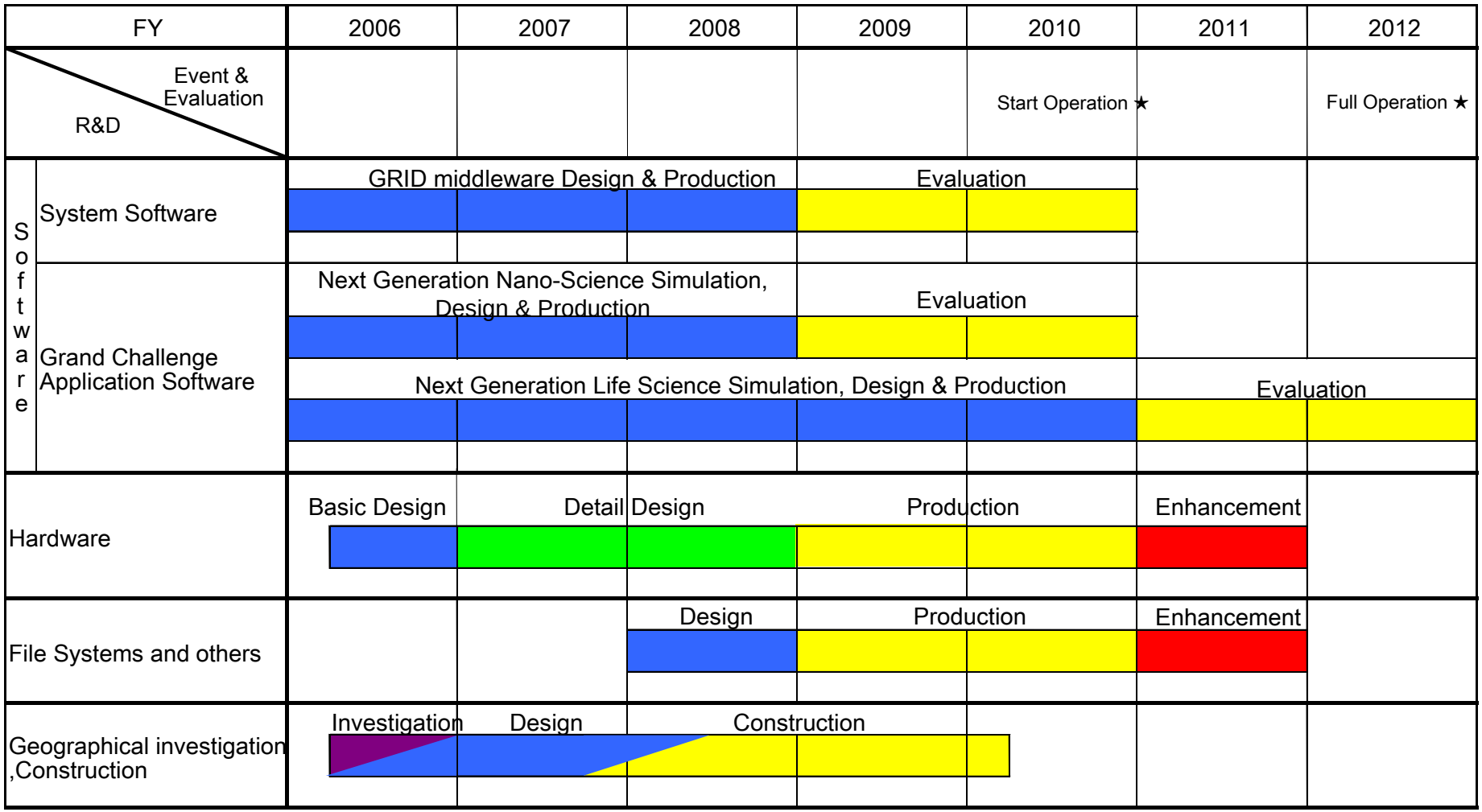
恒河沙、	阿僧祇、	那由他、	不可思議、	無量大数
10^{52}	10^{56}	10^{60}	10^{64}	10^{68}

Number names in Japanese

							Kei
一、	十、	百、	千、	万、	億、	兆、	京、
10^{0}	10^{1}	10^{2}	10^{3}	10^{4}	10^{8}	10^{12}	10^{16}
垓、	杼、	穣、	溝、	澗、	正、	載、	極、
10^{20}	10^{24}	10^{28}	10^{32}	10^{36}	10^{40}	10^{44}	10^{48}
恒河沙、	阿僧祇、	那由他、	不可思議、	無量大数			
10^{52}	10^{56}	10^{60}	10^{64}	10^{68}			

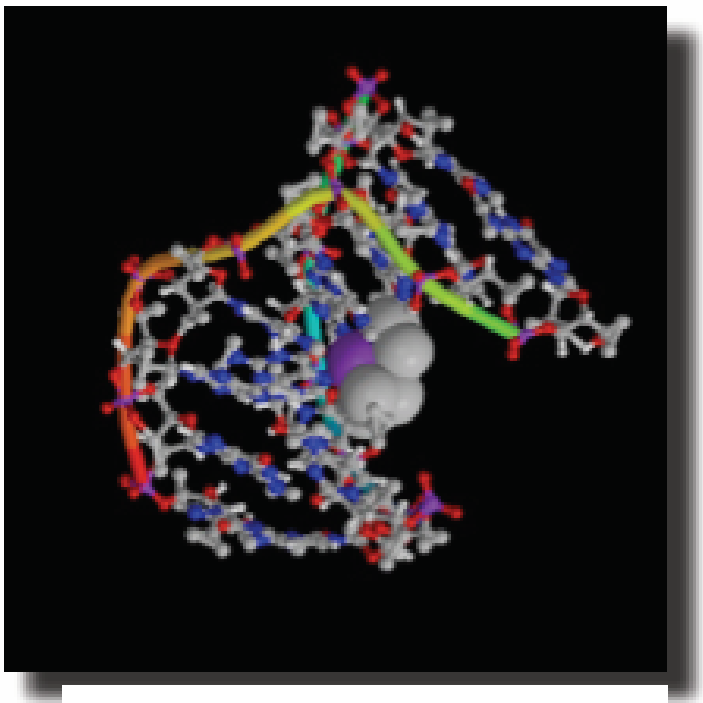
10Peta flops machine is called as Kei-soku（京速）.

Current Schedule

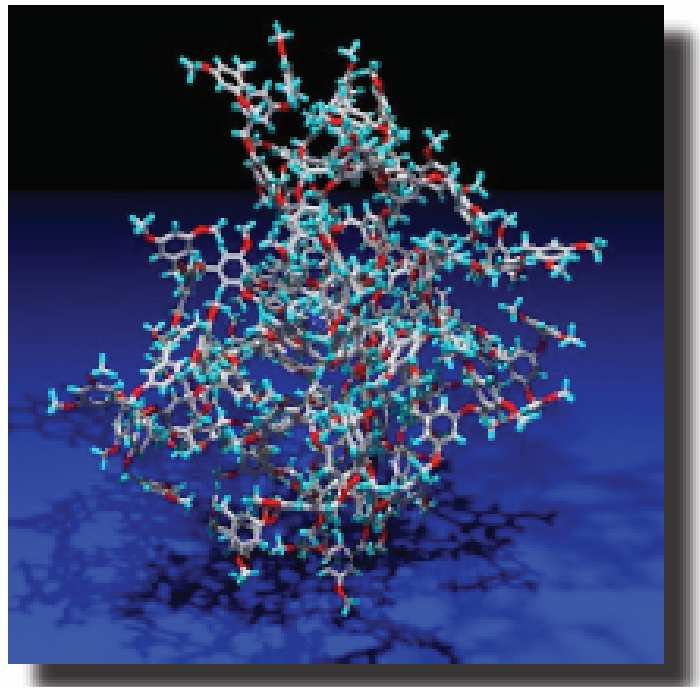


Nano-Bio Simulation

With the emergence of peta-scale computing platforms we are entering a new period of modeling. The computer simulations can be carried out for larger, more complex, and more realistic systems than ever before.



Biomolecules



Nano-scale molecules

Two Approaches to Molecular Theory

- **Wave Function Method**

  Huge and successful efforts in the last 30 years

  State-of-the-art methodology

  Systematic, Converging to exact solution.

  Accurate results for small systems

  Steep N dependence of the computational effort

- **Density Method (Density Functional Theory)**

  Not reached such a mature stage as wave function method

  Not *ab initio*, but best semi-empirical

  Simple and conceptual, Applicable to large systems

  Accuracy depends on *xc* functionals

UT Research Activities

Ab initio Theory

MRMP, MCQDPT
CASVB, QCAS, SPS-SCF&PT
Linear Scaling Method
(PS, RI, Local MP2, Plane wa

DFT

OP Correlation
Parameter-Free Exchange
Long-range corrected functional
van der Waals
TDDFT
Dual-level DFT

UTChem

Relativistic Theory

RESC, DK3
Dirac-Hartree-Fock
Dirac-Kohn-Sham
Ab initio Model Potential
Relativistic Basis Sets

Dynamics

Ab initio dynamics
(TD)DFT dynamics
Hybrid QM/MM
VSCF,VCI

Density Functional Theory (DFT)

DFT may be the only tool that enables us to carry out accurate simulations for larger systems with reasonable computational cost. If practical DFT is developed, which can handle biomolecules and nanomaterials, we can enlarge greatly the scope of computational chemistry.

Outline

- *Fast Coulomb Calculation with Gaussian and Finite Element Coulomb Method*
- *Dual-Level Approach to DFT*
- *Accurate Description of Van der Waals Interactions*

Gaussian and Finite Element Coulomb (GFC) Approach

Kurashige, Nakajima, Hirao, submitted to J.Chem.Phys.

Three time-consuming steps for DFT with GGA

- **Numerical integration of exchange-correlation (*xc*) part**

  can be implemented in linear-scaling fashion using Becke's weighting scheme

- **Coulomb part**

  is very often the most time consuming one, in particular with GGA functionals

- **Diagonalization of Fock matrix**

  scales cubically but insignificant compared to that of the computation of either *xc* or Coulomb for systems with up to several thousands basis functions

Fast Numerical Methods

Much effort has been made to develop efficient methods in evaluation of Coulomb integrals. Integral prescreening technique reduces the scaling from *O(N⁴) to O(N²)*. Furthermore several efficient computational methods have been proposed.

Auxiliary functions

Gaussians Fast Multipole Moment Method (FMM), ~ *O(NlogN)*.

White, Johnson, Gill, Head-Gordon, JCP(1996)

Resolution of the Identity (RI) Approach, *Vahtras, Almlof, Feyerisen, CPL(1993)*

$$J_{pq} = \sum_{A}^{M} (pq|A) \sum_{B}^{M} \left[(A|B)^{-1} \left(\sum_{rs}^{N} D_{rs} (B|rs) \right) \right]$$

4c ERIs are decomposed into 3c and 2c ERIs.

Plane waves Mixed Basis Method, ERI with PW scales as ~ *O(M)*.

Lippert, Hunter, Parrinello, MP(1997), Fusti-Molnar, Pulay, JCP(2002), Kurashige, Nakajima, Hirao, CPL(2006)

$$J_{pq} = J_{pq}^{Gauss} + J_{pq}^{GAU-PW} = J_{pq}^{Gauss} + \sum_{h}^{M} w_h R_p^S(h) R_q^S(h) J^S(h), \quad J^S(k) \frac{4\pi}{k^2} \rho^S(k) \qquad k-space$$

Grid basis Pseudospectral (PS) Method, ~ *O(N²M)*. PS combines analytical basis sets with numerical grid basis functions, *Friesner CPL(1985), Nakajima, Hirao, JCP(2004)*

$$J_{pq} = \sum_{g}^{M} w_g R_p(g) R_q(g) \left(\sum_{rs}^{N} D_{rs} A_{rs}(g) \right), \quad A_{rs}(g) = \int \frac{\chi_r(g')\chi_s(g')}{|g-g'|} dg'$$

Poisson Equation

- A possible alternative to achieve the linear-scaling is to use the Poisson equation with the real-space method for evaluation of the Coulomb potentials.
- The Poisson equation is represented on the grids with finite-difference, finite-element, or wavelet basis functions.
- A huge number of grids would be demanded to describe the Coulomb potential near the nucleus.
- The finite-difference method with the atom-centered grid is not competitive against the conventional analytical integration.

Gaussian and Finite Element (FE) Coulomb Approach

- A new linear scaling method for the fast evaluation of Coulomb integrals with Gaussian basis functions

- Coulomb potential is expanded in the mixed Gaussian and finite-element auxiliary functions that express the core and smooth Coulomb potentials, respectively

- Coulomb integrals can be evaluated by three-center one-electron overlap integrals

- The computational cost and scaling for large molecules are drastically reduced

Gaussian and Finite Element Coulomb Approach

Coulomb integrals are given with Gaussians and Coulomb potentials

$$J_{pq} = \int dr \chi_p(r) \chi_q(r) v(r)$$

$$v(r) = \int dr' \frac{\rho(r')}{|r'-r|}, \qquad \rho(r') = \sum_{rs} D_{rs} \chi_r(r') \chi_s(r')$$

We expand Coulomb potentials in terms of auxiliary functions, Gaussians and Finite Element Basis as

$$v(r) = \sum_i c_i^{FE} f_i^{FE}(r) + \sum_i c_i^{Gauss} f_i^{Gauss}(r)$$

The atom-centered Gaussian functions represent the spherical core potential near a nucleus, while uniform finite-element functions, a tensor product of one-dimension Lagrange interpolate polynomials, represent the residual, which would be smooth across the board.

Coulomb Integrals

Coulomb integrals can be evaluated by overlap integrals among two Gaussian basis functions and one auxiliary function

$$
\begin{aligned}
J_{pq} &= \int dr \chi_p(r) \chi_q(r) v(r) \\
&= \sum_i c_i^G \int dr \chi_p(r) \chi_q(r) f_i^G(r) + \sum_i c_i^{FE} \int dr \chi_p(r) \chi_q(r) f_i^{FE}(r)
\end{aligned}
$$

No four-center two-electron integrals

Drastically reduces the computational cost

Poisson Equation

The expansion coefficients $\{c_i\}$ can be obtained by solving the Poisson equation, which is solved algebraically by the Galerkin method

$$-\nabla^2 V(r) = 4\pi\rho(r) \xrightarrow[\text{functions}]{\text{auxiliary}} \begin{bmatrix} \mathrm{O} & & 0 \\ & \mathbf{A} & \\ 0 & & \mathrm{O} \end{bmatrix} \times \begin{bmatrix} \mathrm{M} \\ \mathbf{c} \\ \mathrm{M} \end{bmatrix} = \begin{bmatrix} \mathrm{M} \\ \mathbf{b} \\ \mathrm{M} \end{bmatrix}, \quad \begin{cases} \mathbf{A}_{\mathrm{ij}} = \underbrace{\int \nabla f_i(r) \cdot \nabla f_j(r) dr}_{\text{kinetic integral}} \\ \mathbf{b}_i = \underbrace{\int \rho(r) \cdot f_i(r) dr}_{\text{overlap integral}} \\ V(r); \ \sum_i \mathbf{c}_i \cdot f_i(r) \end{cases}$$

- The linear equation is solved by using CG method.
- The matrix A is independent of $\rho(r)$ and contains only kinetic-like integrals. It is extremely sparse in the localized auxiliary basis functions.
- The Poisson equation scales as $O(N)$.

Auxiliary Functions

To express the atomic core potential, we used Gaussians with even-tempered sequence

$$\alpha_i = a \cdot b^{i-1} \qquad i = 1, 2, \cdots, n$$

Parameters *a* and *b* are determined to maximize the Coulomb energy with the density obtained from an atomic UHF calculation.
The number of Gaussians was incremented until the error in the Coulomb energy becomes lower than 0.02 mEh.

We used the auxiliary Gaussians,

H(3s), Li (5s), Be(7s), C(8s3d), N(9s), O(10s4d), and F(10s4d)

C, O and F contain d-type functions since their UHF densities are not totally symmetric.

A cubic uniform finite-element, 1.8 bohr on a side, with third-order Lagrange interpolation polynomials.

Performance

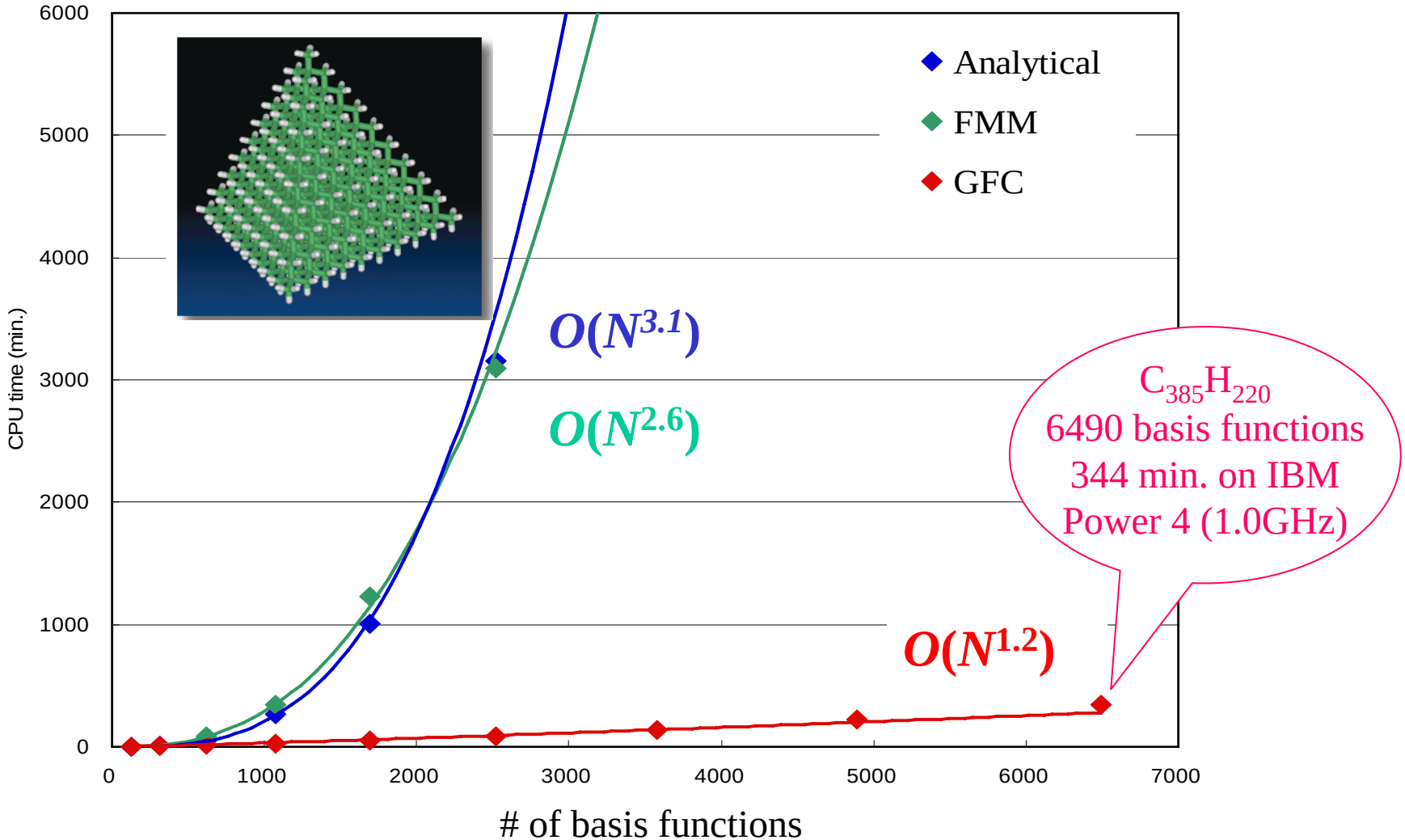
CPU：IBM Power 4 1.0GHz BLYP / SVP

Analytical and FMM results are computed using GAMESS

Table I Equilibrium bond lengths (R_e), harmonic frequencies (ω_e), atomization energies (D_e) and equilibrium total energies (E_{KS}) for diatomic molecules using BLYP/6-31G**

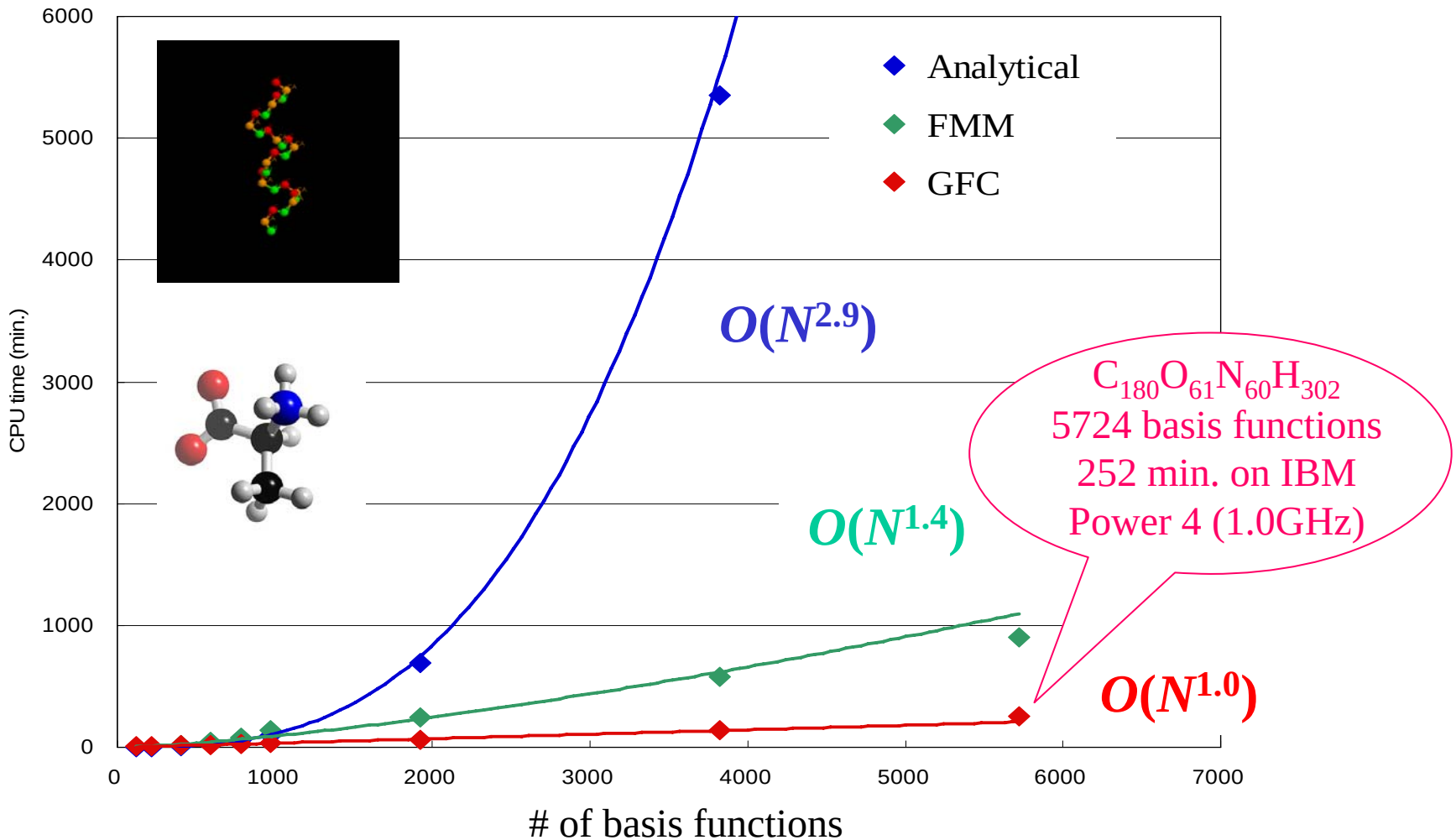
	Analytical integrals				GFC method			
	R_e(Å)	ω_e(cm^{-1})	D_e(kcal/mol)	E_{KS}(hartree)	R_e(Å)	ω_e(cm^{-1})	D_e(kcal/mol)	E_{KS}(hartree)
BeH	1.354	2036	57.1	-15.2422	1.354	2035	57.1	-15.2423
CH	1.146	2725	84.7	-38.4604	1.145	2727	84.8	-38.4607
CN	1.187	2112	189.2	-92.6979	1.187	2111	189.3	-92.6981
CO	1.150	2153	260.6	-113.2907	1.150	2153	260.7	-113.2911
F_2	1.437	994	55.4	-199.4907	1.436	991	55.6	-199.4912
FH	0.936	4016	133.7	-100.4097	0.938	4012	134.0	-100.4103
H_2	0.747	4485	111.1	-1.1679	0.747	4488	111.1	-1.1680
Li_2	2.732	335	20.2	-14.9922	2.734	335	20.2	-14.9923
LiF	1.561	1051	137.8	-107.4008	1.561	1051	137.8	-107.4010
LiH	1.623	1377	57.1	-8.0664	1.622	1376	57.1	-8.0665
N_2	1.118	2388	234.4	-109.5063	1.116	2387	235.2	-109.5076
NH	1.060	3163	87.7	-55.2015	1.059	3164	88.0	-55.2021
NO	1.176	1909	165.3	-129.8753	1.175	1914	166.1	-129.8767
O_2	1.239	1551	138.4	-150.311 6	1.239	1552	138.7	-150.3123
OH	0.992	3617	106.2	-75.7102	0.991	3619	106.3	-75.7105
MAE	-	-	-	-	0.0006	2	0.24	0.0004

Gaussian and FE Coulomb Approach



3D diamond/ SVP

Gaussian and FE Coulomb Approach



1D Alanine α-helix chain/ SVP

Energy Derivative

The Coulomb force can be evaluated efficiently by using the GFC method. The first derivative of the Coulomb energy E_J with respect to a nuclear coordinate X can be written as

$$\frac{\partial E_J}{\partial X} = \frac{\partial}{\partial X}\left\{\int dr_1 \int dr_2 \rho(r_1)\frac{1}{r_{12}}\rho(r_2)\right\} = 2\int dr_1 \left(\frac{\partial \rho(r_1)}{\partial X}\right) v(r_1)$$

In the GFC method, the Coulomb potential $v(r)$ is expanded in the auxiliary functions. Thus the energy derivative can be evaluated approximately by

$$\frac{\partial E_J}{\partial X} = 2\sum_i c_i^G \int dr \left(\frac{\partial \rho(r)}{\partial X}\right) f_i^G(r) + 2\sum_i c_i^{FE} \int dr \left(\frac{\partial \rho(r)}{\partial X}\right) f_i^{FE}(r)$$

Gaussian and FE Coulomb Approach

Gaussian and FE Coulomb approach offers the best performance for evaluating Coulomb integrals without loss of accuracy. The algorithm is found to scale as $N^{1.0} \sim N^{1.2}$ with system size.

Gaussian and FE Coulomb approach makes the molecular quantum calculations affordable for very large systems involving several thousands of basis functions.

Dual-Level Approach to DFT

J. Chem.Phys., **124**, 184108 (2006)

Hybrid GGA Functionals

The success of Kohn-Sham DFT was the development of *xc* functionals depending on density gradients in addition to the density itself (GGA).

A further advance is the mixing of a small fraction of exactly computed HF exchange with GGA exchange such as B3LYP, LC-GGA, etc.

Although hybrid GGA improves the accuracy, it also makes the calculation more expensive.

Fast algorithms for Coulomb integrals cannot be employed for HF exchange because its algebraic structure is not compatible with them.

Exchange Integrals

Fast algorithms for Coulomb interaction cannot be employed for HF exchange because its algebraic structure is not compatible with them.

Only the pseudospectral method can be applied to HF exchange but it scales as $O(N^2M)$.

$$J_{pq} = \sum_{g}^{M} w_g \chi_p^*(g)\chi_q(g)\left(\sum_{rs}^{N} D_{rs}A_{rs}(g)\right), \quad K_{pq} = \sum_{g}^{M} w_g \chi_p^*(g)\left[\sum_{r}^{N}\left(\sum_{s}^{N} D_{rs}\chi_s(g)\right)A_{rq}(g)\right]$$

$$A_{pq}(g) = \int \chi_p^*(1)\frac{1}{|r_1 - r_g|}\chi_q(1)dr_1$$

Friesner CPL(1985), Nakajima, Hirao, JCP(2004)

The GGA is efficient but less accurate while the hybrid GGA is more accurate but less efficient.

Dual-Level Approach to DFT

J. Chem.Phys., **124**, 184108 (2006)

To perform hybrid GGA DFT calculations for large systems, we have developed the dual-level approach. The approach is based on the low sensitivity of the density to the choice of the functional and the basis set. The total electron density in the ground state can be well represented in terms of the density evaluated using the low-quality basis set and the low-cost *xc* functional.

The large reduction of the computer resources can be achieved at an affordable loss of accuracy.

Dual-Level DFT

Solve KS equation
with low-quality basis set & low-level functional
and obtain a total density

⇩

Use a frozen density approximation
and evaluate the total energy
with high-quality basis set & high-level functional

KS Orbitals in the Extended Basis Function Space

We obtain a set of KS orbitals which are expanded in terms of a finite basis set

$$\{\chi_1, \chi_2, \cdots, \chi_N\}$$

Then we expand the basis set space by adding a few more of basis functions

$$\{\chi_1, \chi_2, \cdots, \chi_N, \chi_{N+1}, \chi_{N+2}, \cdots, \chi_M\}$$

The additional members of the basis set are subject to the condition that they be orthogonal to the original KS occupied orbitals, ϕ_i

$$\langle \phi_i | \chi_k \rangle = 0, \qquad k = N+1, N+2, \cdots, M$$

This can be achieved easily by using the KS density operator P if

$$(1-P)|\chi_k\rangle \rightarrow |\chi_k\rangle \qquad where \qquad P = \sum_i |\phi_i\rangle\langle\phi_i|$$

KS matrix in a higher-level *xc* functional

KS matrix is projected onto the occupied space

$$F_{pq}^{occ} = \sum_{rs} (SD)_{pr} F_{rs} (SD)_{sq}$$

$$F_{pq} = h_{pq} + 2\sum_{PQ} D_{PQ}^{L} (pq|PQ) - t_{EX} \sum_{PQ} D_{PQ}^{L} (pQ|Pq) + t_{XC} v_{pq}^{XC} (\rho^{L})$$

P,*Q* are labels of AO of low-level basis functions and *p*,*q* are labels for the extended basis.

The reference total electron density is defined as

$$\rho^{L}(r) = 2\sum_{i}^{occ} \phi_i^{L*}(r)\phi_i^{L}(r) = 2\sum_{PQ} D_{PQ}^{L} \chi_P^{L*}(r)\chi_Q^{L}(r) = 2\sum_{pq} D_{pq} \chi_p^{*}(r)\chi_q(r)$$

$$with \qquad D_{pq} = \sum_{PQ} \tilde{\Omega}_{pP}^{+} D_{PQ}^{L} \tilde{\Omega}_{Qq}, \quad \tilde{\Omega}_{Pp} = \sum_{q} \tilde{S}_{Pq} S_{qp}^{-1}$$

Total Electronic Energy

The reference energy of the KS total electronic energy is given by

$$E_{KS}^{(0)} = 2\sum_{pq} D_{pq} h_{pq} + 2\sum_{pq}\sum_{PQ} D_{pq} D_{PQ}^{L} \left(pq|PQ\right) - t_{EX} \sum_{pq}\sum_{PQ} D_{pq} D_{PQ}^{L} \left(pQ|Pq\right) + t_{XC} E_{XC}$$

Since no rotations between occupied and virtual orbitals are allowed, Brillouin theorem is not satisfied. The correction to the KS energy is evaluated perturbatively by

$$E_{KS}^{(1)} = 2\sum_{i}^{occ}\sum_{a}^{vir} \frac{\left|F_{ia}\right|^2}{\varepsilon_i - \varepsilon_a}$$

The KS total electronic energy is given by

$$E_{KS} = E_{KS}^{(0)} + E_{KS}^{(1)}$$

Mean Absolute Errors (relative to full B3LYP) for H_2, LiH, BeH, CH, NH, OH, FH, Li_2, LiF, CN, CO, N_2, NO, O_2, and F_2

	r_{CN} (Å)	ν (cm^{-1})	E_{total} (a.u.)	E_{atom} (kcal/mol)
B3LYP:LDA	0.004	28	0.0053	5.5
B3LYP:BLYP	0.003	21	0.0057	6.2

Basis sets, High-level: cc-pVQZ, Low-level: cc-pVQZ(-pol)

E_{total}: total energy (a.u.) E_{Atom}: atomization energy (kcal/mol)

Mean Absolute Errors (relative to full B3LYP) for H_2, LiH, BeH, CH, NH, OH, FH, Li_2, LiF, CN, CO, N_2, NO, O_2, and F_2

	r_{CN} (Å)	ν (cm^{-1})	E_{total} (a.u.)	E_{atom} (kcal/mol)
B3LYP:LDA	0.004	28	0.0053	5.5
B3LYP:BLYP	0.003	21	0.0057	6.2

Basis sets, High-level: cc-pVQZ, Low-level: cc-pVQZ(-pol)
E_{total}: total energy (a.u.) E_{Atom}: atomization energy (kcal/mol)

The CPU (in second) of self-consistent B3LYP/cc-pVQZ and B3LYP/cc-pVQZ:LDA/cc-pVQZ(-pol) calculations of CN

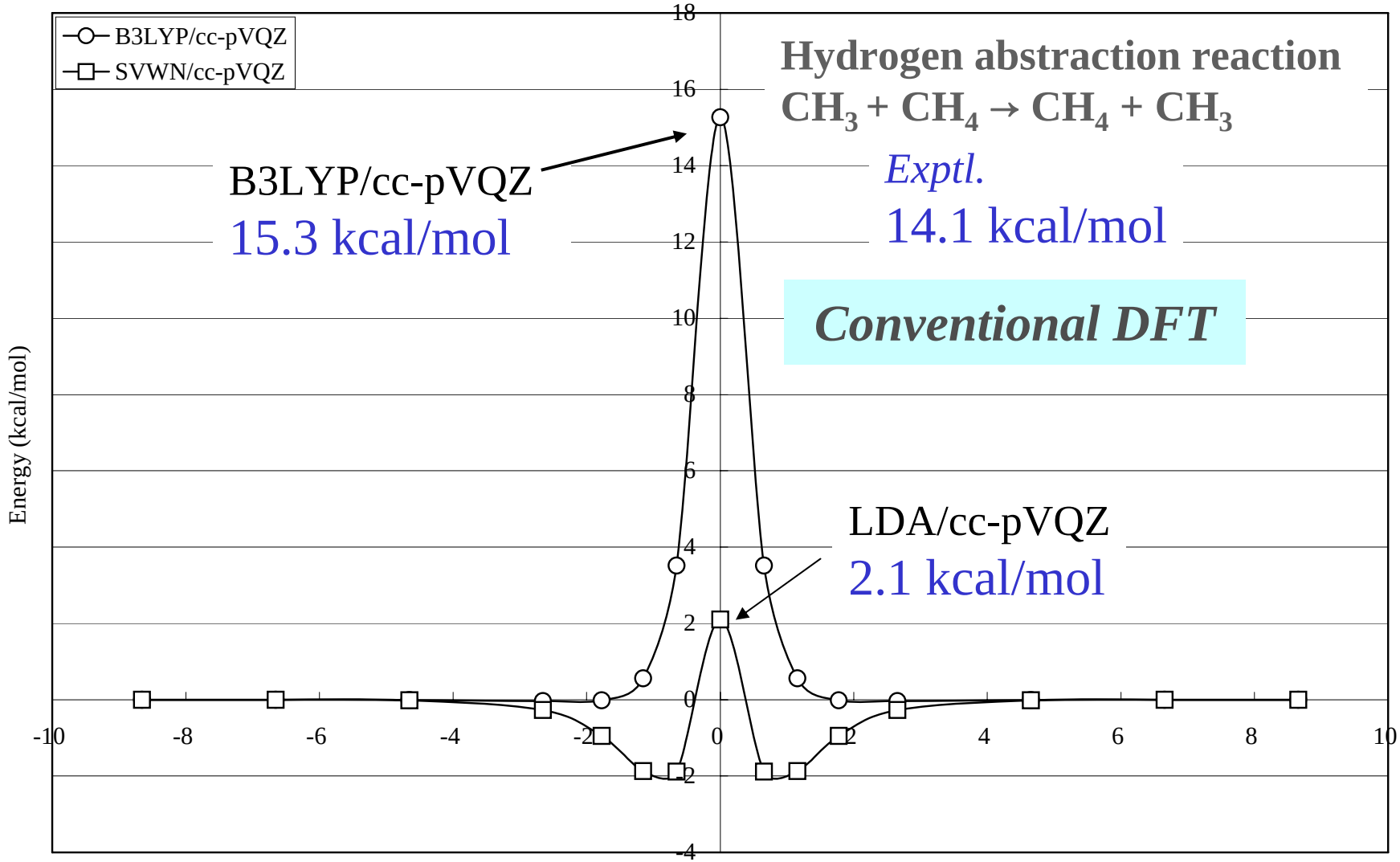
	SCF	KS matrix construction	Energy correction	Total
full B3LYP	2169	------	------	2169
B3LYP:LDA	23	62	0	85

25.5 times faster

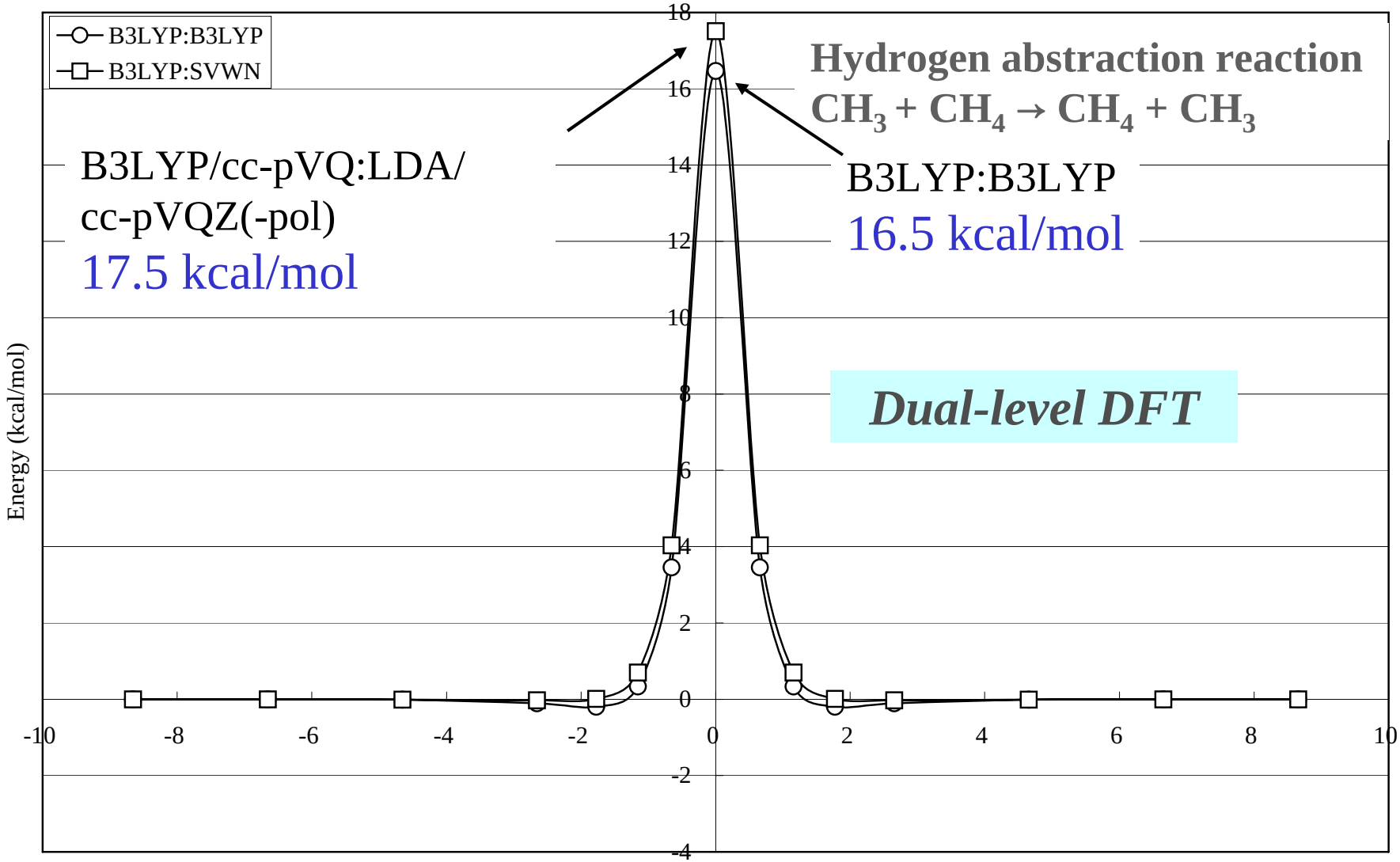
Calculated barrier height with ZPVE (kcal/mol) for the reaction $H_2CO \rightarrow H_2 + CO$

CCSD(T)/cc-pVQZ	81.3
HF/cc-pVQZ(-pol)	105.2
LDA/cc-pVQZ	95.9
B3LYP/cc-pVQZ(-pol)	85.6
B3LYP/cc-pVQZ	78.8
B3LYP/cc-pVQZ: LDA/cc-pVQZ(-pol)	77.7
Exptl.	72.2±0.8

Target value

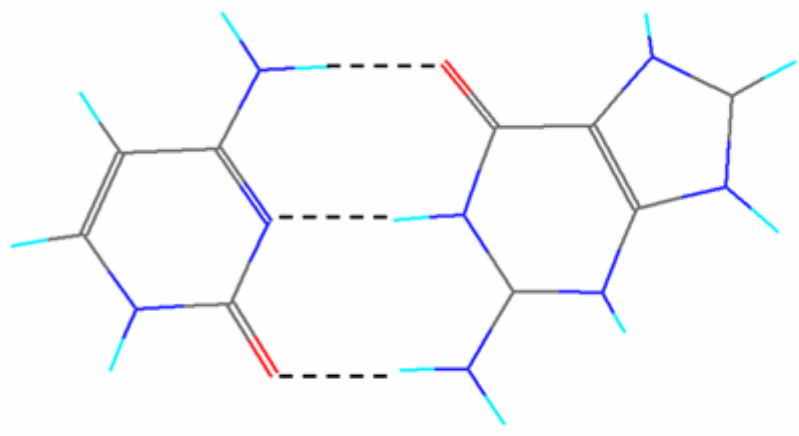


The potential energy curve computed by self-consistent B3LYP/cc-pVQZ (○) and LDA/cc-pVQZ (□) approaches in the hydrogen abstraction reaction $CH_3 + CH_4 = CH_4 + CH_3$.

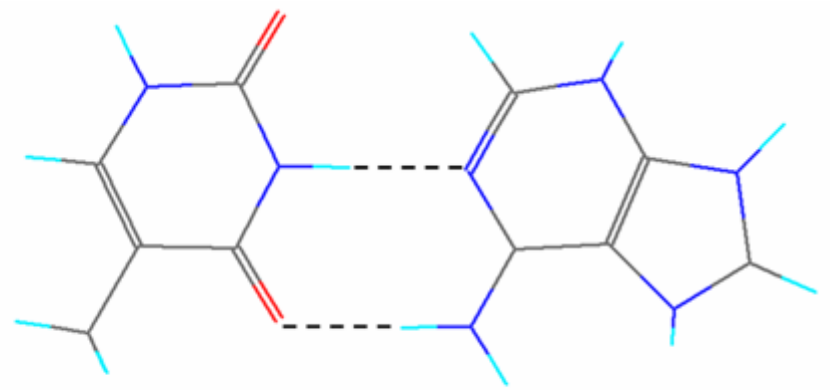


The potential energy curve computed by dual-level B3LYP:B3LYP (○) and B3LYP:LDA (□) approaches in the hydrogen abstraction reaction $CH_3 + CH_4 = CH_4 + CH_3$.

Calculated interaction energies (kcal/mol) between two DNA bases



Cytosine-Guanine



Thymine-Adenine

Calculated interaction energies (kcal/mol) between two DNA bases

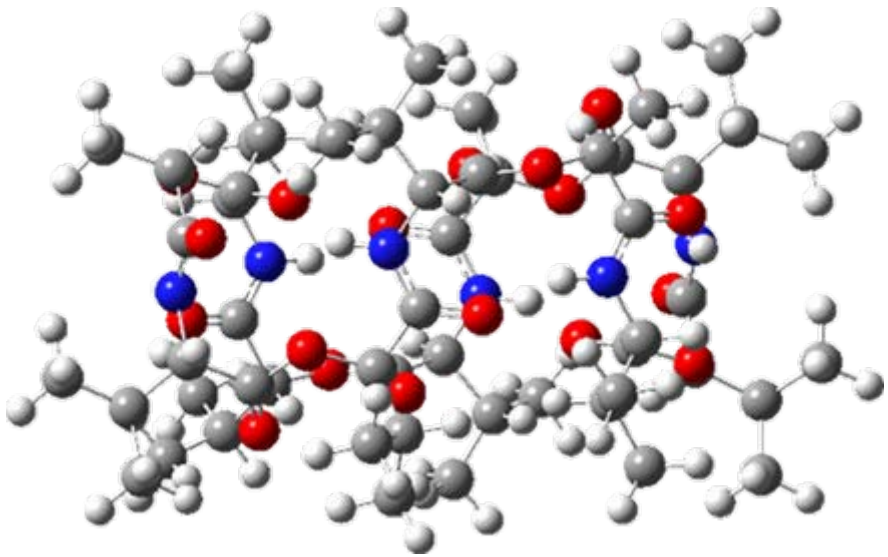
Cytosine-Guanine pair		
MP2/6-31G**	31.8	
B3LYP/6-31G	35.4	Target value
LDA/6-31G	49.3	
B3LYP/6-31G**	31.1	
B3LYP/6-31G**: LDA/6-31G	31.1	
Adenine-Thymine pair		
MP2/6-31G**	17.4	
B3LYP/6-31G	19.9	Target value
LDA/6-31G	30.4	
B3LYP/6-31G**	16.5	
B3LYP/6-31G**: LDA/6-31G	16.5	

Timing of Dual-Level DFT

Timing of dual-level DFT

	# of basis functions	Lower-level (h:m:s)	Higher-level (h:m:s)	Total cpu (h:m:s)	Total Energy (au)
B3LYP/6-31+G**	1854	------	147:00:36	147:00:36	-3794.9322
B3LYP:LDA/6-31G	**882**	**2:51:58**	**7:22:36**	**10:15:35**	**-3794.9481**

14.3 times faster



Valinomycin
($C_{54}H_{90}N_6O_{18}$)

Conclusions

The dual-level DFT approach works quite well and the large reduction of the computer resources can be achieved at an affordable loss of accuracy. Hybrid functionals can now be applied to bio and nano-scale systems.

Accurate Description of van der Waals Interactions

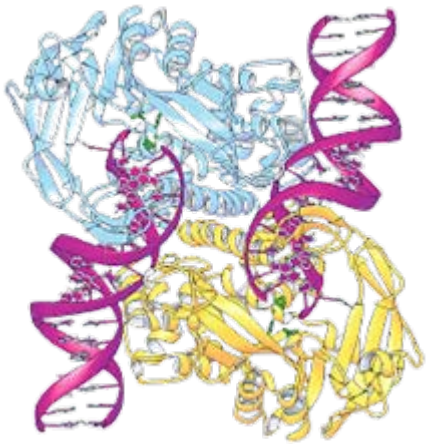
J. Chem.Phys., **123**, 104307 (2005)

Mol.Phys. (Handy Issue) **103**, 1151 (2005)

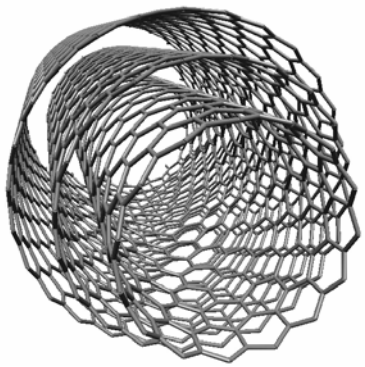
J. Chem.Phys., **117**, 6010 (2002)

Van der Waals interactions

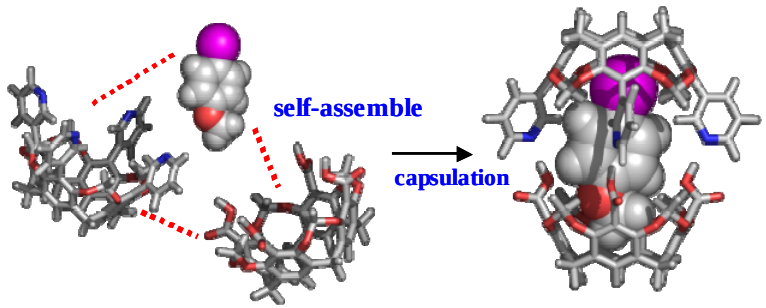
Due to this failure, application of DFT is limited.



Biomolecules

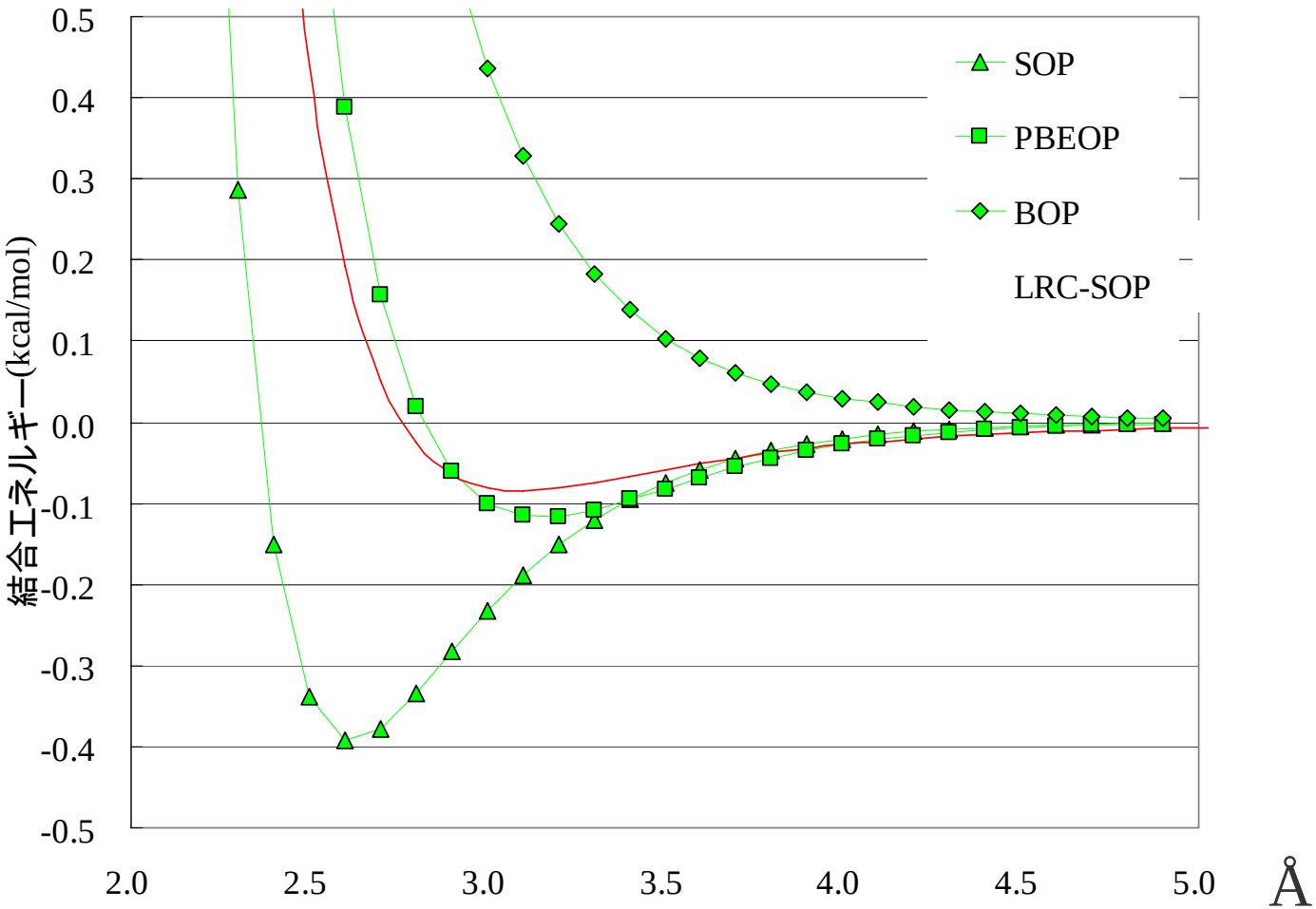


Nanotubes



Host-Guest systems

Potential Energy Curves of Ne-Ne



A strong functional dependency appeared in the pure DFT potential

Van der Waals interactions

$$\Delta E^{vdW} = E \text{ (Pauli repulsion) - } E \text{ (Dispersion attraction)}$$

LDA predicts the binding character of vdW interactions. However, LDA severely overestimates the binding energy and yields a too short vdW bond. The GGA predicts repulsive vdW interactions. Thus, none of the functionals account successfully for vdW interactions.

MP2 significantly overestimates the binding energies and MP2 results have strong basis set dependence.
Only CCSD(T) with a large basis set gives the accurate estimation.

Pauli repulsion

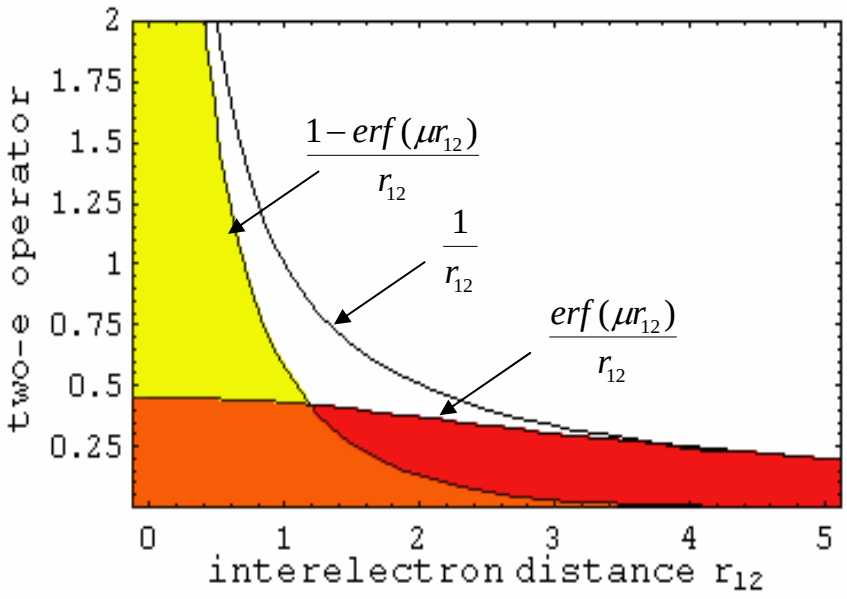
The Pauli repulsion can be described accurately only after the correct long range electron-electron interactions are taken into account in the exchange functional.

By splitting the Coulomb interaction into short-range and long-range components, LC-GGA can be expressed as

$$\frac{1}{r_{12}} = \frac{1-erf\left(\mu r_{12}\right)}{r_{12}} + \frac{erf\left(\mu r_{12}\right)}{r_{12}}$$

⇧ ⇧

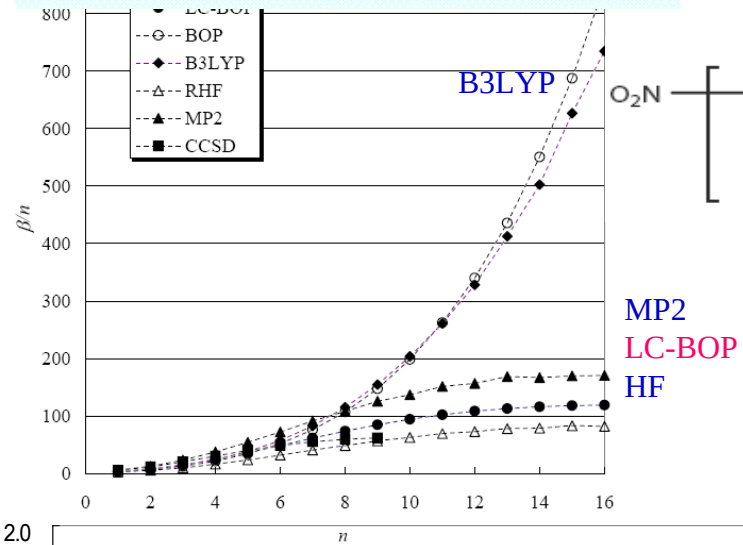
GGA exchange *HF exchange*

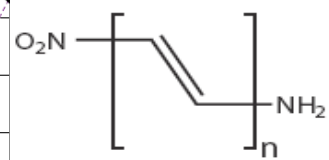


μ *is a parameter controlling the separation*

Hybrid GGA has good energetics, good Rydberg behavior, good CT predictions, and good optical response.

Hyperpolarizability





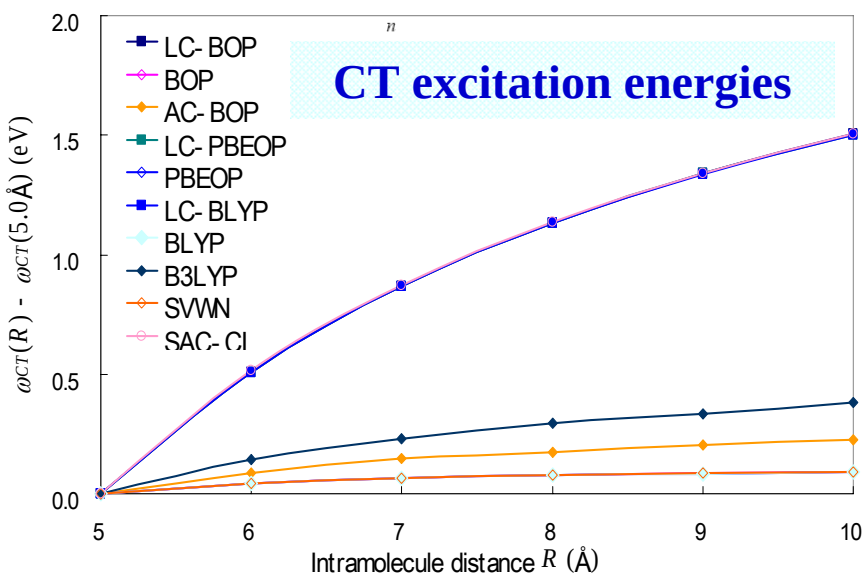
Rydberg excitation energies

	MAE
BLYP	1.54
B3LYP	0.89
LC-BLYP	0.41
SAC-CI	0.19

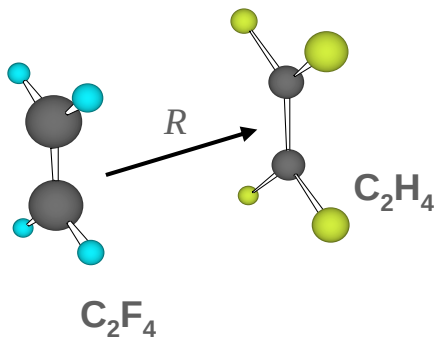
Reaction barrier heights (kcal/mol)

		MAE
BOP	8.7	
B3LYP	5.6	
LC-BOP	2.6	

CT excitation energies



ΔE_{CT}(eV)	
BLYP	5.40
B3LYP	7.49
LC-BLYP	12.49
Exptl.	12.5



Dispersion Attraction

Van der Waals functional (Andersson et al, PRL 1996)

$$E^{dispersion} = -\frac{6}{4(4\pi)^{3/2}} \int_{V_1} d^3\mathbf{r}_1 \int_{V_2} d^3\mathbf{r}_2 \frac{\sqrt{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}}{\sqrt{\rho(\mathbf{r}_1)} + \sqrt{\rho(\mathbf{r}_2)}} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|^6}$$

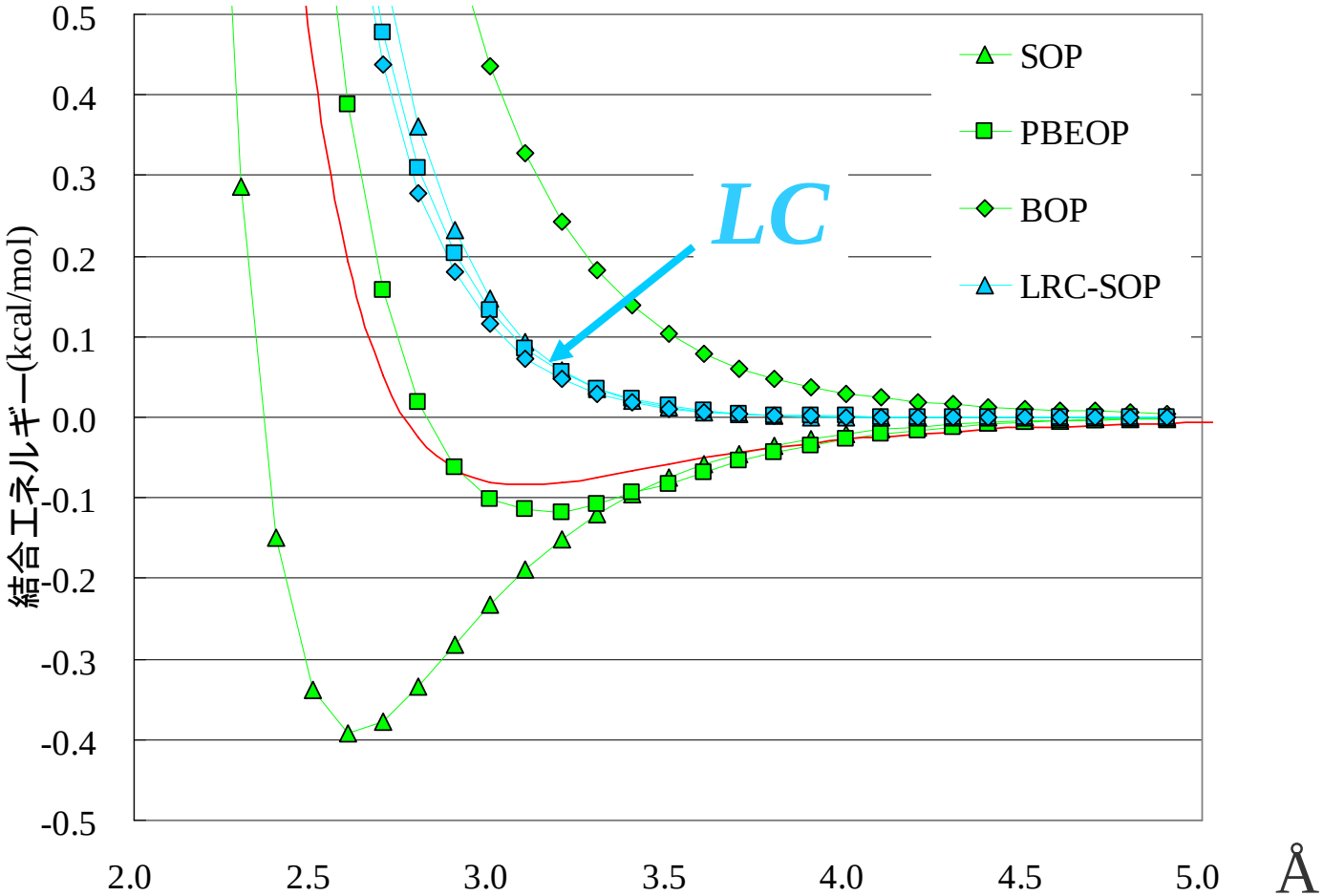
The functional provides an accurate dispersion energy between well-separated electron distributions.

multiplied by a damping factor

$$f_{damp} = \exp\left[-\left(\frac{\alpha_{AB}}{r_{12}}\right)^6\right] \qquad (\alpha_{AB} = C_1 R_{AB} + C_2)$$

R_{AB} : Clementi's atomic radii (1963)

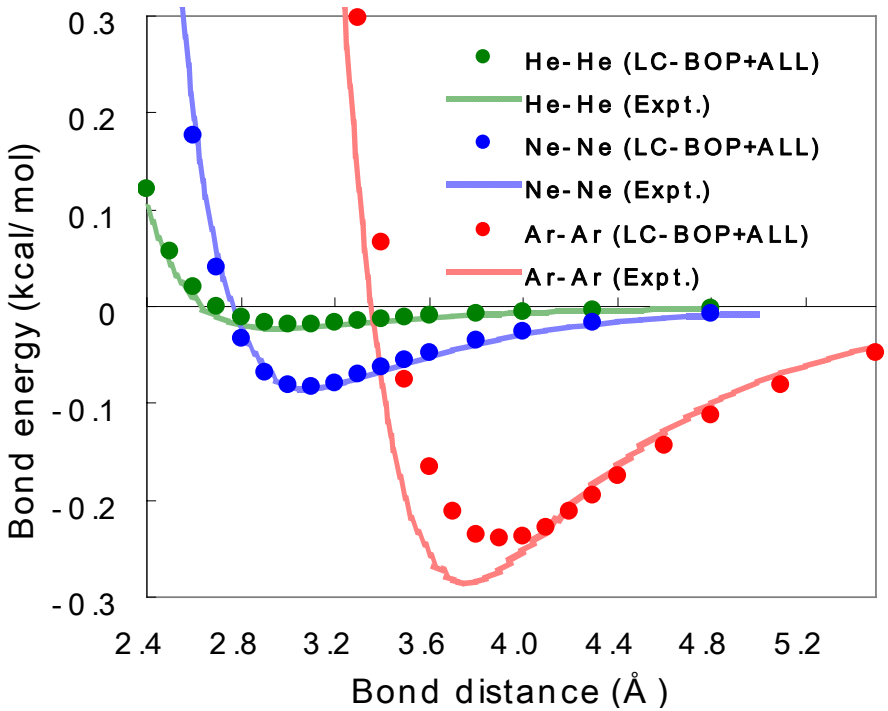
Potential Energy Curves of Ne-Ne



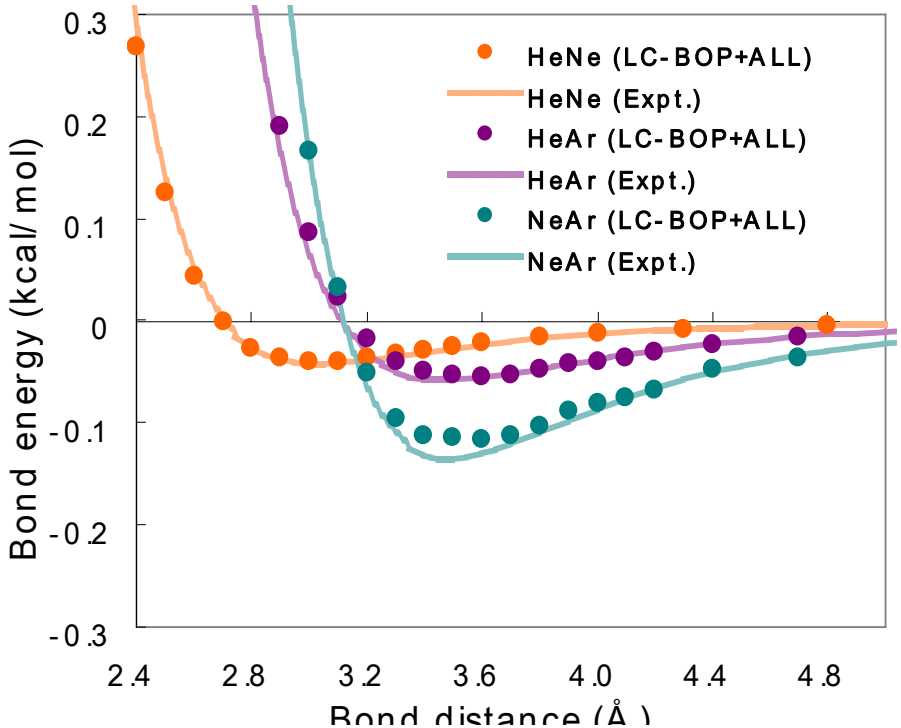
Applying the LC scheme to the exchange functionals leads to similar repulsive potentials

Potential Energy Curves of Rare-gas Dimers

Basis set: aug-cc-pVTZ, BSSE-corrected



Homo-nuclear

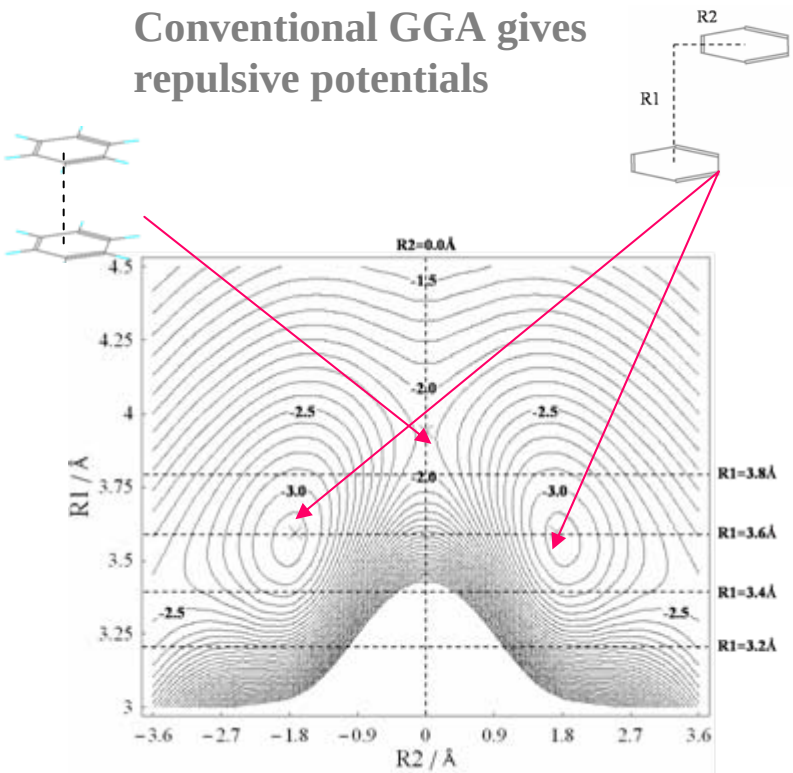


Hetero-nuclear

Van der Waals Interactions

Benzene Dimer

Conventional GGA gives repulsive potentials

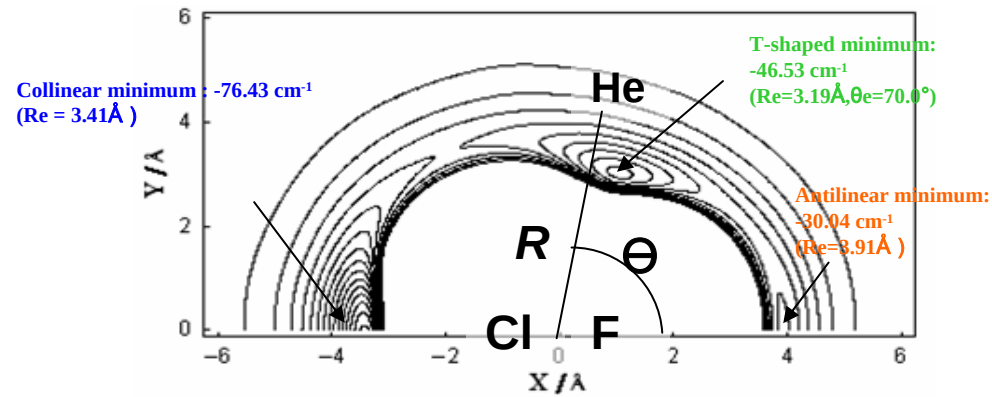


	D_e (kcal/mol)	R_1 (Å)	R_2 (Å)
Present/aug-cc-pVDZ	3.17		3.6
	1.7		
MP2/aug-cc-pVQZ	4.79		3.4

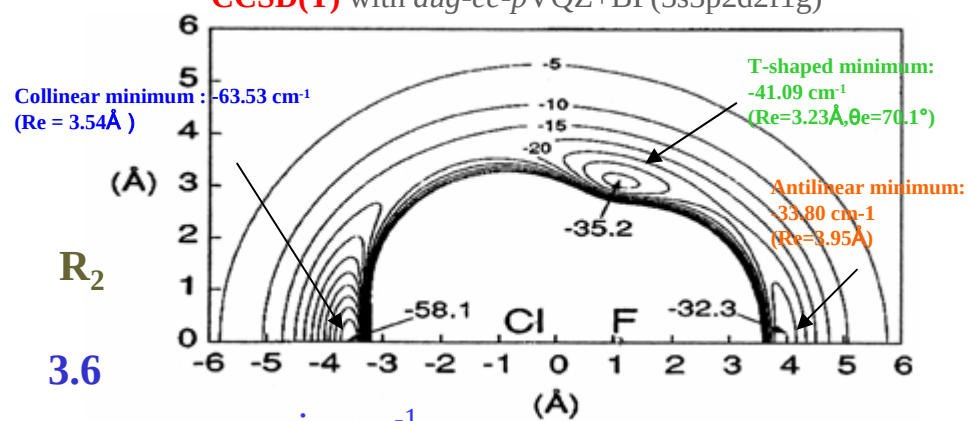
ClF…He

Conventional GGA gives only linear structures

Our DFT with *aug-cc-p*VQZ

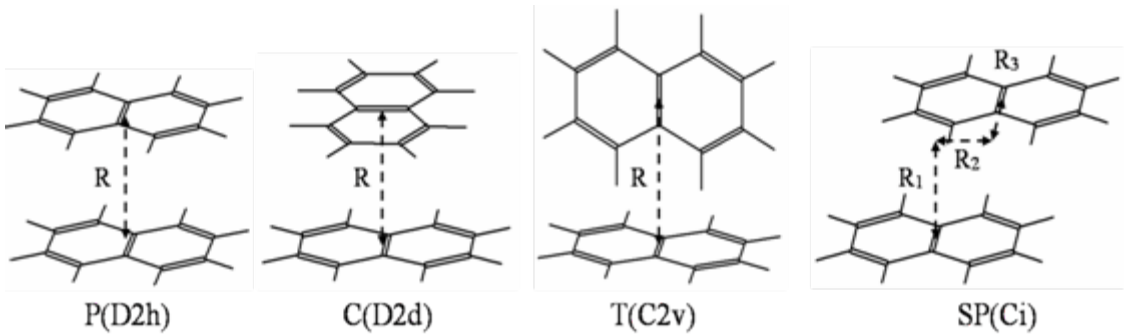


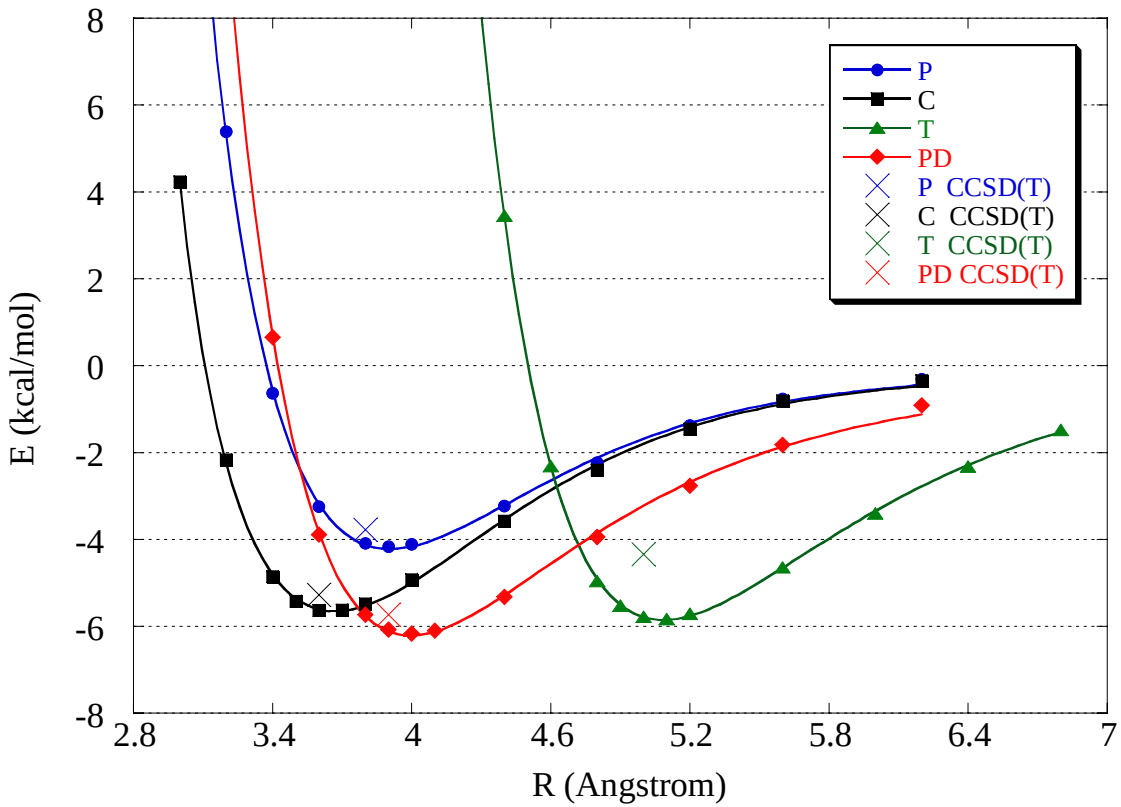
CCSD(T) with *aug-cc-p*VQZ+BF(3s3p2d2f1g)



in cm^{-1}

Naphthalene Dimer





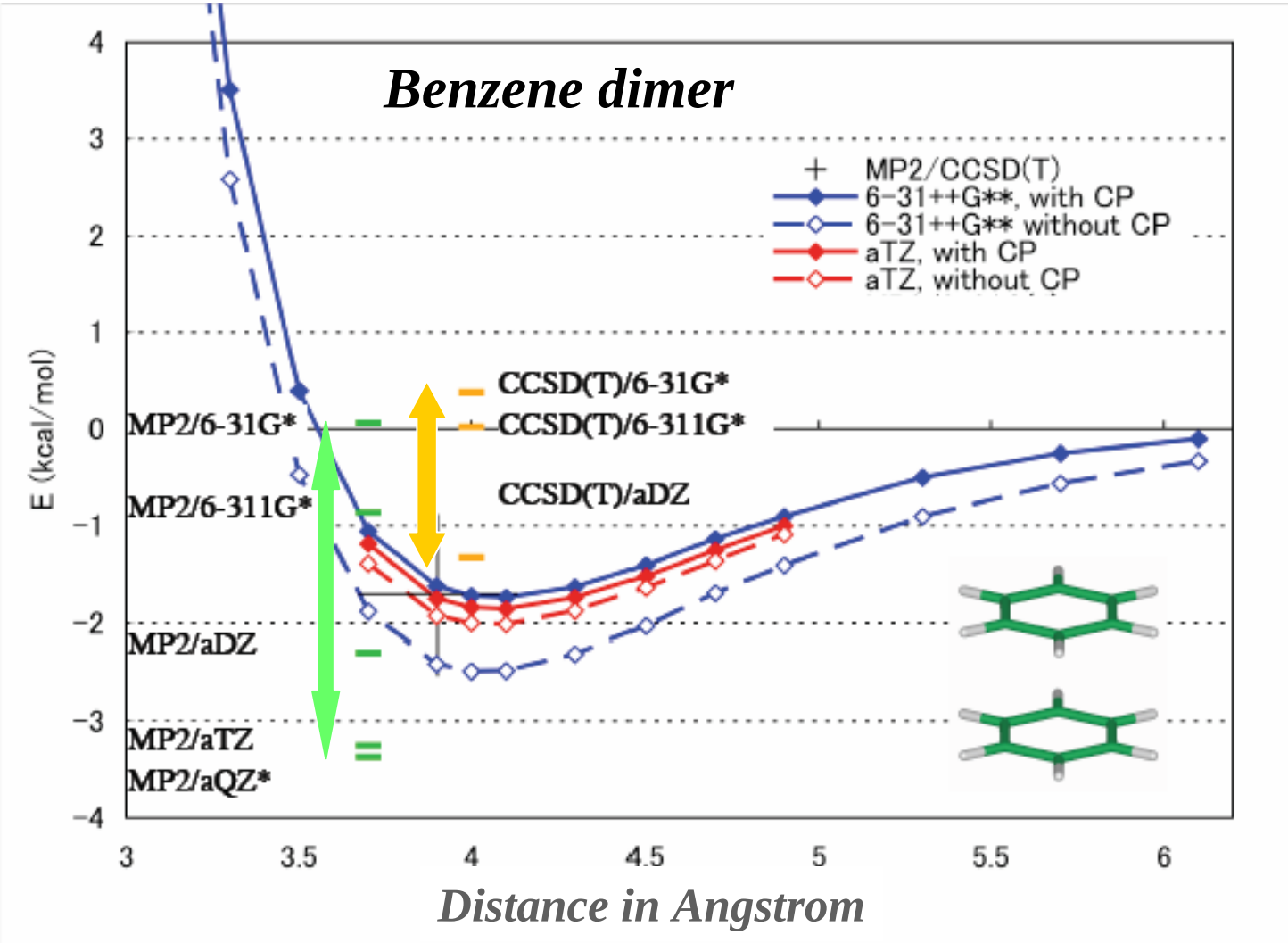
Our DFT 6.16 kcal/mol
CBS CCSD(T) 5.73 kcal/mol

Mean absolute percentage error (%)

	Present	Becke	MP2
Binding energies			
Dispersion	9.9	14.9	46.6
Dipole-induced dipole	5.6	11.8	8.6
Dipole-Dipole	9.2	8.5	9.5
Hydrogen bonded	7.1	8.6	2.4
Overall	7.5	11.7	19.2
Separations			
Dispersion	0.11	0.08	
Dipole-induced dipole	0.08	0.02	
Dipole-Dipole	0.05	0.05	
Hydrogen bonded	0.06	0.08	
Overall	0.08	0.06	

Becke & Johnson, JCP (2005) ：the model requires the polarizability of each monomer.

Basis set dependence



The calculated results have little dependence on the basis set used.

Basis Set Errors

Slow convergence of the partial wave expansion due to a consequence of the Coulomb-singularity of electron interaction

Electron cusp condition

$$\lim_{r_{12}\to 0}\left(\frac{\partial\Psi}{\partial r_{12}}\right)_{av} = \frac{1}{2}\Psi\left(r_{12}=0\right)$$

$$\Psi = \Psi\left(r_{12}=0\right)\left(1+\frac{1}{2}r_{12}+\cdots\right)$$

Expansion method

$$\frac{1}{2}r_{12}\Phi_0 = \sum_i C_i\Phi_i$$

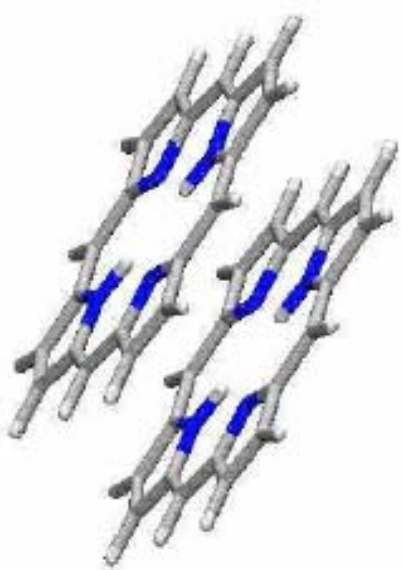
$$\left\langle\Phi_0\left|\frac{1}{r_{12}}\cdot r_{12}\right|\Phi_0\right\rangle == \sum_i C_i^2\left\langle\Phi_0\left|\frac{1}{r_{12}}\right|\Phi_i\right\rangle\left\langle\Phi_i\left|r_{12}\right|\Phi_0\right\rangle = 1$$

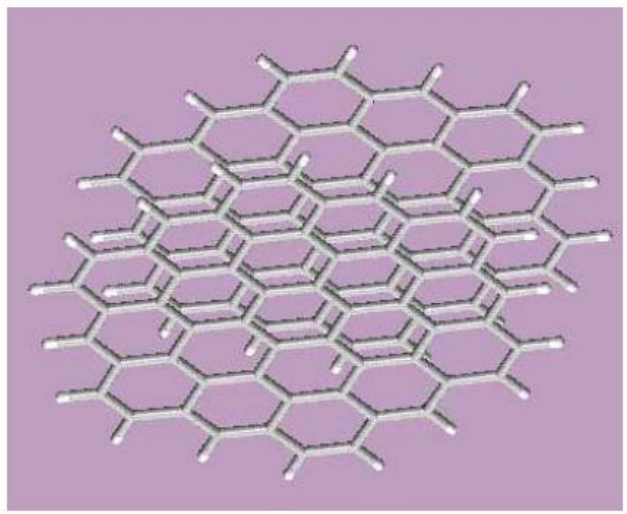
Error

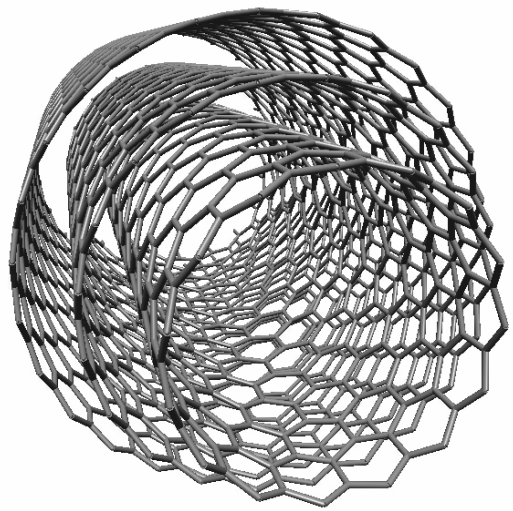
$$\Delta E \approx \left(1-\sum_i C_i^2\left\langle\Phi_0\left|\frac{1}{r_{12}}\right|\Phi_i\right\rangle\left\langle\Phi_i\left|r_{12}\right|\Phi_0\right\rangle\right)$$

Application to π-stacking energies

- Large planar aromatic systems (graphene sheets, porphyrins, DNA bases) are attracted by a considerable dispersion force.
- The dimerization energy is difficult to measure because of decomposition
- In the limit of large parallel sheets, the dispersion force diminishes as $1/r^4$, not as $1/r^6$

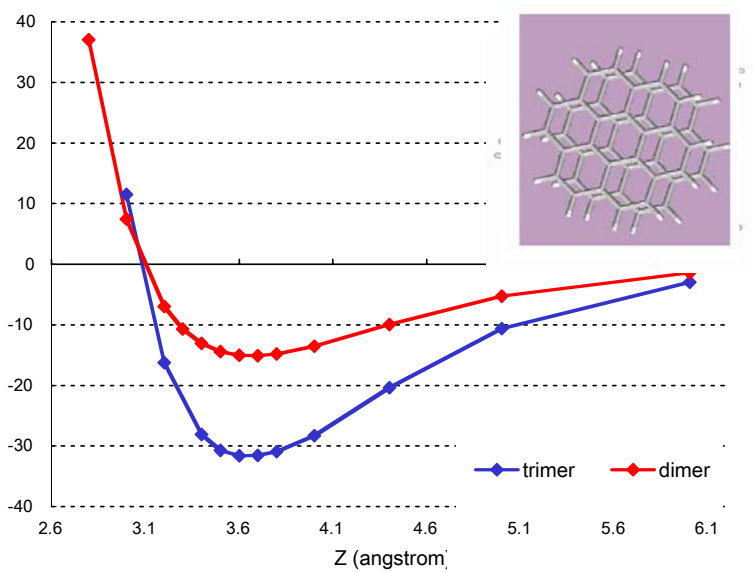




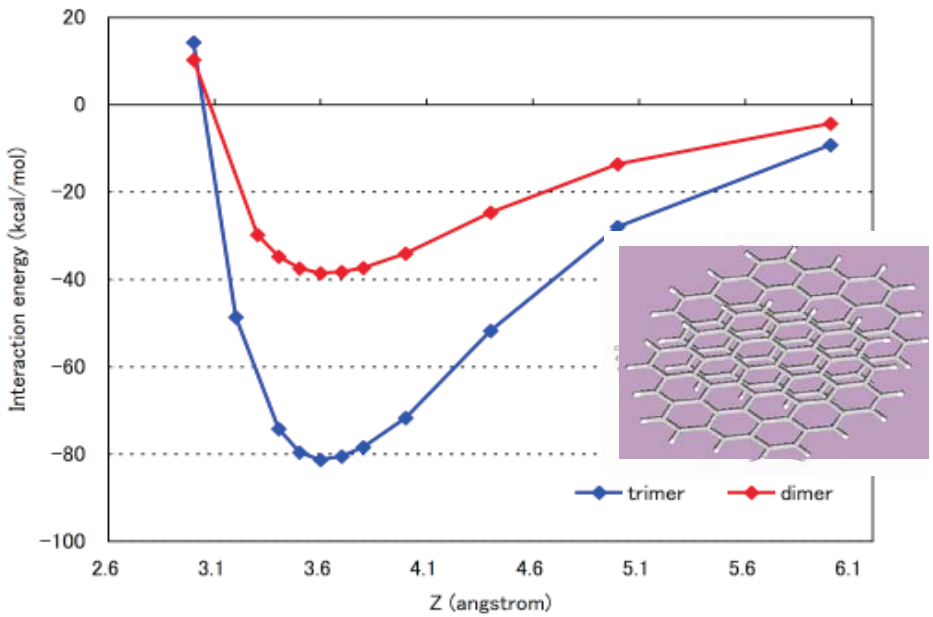


π-stacking energies

Coronene dimer ($C_{48}H_{24}$) & trimer ($C_{72}H_{36}$)



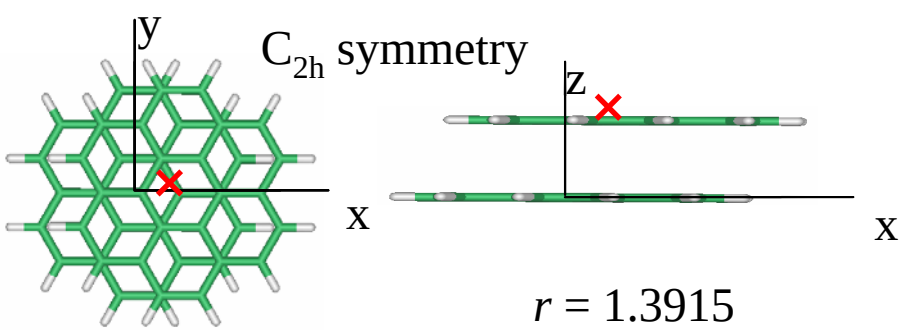
Circumcoronene dimer ($C_{108}H_{36}$) & trimer ($C_{162}H_{54}$)



LC-BOP/ 6-31++G**

1008 BFs

z = 3.6 Å
15.0 kcal/mol
(MP2 35.0 kcal/mol)



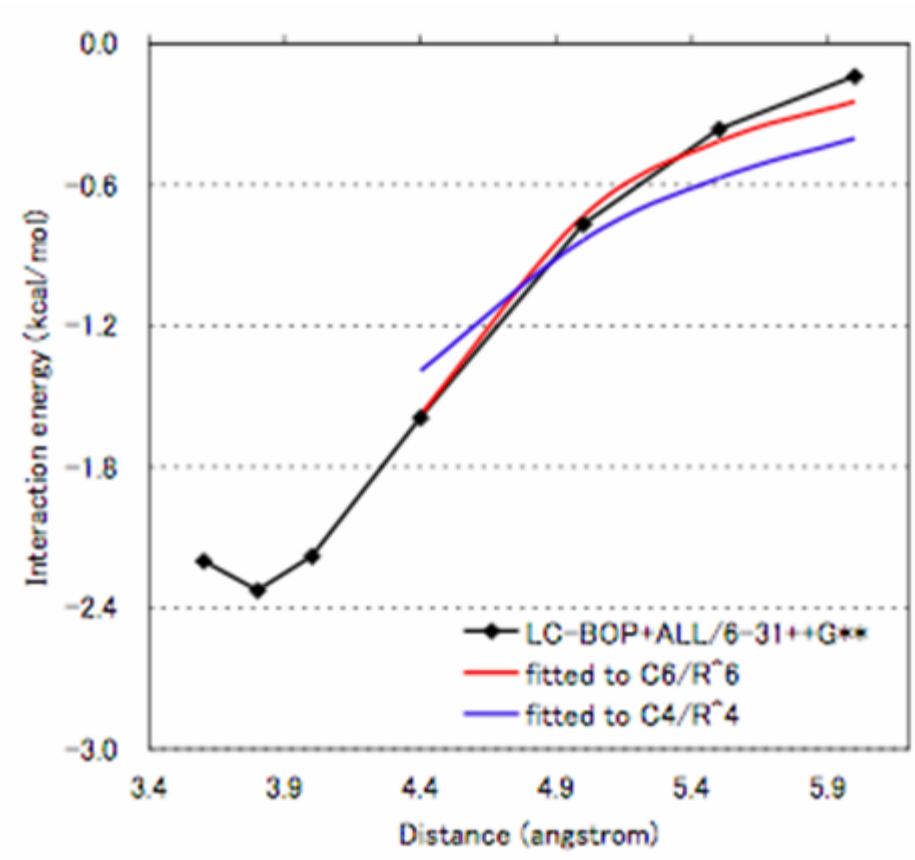
C_{2h} symmetry

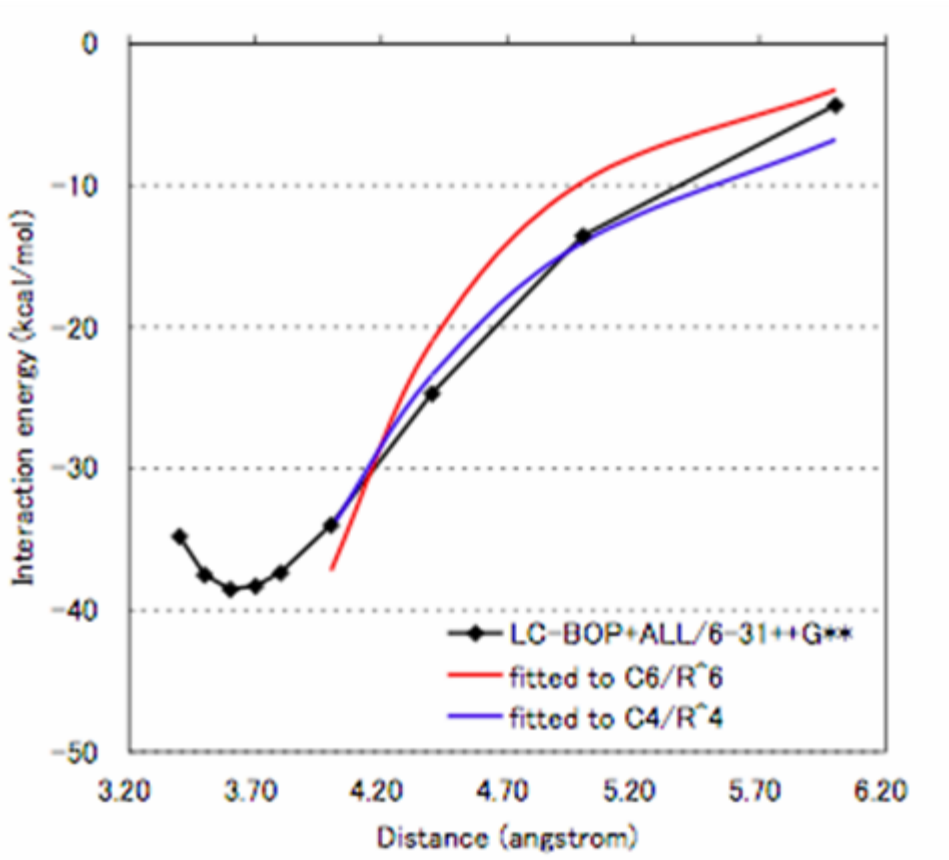
r = 1.3915

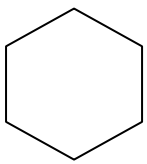
2160 BFs

z = 3.6 Å
38.5 kcal/mol
(MP2 98.5 kcal/mol)

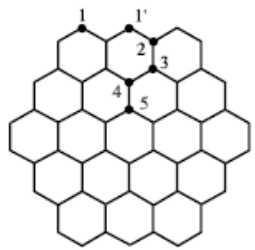
The dispersion force diminishes as $1/R^4$, not as $1/R^6$?



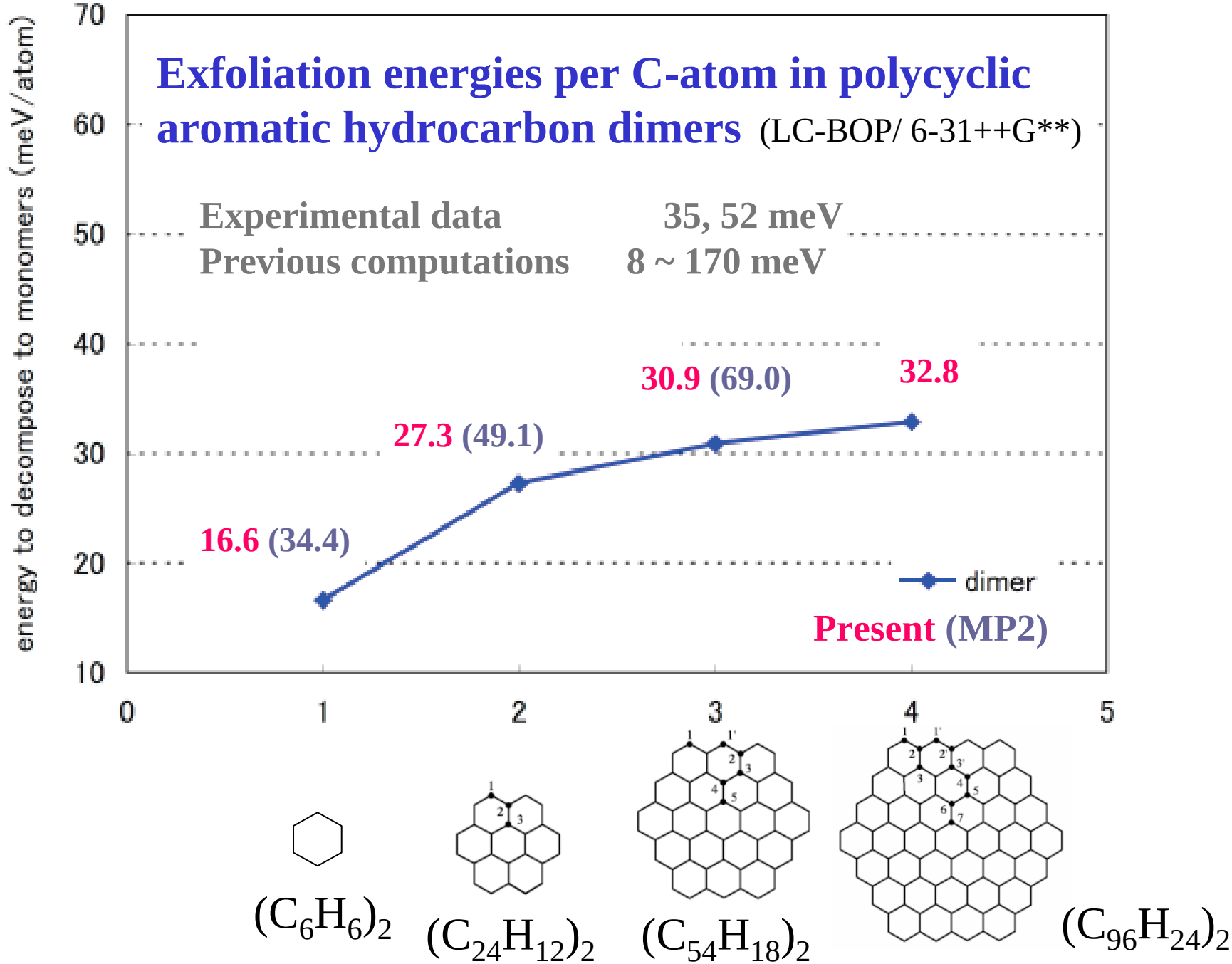




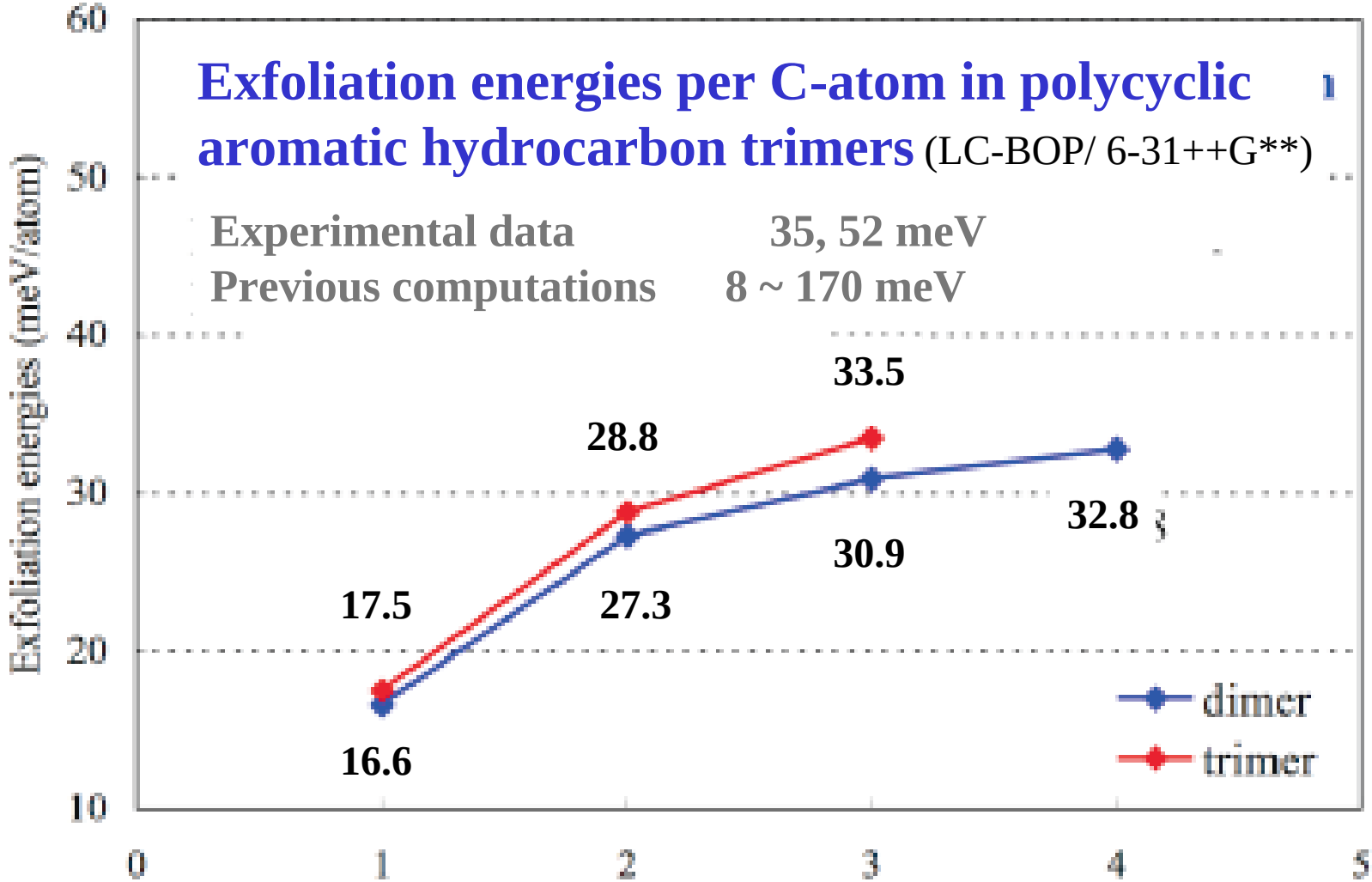
$(C_6H_6)_2$

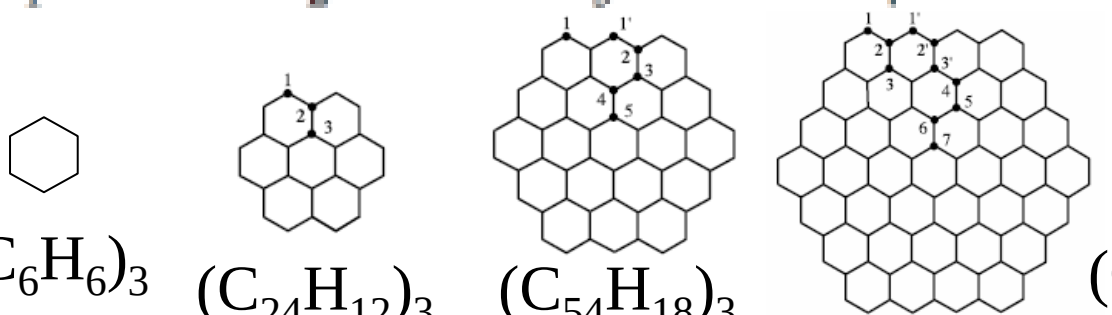


$(C_{54}H_{18})_2$


Exfoliation energies per C-atom in polycyclic aromatic hydrocarbon dimers (LC-BOP/ 6-31++G**)
Experimental data 35, 52 meV
Previous computations 8 ~ 170 meV
16.6 (34.4)
27.3 (49.1)
30.9 (69.0)
32.8
dimer
Present (MP2)
energy to decompose to monomers (meV/atom)
70
60
50
40
30
20
10
0
1
2
3
4
5
$(C_6H_6)_2$
$(C_{24}H_{12})_2$
$(C_{54}H_{18})_2$
$(C_{96}H_{24})_2$

Exfoliation energies per C-atom in polycyclic aromatic hydrocarbon trimers (LC-BOP/ 6-31++G**)

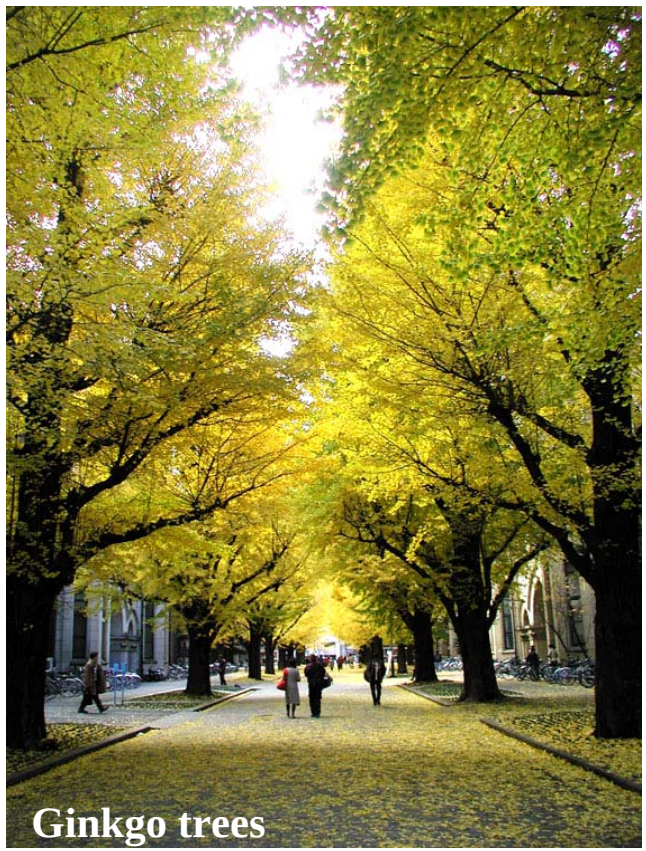




$(C_6H_6)_3$ $(C_{24}H_{12})_3$ $(C_{54}H_{18})_3$ $(C_{96}H_{24})_3$

Van der Waals Interactions

The proposed method is expected to be a promising alternative for calculating accurate van der Waals interactions in larger molecules, since this method requires much less computational cost compared to high-level *ab initio* wave function methods, such as CCSD(T).



Hirao Group (UT)

Thanks to Ministry of Education, Japan Science and Technology Agency (JST), past and present group, and many others.