SAND2013-6826C

Quantum Monte Carlo Simulation of van der Waals Systems

Anouar Benali[1], Luke Shulenburger[2], Nichols A. Romero[1], Jeongnim Kim[3], O. Anatole von Lilienfeld[1]

1- Argonne National Laboratory - Argonne, IL, USA
2- Sandia National Laboratories - Albuquerque, NM, USA
3- Oak Ridge National Laboratory - Oak Ridge, TN, USA

Contact: benali@anl.gov

Managed by UT-Battelle for the Department of Energy

QMCPACK Simulation Suite

- Jeongnim Kim.	ORNL (Formerly UIUC)
- David Ceperley	UIUC
- Luke Shulenburger	SNL
- Ken Esler	Stoneridge(Formerly UIUC)
- Miguel Morales	LLNL
- Jeremy McMinis	LLNL
- Nichols Romero	ANL
- Anouar Benali	ANL

[1] J. Kim *et al.* J. of Phy. - Conf. Series. (2012) vol. 402, no. 1, 012008
[2] K. P. Esler *et al.* Comp. in Sci. and Eng. (2012) vol. 14, no. 1, 40.
[3] J. Kim et al. "Qmcpack simulation suite."

IBM Blue Gene/Q
Mira @ Argonne National Laboratory

Mira: next-generation supercomputer

48 racks
1,024 nodes per rack
16 cores per node
64 threads per node
16GB memory/node
1.6GHz 16-way core processor
240 GB/s, 35 PB storage

786k cores
8.15 PF/s

BGQ - High Performance features

Quad FPU (QPU) →

DMA unit

List-based prefetcher

TM (Transactional Memory)

SE (Speculative Execution)

Wakeup-Unit

Scalable Atomic Operations

Instruction Extensions (QPX) to PowerISA

4-wide double precision FPU SIMD (BG/L,P are 2-wide) usable as:

- scalar FPU
- 4-wide FPU SIMD
- 2-wide complex arithmetic SIMD

Attached to AXU port of A2 core – A2 issues one instruction/cycle to AXU

8 concurrent floating point operations (FMA) + load +store

- 6 stage pipeline

Permute instructions to reorganize vector data

- supports a multitude of data alignments

4R/2W register file

- 32x32 bytes per thread

32B (256 bits) data path to/from L1 cache

Intrinsic:

TYPE: vector4double A;

Loads and stores

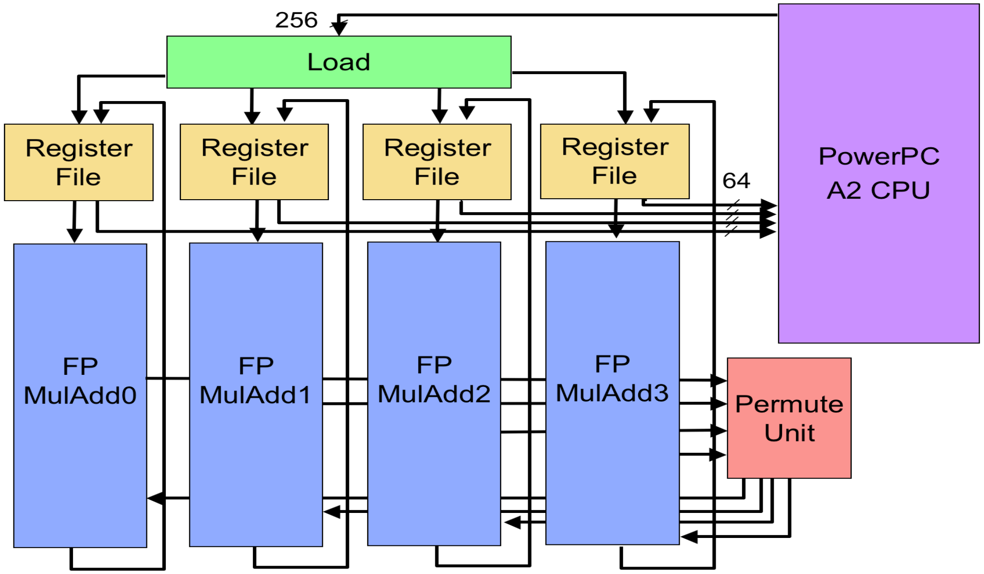
Unary operations

Binary operations

Multiply-add operations

Special functions

With four multiply-add units, the quad FPU can perform eight double-precision floating-point operations per cycle. The processor can execute up to eight double-precision floating-point operations, based on a fused multiply-add (FMA), along with an FP load and an FP store in a single cycle.



- The performance comes from the quad pipe Floating point unit.
- Each cycle, the quad FPU, can serve as a simple scalar FPU or a four wide SIMD FPU, or it can perform two complex arithmetic SIMD operations.
- All of these operations can be single or double precision.

Einspline

Eval_z:

Evaluation of spline coefficients (complex)

```
for ( i = 0; i < 64; i++ ){
   s = d[i];
   p = (double *)coefs[i];
   for ( n = 0; n < M - rem ; n = n + 8){
      double a0, a1,.., a7;
      double b0, b1, .., b7;

      a0 = v[n+0];
      //code
      a7 = v[n+7];

      b0 = p[n+0];
      //code
      b7 = p[n+7];

     //operations
      a0=a0+s*b0;
      ..
      v[n+0]=a0;
       ..
}}
```

Using QPX

```
for ( i = 0; i < 64; i++ ){
   s = d[i];

   p = (double *)coefs[i];
   vector4double t = { s, s, s, s };
   for ( n = 0; n < M - rem ; n = n + 8){
    vector4double f0, f1;
    vector4double g0, g1;



          g0 = vec_ld( j,    p );
          g1 = vec_ld( j+32, p );

          f0 = vec_ld( j,    v );
          f1 = vec_ld( j+32, v );

          f0 = vec_madd( t, g0, f0 );
          f1 = vec_madd( t, g1, f1 );

          vec_st( f0, j,    v );
          vec_st( f1, j+32, v );
}}
```

QMC Modelization

The many-body trial wavefunction

$$\Psi_T(R) = J(R)\Psi_{AS}(R) = e^{J_1+J_2+..}\sum_k^M C_k D_k^{\uparrow}(\phi) D_k^{\downarrow}(\phi)$$

Correlation (Jastrow)

$$J_1 = \sum_i^N \sum_l^{N_{ions}} u_1\left(\left|r_i - r_l\right|\right)$$

$$J_2 = \sum_{i \neq j}^N u_2\left(\left|r_i - r_j\right|\right)$$

Anti-symmetric function (Pauli principle)

$$D_k^{\sigma} = \begin{vmatrix} \phi_1(r_1) & \cdots & \phi_1(r_{N^\sigma}) \\ \vdots & \ddots & \vdots \\ \phi_{N^\sigma}(r_1) & \cdots & \phi_{N^\sigma}(r_{N^\sigma}) \end{vmatrix}$$

Single-particle orbitals

$$\phi_i = \sum_l^{l=N_b} C_l^i \Phi_l$$

Basis sets: molecular orbitals,
plane-wave, grid-based orbitals…

$$\Phi_l$$

Many methods of approximating the plane-wave-represented single-particle orbitals with polynomials:

-B-spline approximation in QMC, report significant reduction in time of calculation while maintaining plane-wave-level accuracy

Profiling

System:

- Ar Solid – 32 atoms – 256 electrons – B-splines representation of WF (1.9Gb) :
- 256 nodes – 32 threads – 2 Walkers per thread

Profile with original version of QMCPACK

```
Flat profile:
Total run time: 53min40
Each sample counts as 0.01 seconds.
%   cumulative   self     self     total
 time   seconds   seconds    calls  Ts/call  Ts/call  name
56.95  58369.57 58369.57                              .eval_multi_UBspline_3d_z_vgh
14.02  72738.82 14369.25                              .eval_multi_UBspline_3d_z
2.11  77918.51  2161.01                               SymmetricDTD
1.70  79663.07  1744.56                               EinsplineSetExtended::evaluate
```

Evaluation of spline, gradient and hessian coefficients (complex)

Evaluation of spline coefficients (complex)

71% of the application time spent in the Spline evaluation of the Wave Function

Optimization

-> 2 algorithms accessing memory horizontally or vertically over the {Points in space; minicube around the point}

Important reduction of the number of arithmetic operations

-> Complete rewriting of the functions with QPX intrinsics

- Increase of the number of floating point operations
- Reduction of the number of cycles per operation

-> Manual memory prefetching when possible

Increase the availability of data in the L1 cache

Profiling

Coefficients (type)

Eval_Z	Complex Double
Eval_D	Double
Eval_S	Float

Coefficients, Gradients, Hessian (type)

Eval_Z_VGH	Complex Double
Eval_D_VGH	Double
Eval_S_VGH	Float

Speed up	Eval_Z	Eval_D	Eval_S	Eval_Z_VGH	Eval_D_VGH	Eval_S_VGH
Algorithm B	0.38	0.81	0.39	1.59	0.93	1.62
Algorithm M	2.48	0.91	1.02	2.15	1.01	0.95

Algorithm (X) with QPX	3.94 (Algo. M)	1.08 (Algo. M)	1.26 (Algo. M)	7.62 (Algo. B)	1.58 (Algo.B)	1.31 (AlgoB)
QPX + Prefetch	4.25	1.23	1.81	-	-	-

Profiling

System:

- Ar Solid – 32 atoms – 256 electrons – Bsplines WF (1.9Gb) :
- 256 nodes – 32 threads – 2 Walkers per thread

Profile with QPX and Prefetch

Flat profile:

Total run time: 20min03

Each sample counts as 0.01 seconds.

Profile with Original Algoritm

Flat profile:

Total run time: 53min40

Each sample counts as 0.01 seconds.

QPX and Prefetch % time	QPX and Prefetch cumulative seconds	QPX and Prefetch self seconds		Original Algoritm % time	Original Algoritm cumulative seconds	Original Algoritm self seconds
14.08	5380.43	5380.43	**.eval_multi_UBspline_3d_z_vgh**	**56.95**	58369.57	58369.5
8.25	12270.83	3152.52	**.eval_multi_UBspline_3d_z**	**14.02**	72738.82	14369.25
5.68	14441.45	2170.62	.SymmetricDTD	2.11	77918.51	2161.01
4.85	16292.97	1851.52	EinsplineSetExtended::evaluate	1.70	79663.07	1744.5

Total run time Speedup of 2.68 times

HPM PROFILING

Original Code

```
27.644.290.379.027   All XU Instruction
22.786.190.220.714   All AXU Instruction
43.043.218.198.088   FP Operations Group 1

Derived metrics for code block "mpiAll" averaged
over process(es) on node <0,0,0,0,0>:
Instruction mix:  FPU = 45.18 %,  FXU = 54.82 %
Instructions per cycle completed per core =
0.6138
Per cent of max issue rate per core = 33.65 %
Total weighted GFlops for this node = 13.412
Loads that hit in L1 d-cache =  94.03 %
                  L1P buffer =   5.36 %
                  L2 cache   =   0.35 %
                  DDR        =   0.26 %
DDR traffic for the node: ld = 1.508, st =
0.540, total = 2.049 (Bytes/cycle)
```

Percentage of peak= 6.55%

BGQ optimized Code

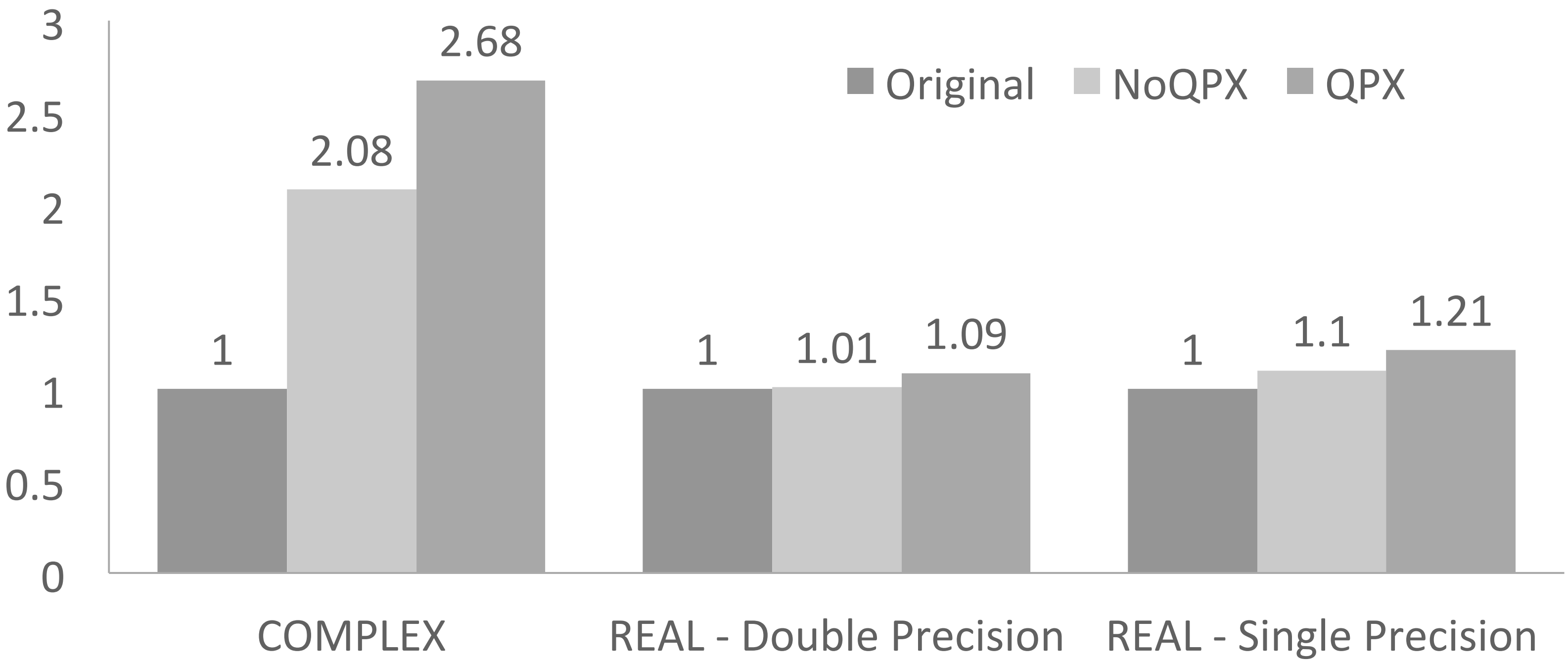
```
 8.581.366.867.332   All XU Instruction
 4.896.512.230.816   All AXU Instruction
13.017.533.928.058   FP Operations Group 1

Derived metrics for code block "mpiAll" averaged
over process(es) on node <0,0,0,0,0>:
Instruction mix:  FPU = 36.33 %,  FXU = 63.67 %
Instructions per cycle completed per core =
0.4417
Per cent of max issue rate per core = 28.12 %
Total weighted GFlops for this node = 10.922
Loads that hit in L1 d-cache =  88.60 %
                  L1P buffer =   5.92 %
                  L2 cache   =   4.50 %
                  DDR        =   0.98 %
DDR traffic for the node: ld = 3.503, st =
1.101, total = 4.604 (Bytes/cycle)
```

Percentage of peak= 5.33%

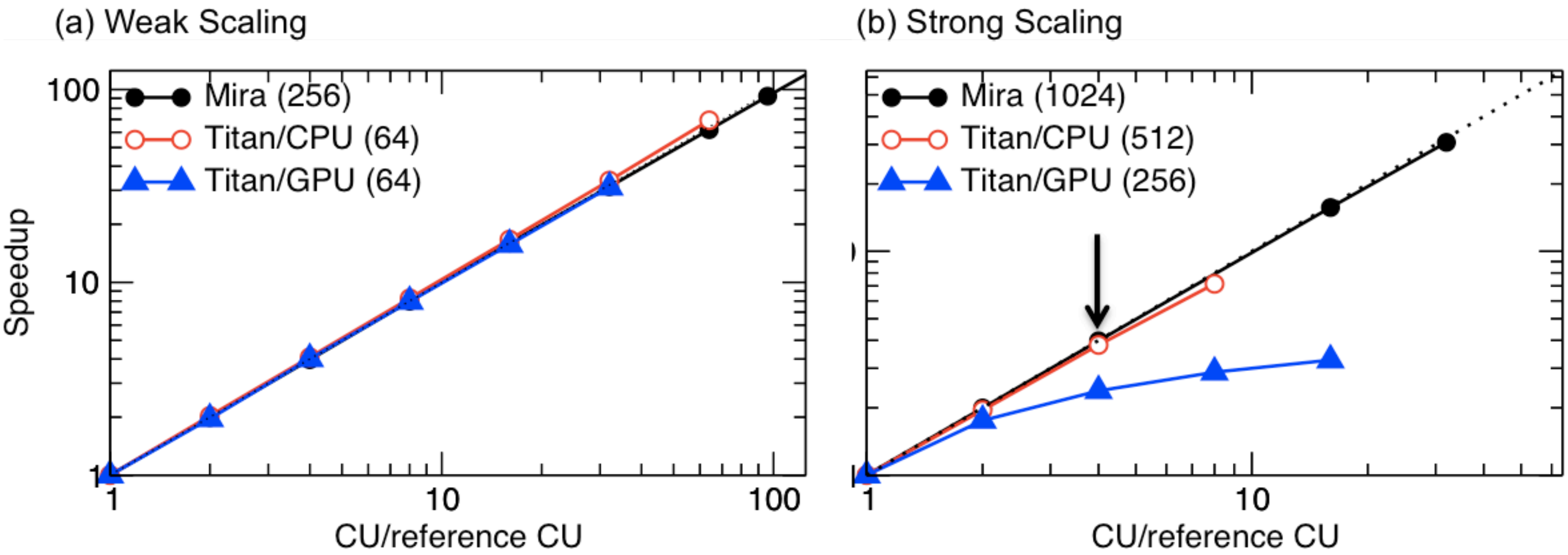
Total run time Speedup of 2.68 times

QMCPACK - performance on Blue Gene/Q



Application speedup using QPX and prefetching is **2.68** folds from original Algorithm.

Overall Scaling on LCF Architectures



Applications on van der Waals dominated systems

Objectives

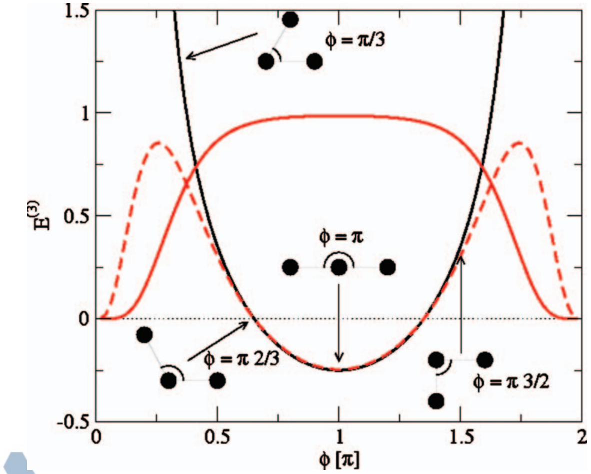
Van der Waals forces are important
-> Noble gases are proto-typical! We use Ar as a case of principle.

- London (C_6/R^6) widely used in force fields (Lennard-Jones tail)
- Axilrod-Teller-Muto (C_9/R^9) is 3-body analogue

Dispersion Coefficients (C_6, C_9)

London[1,2] $W(R)=-C_6/R^6$

Axilrod-Teller-Muto[3] $W(R)=C_9\,(3\cos[\varphi]\cos[\varphi]\cos[\varphi]+1)/R^9$



[1,] W. Heitler and F. London, Z. Phys. **44**, 455 (1927)
[2] R. Eisenschitz and F. London, Z. Phys. **60**, 491 (1930)
[3] B. M. Axilrod and E. Teller, J. Chem. Phys. **11,** 299 (1943)

Objectives

Van der Waals forces are important

-> Noble gases are proto-typical! We use Ar as a case of principle.

- London (C_6/R^6) widely used in force fields (Lennard-Jones tail)
- Axilrod-Teller-Muto (C_9/R^9) is 3-body analogue

-> Argon EOS

- Evaluation of the 2body, 3 body and MBC to the crystal solid (in progress)

-> We apply the method to Ellipticine and DNA

- Binding Energy of the drug Ellipticine to DNA

QMC Modelization

- We Solve the many-body Schrodinger equation and we express the wavefunction as follow;

$$\Psi_T\left(x_1,x_1,\ldots,x_N\right) = J\left(x_1,x_1,\ldots,x_N\right)\Psi_{AS}\left(x_1,x_1,\ldots,x_N\right)$$

DFT Calculation (LDA funtional)

Trial Wavefunction (PWSCF)

$$J_1(\vec{R}) = \prod_{ia}\exp\left[\sum_k \left(b_{ak} r_{ia} + c_{ak}\right) v_{ak}\left(r_{ia}\right)\right]$$

$$J_2(\vec{R}) = \prod_{i<j}\exp\left[\sum_k \left(b_k r_{ij} + c_k\right) v_k\left(r_{ij}\right)\right]$$

One-Body + Two-body Jastrow

Optimization of the factors (convergence using VMC)

$$E_{VMC} = \min_{\alpha}\left\langle \Psi_T\left(\vec{R};\alpha\right)\middle|\hat{H}\middle|\Psi_T\left(\vec{R};\alpha\right)\right\rangle$$

Variational Monte Carlo

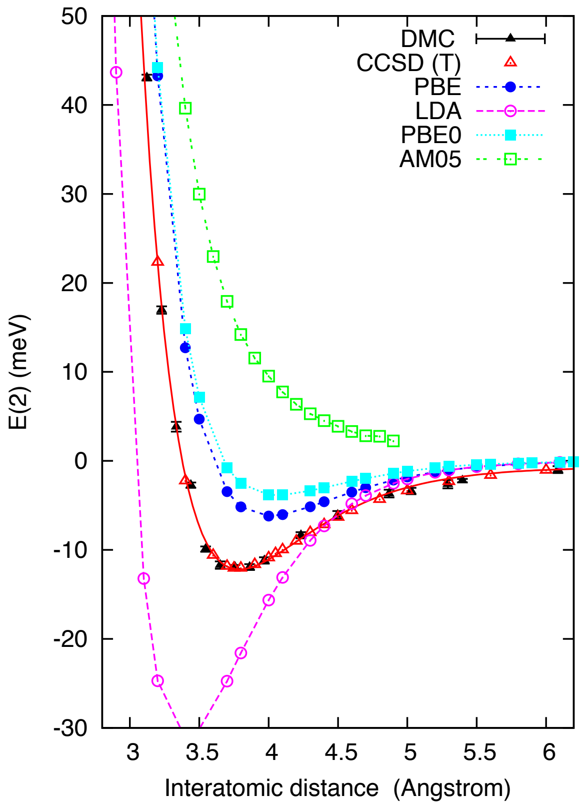
New Trial Wavefunction

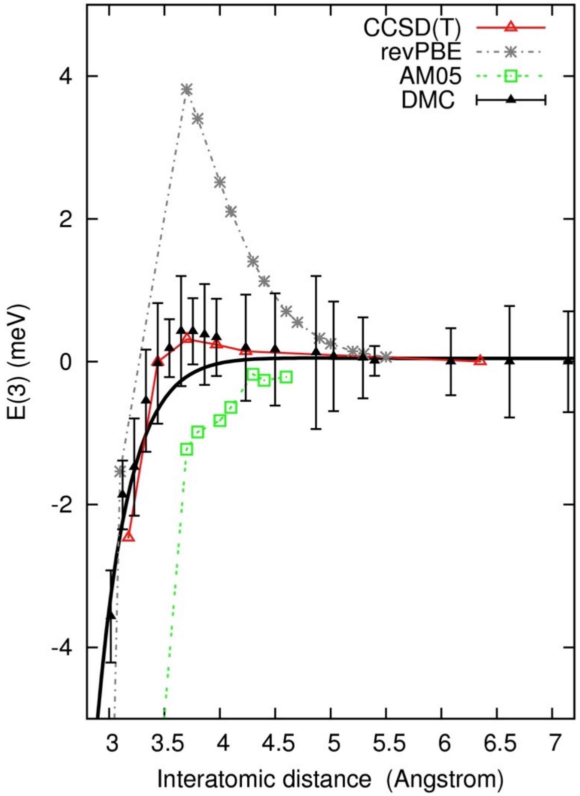
$$E_{DMC} = \left\langle \phi_0 \middle| \hat{H} \middle| \Psi_T \right\rangle, \phi_0 = \lim_{\beta\to\infty} \exp^{-\beta\hat{H}} \Psi_T$$

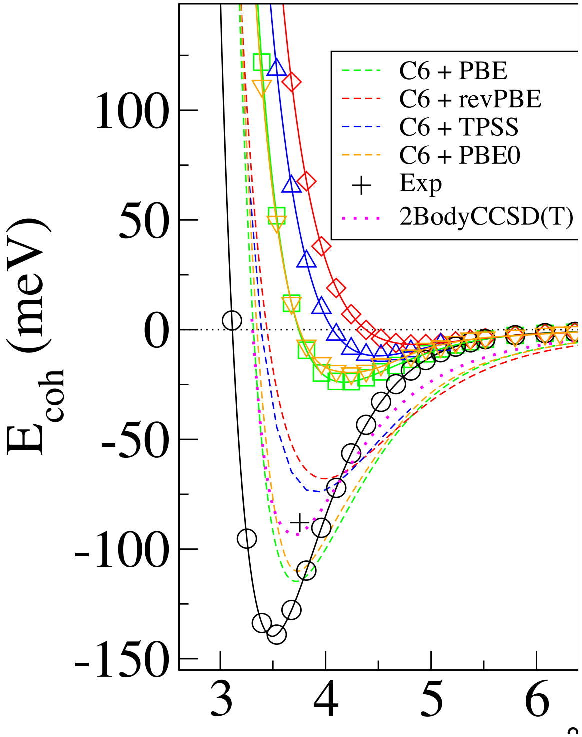
Diffusion Monte Carlo

- Solid: Corrections to finite sizes effects, Kinetic and MPC, twists averaging

Argon Systems







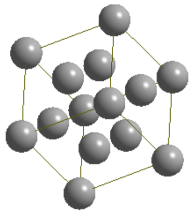
	$d_{Ar\text{-}Ar}$ (Å)	$E_{(2)}$ (meV)	$E_{(3)}$ (meV)	C_6	C_9
This Work	3.757	-12.232 ±0.987	0.289 ±0.567	63.1	517.6
Ref.	3.76[a]	-12.3[a]	0.3[b]	64.3[c]	518[c]

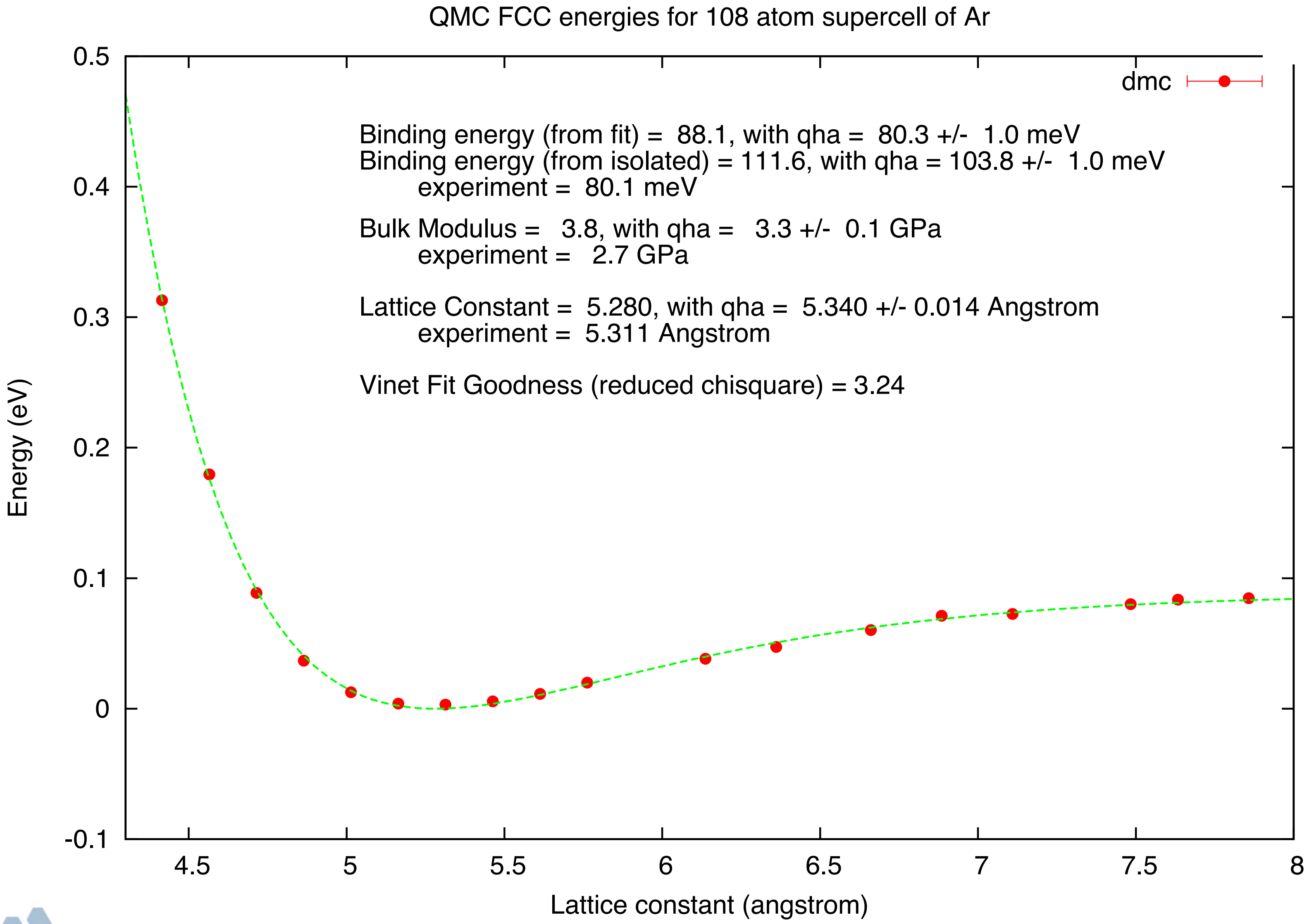
[a] Experimental, P. R. Herman, P. E. LaRocque, and B. P. Stoicheff, J. Chem. Phys. **88**, 4535 (1988)

[b] R. Podeszwa and K. Szalewicz, J. Chem. Phys. **126**, 194102 (2007)

[c] O. A. von Lilienfeld and A. Tkatchenko, J. Chem. Phys. **132**, 234109 (2010)

Argon Solid





QMC FCC energies for 108 atom supercell of Ar

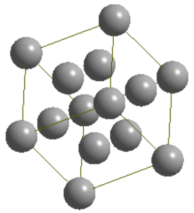
Binding energy (from fit) = 88.1, with qha = 80.3 +/- 1.0 meV
Binding energy (from isolated) = 111.6, with qha = 103.8 +/- 1.0 meV
experiment = 80.1 meV

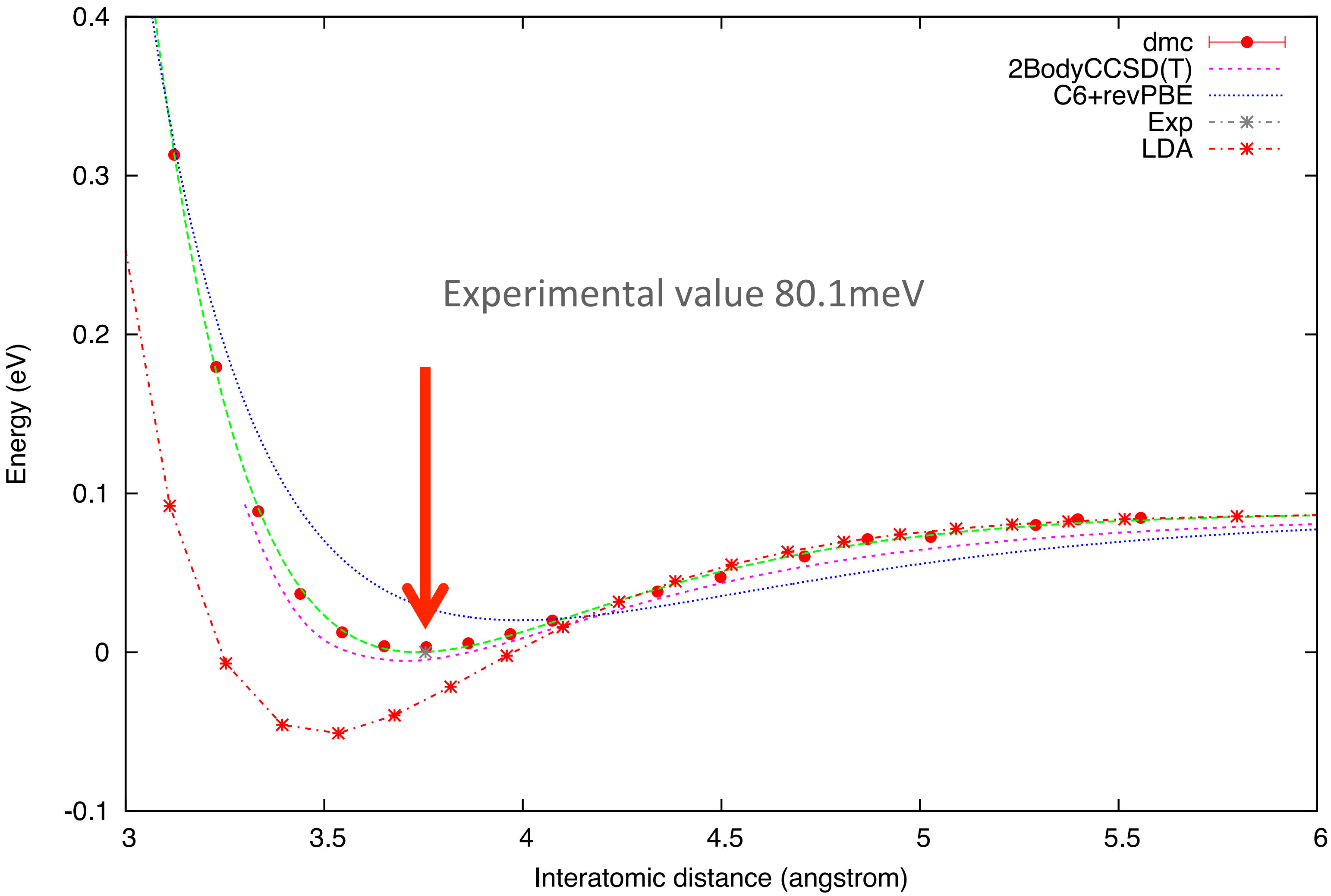
Bulk Modulus = 3.8, with qha = 3.3 +/- 0.1 GPa
experiment = 2.7 GPa

Lattice Constant = 5.280, with qha = 5.340 +/- 0.014 Angstrom
experiment = 5.311 Angstrom

Vinet Fit Goodness (reduced chisquare) = 3.24

Argon Solid

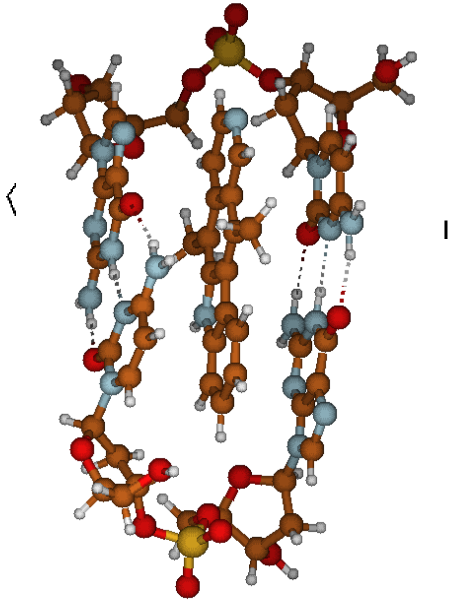




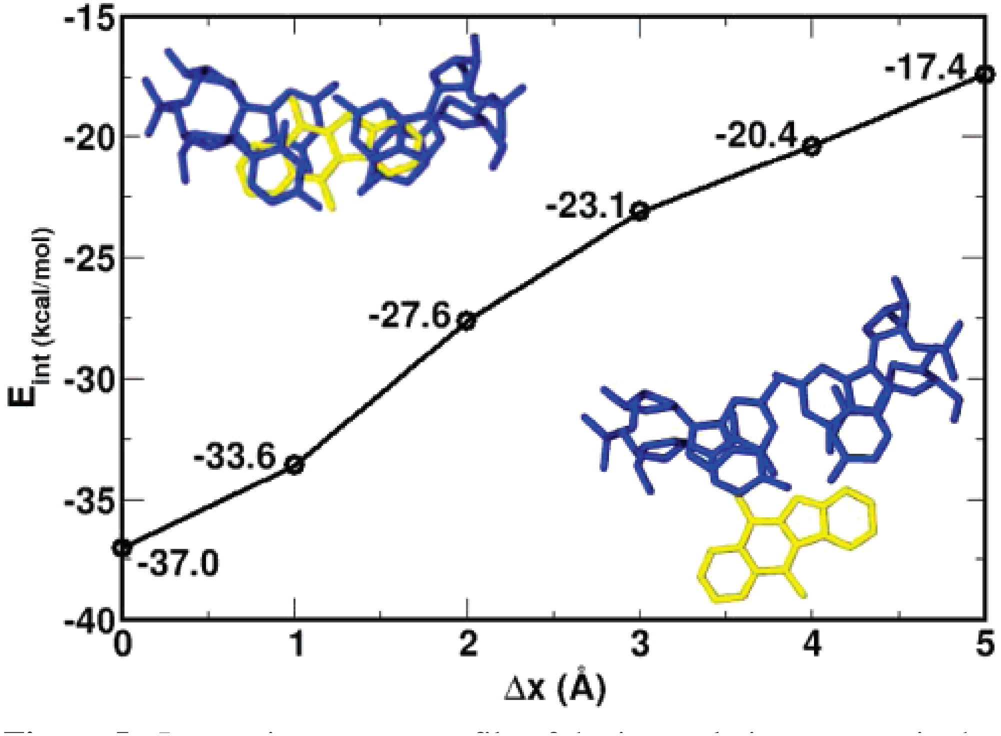
Experimental value 80.1meV

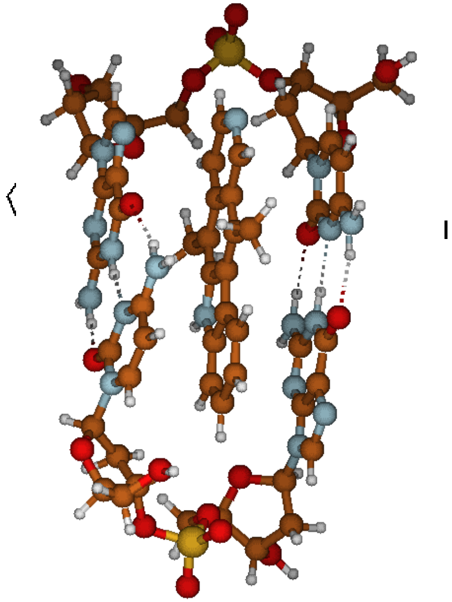
Ellipticine

- Ellipticine is a planar polycyclic aromatic molecule
- Bind to DNA by non-covalent pi-pi-stacking with the nucleic acid Watson-Crick base pairs
- Binding energy is directly correlated to biological activity of the molecule in cancer treatment

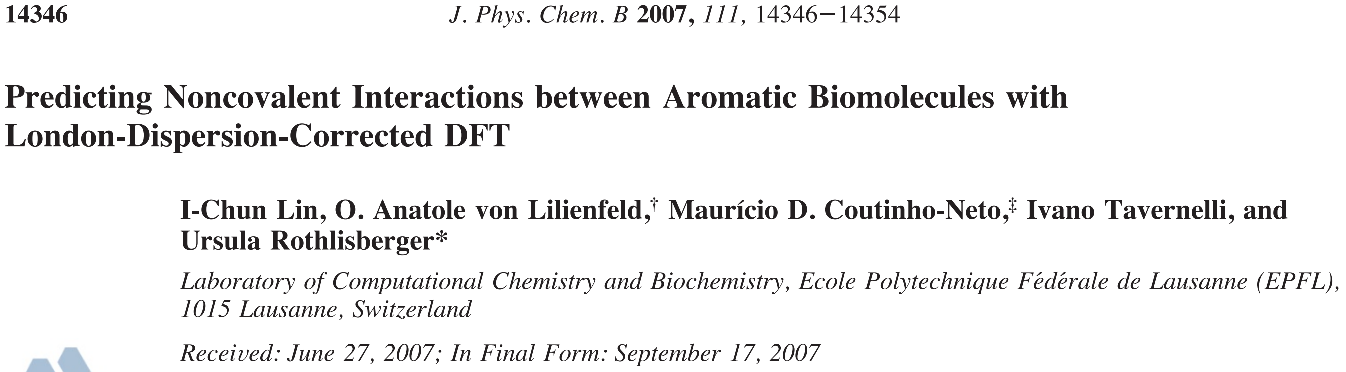


Ellipticine





Dispersion-corrected atom-centered potentials (DCACPs)



14346 *J. Phys. Chem. B* **2007,** *111,* 14346−14354

Predicting Noncovalent Interactions between Aromatic Biomolecules with London-Dispersion-Corrected DFT

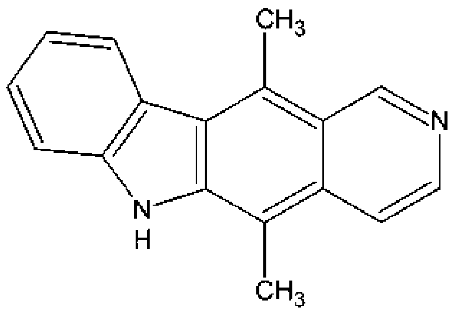
I-Chun Lin, O. Anatole von Lilienfeld,[†] Maurício D. Coutinho-Neto,[‡] Ivano Tavernelli, and Ursula Rothlisberger*

Laboratory of Computational Chemistry and Biochemistry, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

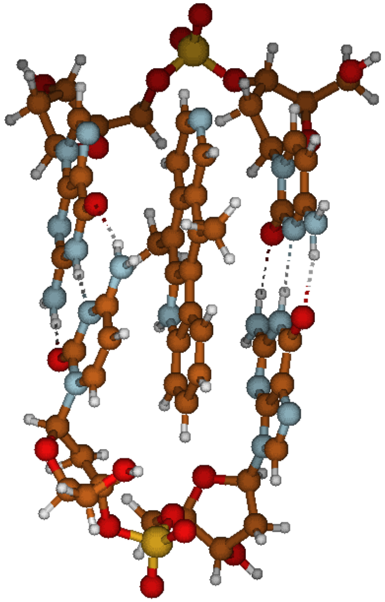
Received: June 27, 2007; In Final Form: September 17, 2007

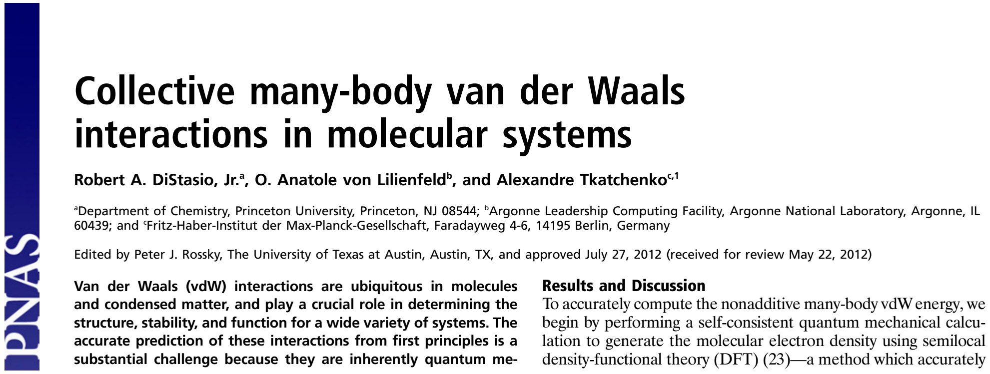
Ellipticine

Level of theory	ΔE_{Bind} (Kcal/mol)
DFT[1]	+5.2
vdW-TS[1]	-46.6
vdW-TB[1]	-39.1
PBE-D3/QZVP[2]	-35.68 (D2) ; -32.84 (D2+D3)
PBE-NL/QZVP[2]	-39.11 (D2) ; -36.27 (D2+D3)
dDsC-PBE/QZ4P[2]	-40.91 (D2) ; -38.07 (D2+D3)
vdW-MB[1]	-34



ellipticine





PNAS

Collective many-body van der Waals interactions in molecular systems

Robert A. DiStasio, Jr.[a], O. Anatole von Lilienfeld[b], and Alexandre Tkatchenko[c,1]

[a]Department of Chemistry, Princeton University, Princeton, NJ 08544; [b]Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439; and [c]Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

Edited by Peter J. Rossky, The University of Texas at Austin, Austin, TX, and approved July 27, 2012 (received for review May 22, 2012)

Van der Waals (vdW) interactions are ubiquitous in molecules and condensed matter, and play a crucial role in determining the structure, stability, and function for a wide variety of systems. The accurate prediction of these interactions from first principles is a substantial challenge because they are inherently quantum me-

Results and Discussion

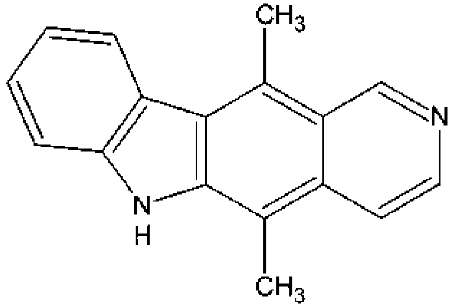
To accurately compute the nonadditive many-body vdW energy, we begin by performing a self-consistent quantum mechanical calculation to generate the molecular electron density using semilocal density-functional theory (DFT) (23)—a method which accurately

[1] R. DiStasio, O. A. von Lilienfeld, A. Tkatchenko, PNAS (2012)

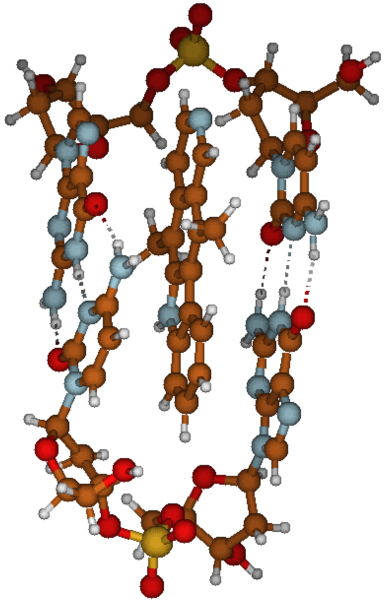
[2] S. Grimme – Private communication

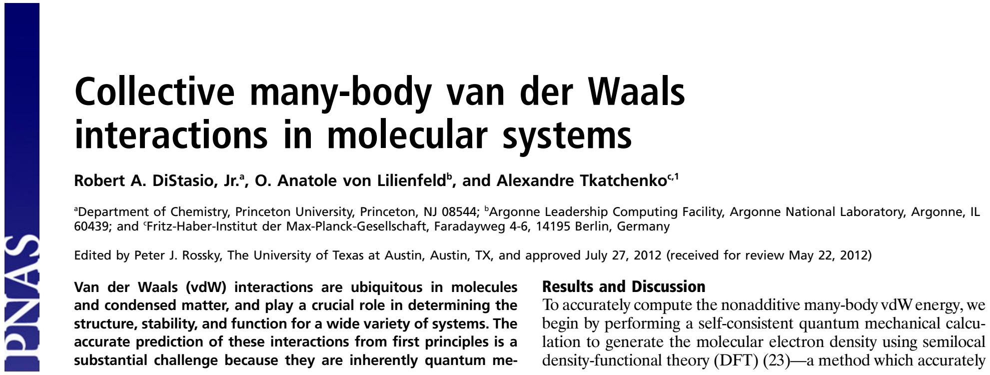
Ellipticine

Level of theory	ΔE_{Bind} (Kcal/mol)
DFT[1]	+5.2
vdW-TS[1]	-46.6
vdW-TB[1]	-39.1
PBE-D3/QZVP[2]	-35.68 (D2) ; -32.84 (D2+D3)
PBE-NL/QZVP[2]	-39.11 (D2) ; -36.27 (D2+D3)
dDsC-PBE/QZ4P[2]	-40.91 (D2) ; -38.07 (D2+D3)
vdW-MB[1]	-34
DMC	-33.6 ± 0.9



ellipticine





Collective many-body van der Waals interactions in molecular systems

Robert A. DiStasio, Jr.[a], O. Anatole von Lilienfeld[b], and Alexandre Tkatchenko[c,1]

[a]Department of Chemistry, Princeton University, Princeton, NJ 08544; [b]Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439; and [c]Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

Edited by Peter J. Rossky, The University of Texas at Austin, Austin, TX, and approved July 27, 2012 (received for review May 22, 2012)

Van der Waals (vdW) interactions are ubiquitous in molecules and condensed matter, and play a crucial role in determining the structure, stability, and function for a wide variety of systems. The accurate prediction of these interactions from first principles is a substantial challenge because they are inherently quantum me-

Results and Discussion

To accurately compute the nonadditive many-body vdW energy, we begin by performing a self-consistent quantum mechanical calculation to generate the molecular electron density using semilocal density-functional theory (DFT) (23)—a method which accurately

[1] R. DiStasio, O. A. von Lilienfeld, A. Tkatchenko, PNAS (2012)

[2] S. Grimme – Private communication

Ellipticine

Level of theory	$\Delta E^{A:T}_{B-A}$	$\Delta E^{C:G}_{B-A}$
DFT	+4.2	+1.9
vdW-TS	+2.5	−3.7
vdW-TB	+2.6	−3.5
vdW-MB	−0.1	−8.2

(Left) Binding energies (ΔE_{bind}) for the DNA–ellipticine complex in kcal/mol. (Right) Relative conformational energies of A-DNA and B-DNA ($\Delta E_{B-A} = E_B - E_A$) consisting of pure adenine–thymine (A:T) and cytosine–guanine (C:G) sequences in kcal/mol per bp. All DFT calculations were performed using the PBE functional (37).

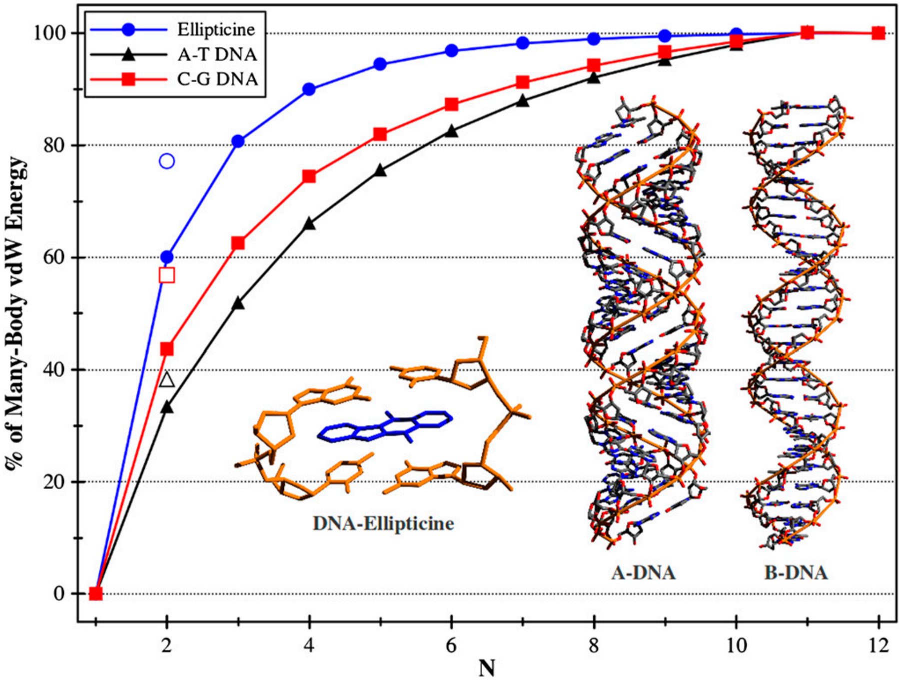
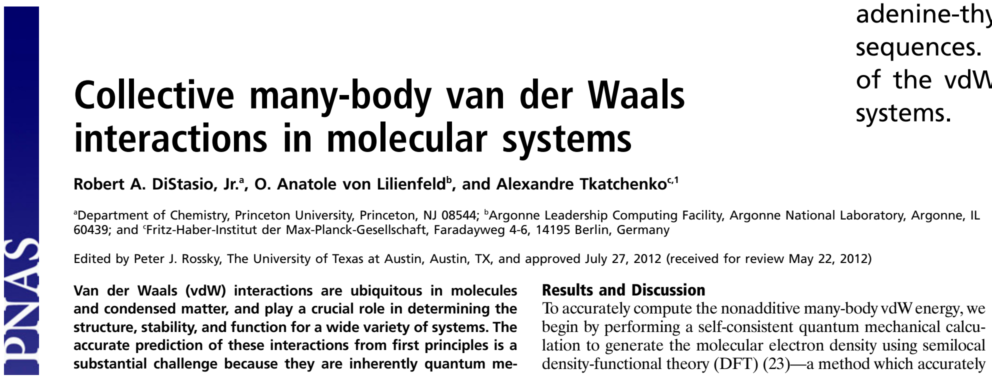


Fig. 2. Percentagewise convergence of the individual vdW-*N*B contributions with respect to the vdW-MB energy. Displayed cases include the binding energy of the DNA-ellipticine complex (blue circles) and the relative binding energies of a single base pair in A-DNA and B-DNA consisting of pure adenine-thymine (black triangles) and pure cytosine-guanine (red squares) sequences. The unfilled markers at $N = 2$ correspond to the predictions of the vdW-TB effective pairwise model for each of the aforementioned systems.



Collective many-body van der Waals interactions in molecular systems

Robert A. DiStasio, Jr.[a], O. Anatole von Lilienfeld[b], and Alexandre Tkatchenko[c,1]

[a]Department of Chemistry, Princeton University, Princeton, NJ 08544; [b]Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439; and [c]Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

Edited by Peter J. Rossky, The University of Texas at Austin, Austin, TX, and approved July 27, 2012 (received for review May 22, 2012)

Van der Waals (vdW) interactions are ubiquitous in molecules and condensed matter, and play a crucial role in determining the structure, stability, and function for a wide variety of systems. The accurate prediction of these interactions from first principles is a substantial challenge because they are inherently quantum me-

Results and Discussion

To accurately compute the nonadditive many-body vdW energy, we begin by performing a self-consistent quantum mechanical calculation to generate the molecular electron density using semilocal density-functional theory (DFT) (23)—a method which accurately

Conclusion

- QMC is a great method but expensive. Requires tuning on supercomputers.

- Rare gas study confirms that the method goes below the Kcal/mol accuracy and reproduces CCSD (T) results.

- Van der Waals corrected DFT methods have improved greatly by the inclusion of Manybody effects. However, are still predicting energies widely spread. With a QMC benchmark, the order of the manybody-vdw correction can be controlled to reproduce DMC energies.

Acknowledgments

- Collaborators: Luke Shulenburger, Nichols A. Romero, Jeongnim Kim and O. Anatole von Lilienfeld

- This research used resources of the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357.

- Sandia National Laboratories is a multiprogram laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-AC04-94AL85000.

INCITE: Innovative and Novel Computational Impact on Theory and Experiment

INCITE promotes transformational advances in science and technology through large allocations of computer time, supporting resources, and data storage at the Argonne and Oak Ridge Leadership Computing Facilities (LCFs) for computationally intensive, large-scale research projects.

Oak Ridge National Laboratory

Argonne National Laboratory

Allocation Programs at the LCFs

	60% INCITE		30% ALCC		10% Director's Discretionary	
Mission	High-risk, high-payoff science that requires LCF-scale resources*		High-risk, high-payoff science aligned with DOE mission		Strategic LCF goals	
Call	1x/year – (Closes June)		1x/year (Closes February)		Rolling	
Duration	1-3 years, yearly renewal		1 year		3m,6m,1 year	
Typical Size	30 - 40 projects	10M - 100M core-hours/yr.	5 - 10 projects	1M – 75M core-hours/yr.	100s of projects	10K – 1M core-hours
Review Process	Scientific Peer-Review	Computational Readiness	Scientific Peer-Review	Computational Readiness	Strategic impact and feasibility	
Managed By	INCITE management committee (ALCF & OLCF)		DOE Office of Science		LCF management	
Availability	Open to all scientific researchers and organizations ***Capability >20% of cores***					

Twofold review process

	New proposal assessment	Renewal assessment
① **Peer review: INCITE panels**	• Scientific and/or technical merit • Appropriateness of proposal method, milestones given • Team qualifications • Reasonableness of requested resources	• Change in scope • Met milestones • On track to meet future milestones • Scientific and/or technical merit
② **Computational readiness review: LCF centers**	• Technical readiness • Appropriateness for requested resources	• Met technical/ computational milestones • On track to meet future milestones
Award Decisions	• INCITE Awards Committee comprised of LCF directors, INCITE program manager, LCF directors of science, sr. management	