Crystal structure prediction of pharmaceutical-like molecules: Fast optimization methods and high-level energies

Gerit Brandenburg <g.brandenburg@ucl.ac.uk> | 9th of March 2018

THOMAS YOUNG CENTER SEMINAR, UNIVERSITY COLLEGE LONDON, UK

Outline of talk

1 Introduction

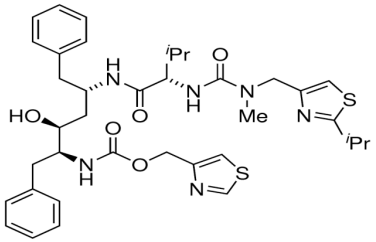
2 Simulation based crystal structure prediction

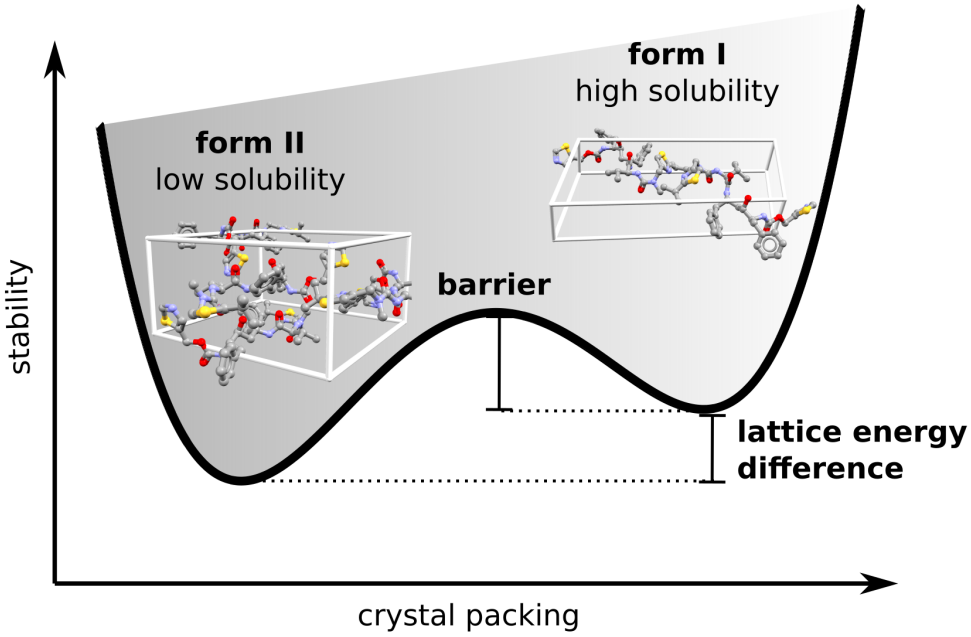
3 Quantum Monte-Carlo for molecular materials

4 Project outlook

Academic and industrial interest in molecular materials

API: ritonavir
- HIV antiviral drug
- industrially produced in form I
- form II found 2 years after market launch
- reformulation required [1]





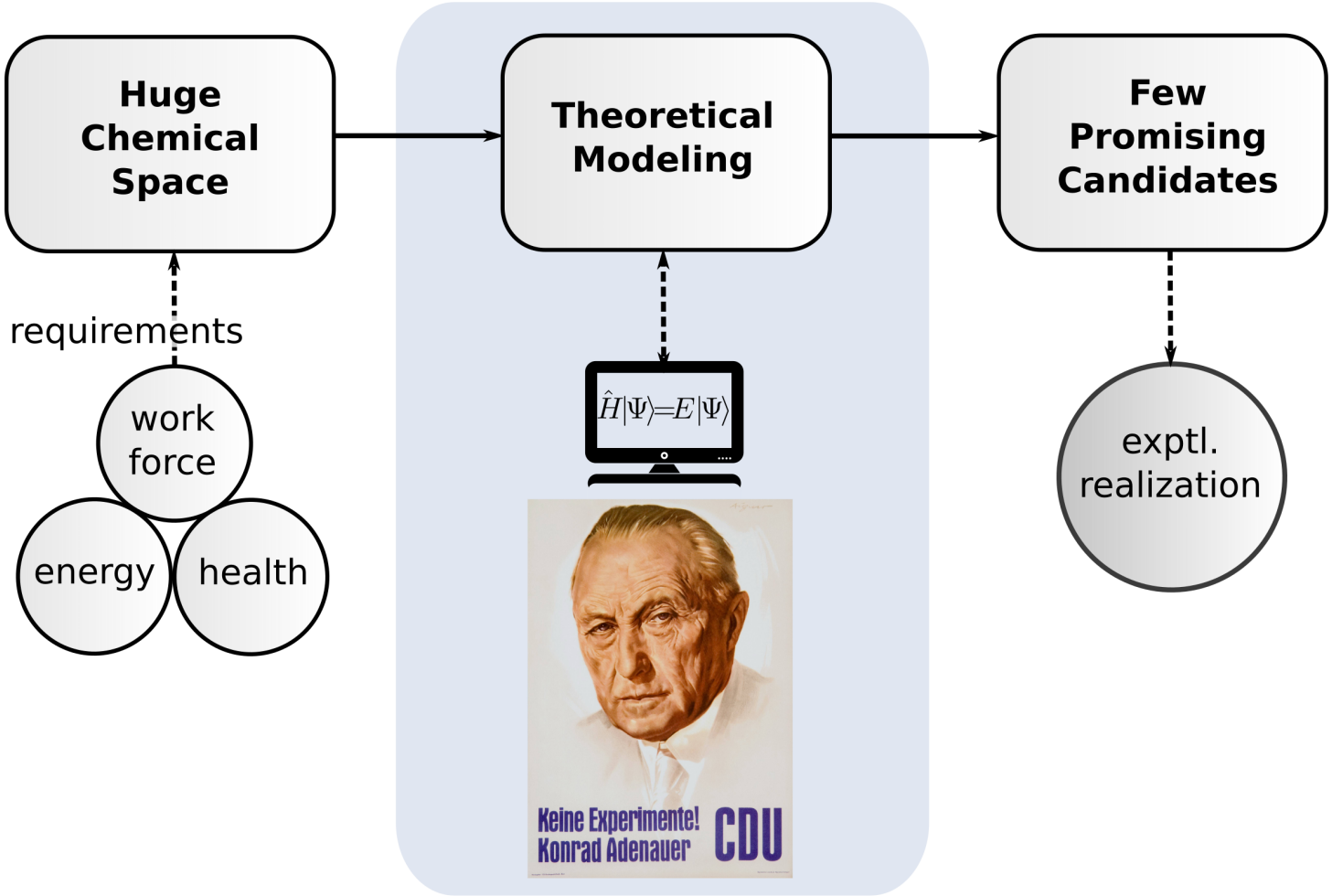
challenge A: optimize huge number of possible configurations

challenge B: predict the stability within (sub) chemical accuracy

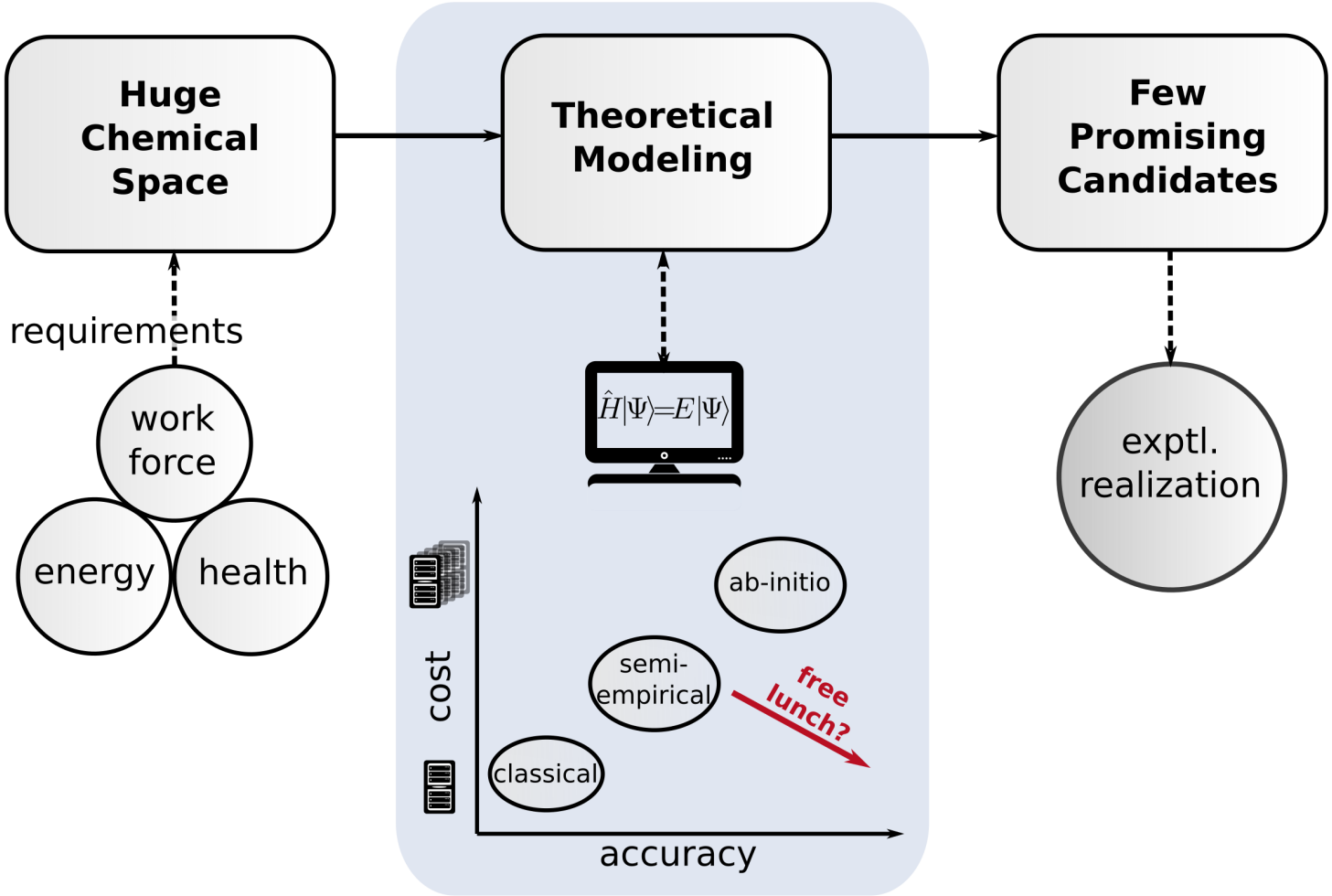
[1] J. Bauer, et al., *J. Pharm. Res.* **18**, 859-866 (2001).

[2] S. L. Price, JGB, *Molecular Crystal Structure Prediction*; Elsevier Australia ISBN: 9780128098356 (2017).

Chemical and material discovery can employ computational models

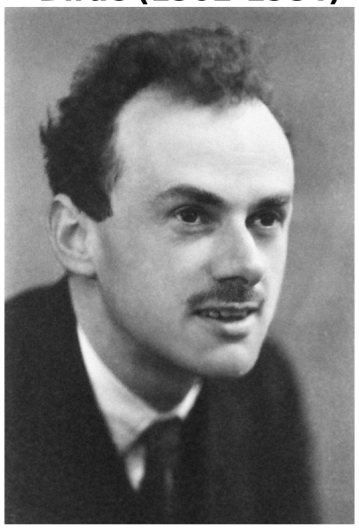


Chemical and material discovery can employ computational models



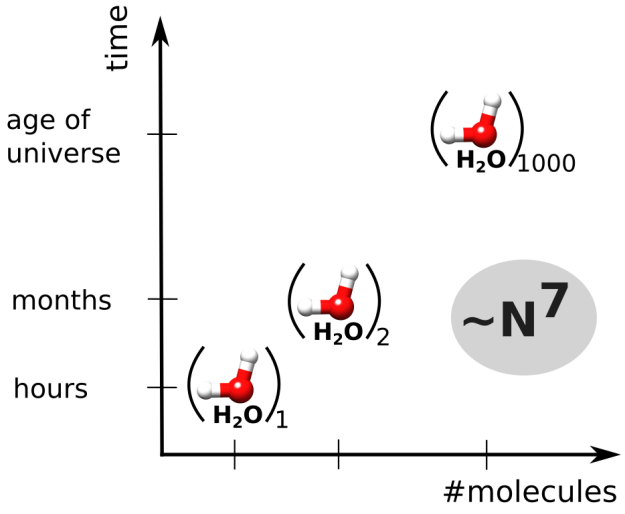
Exact simulation of extended systems computationally very demanding

Paul Adrian Maurice Dirac (1902-1984)



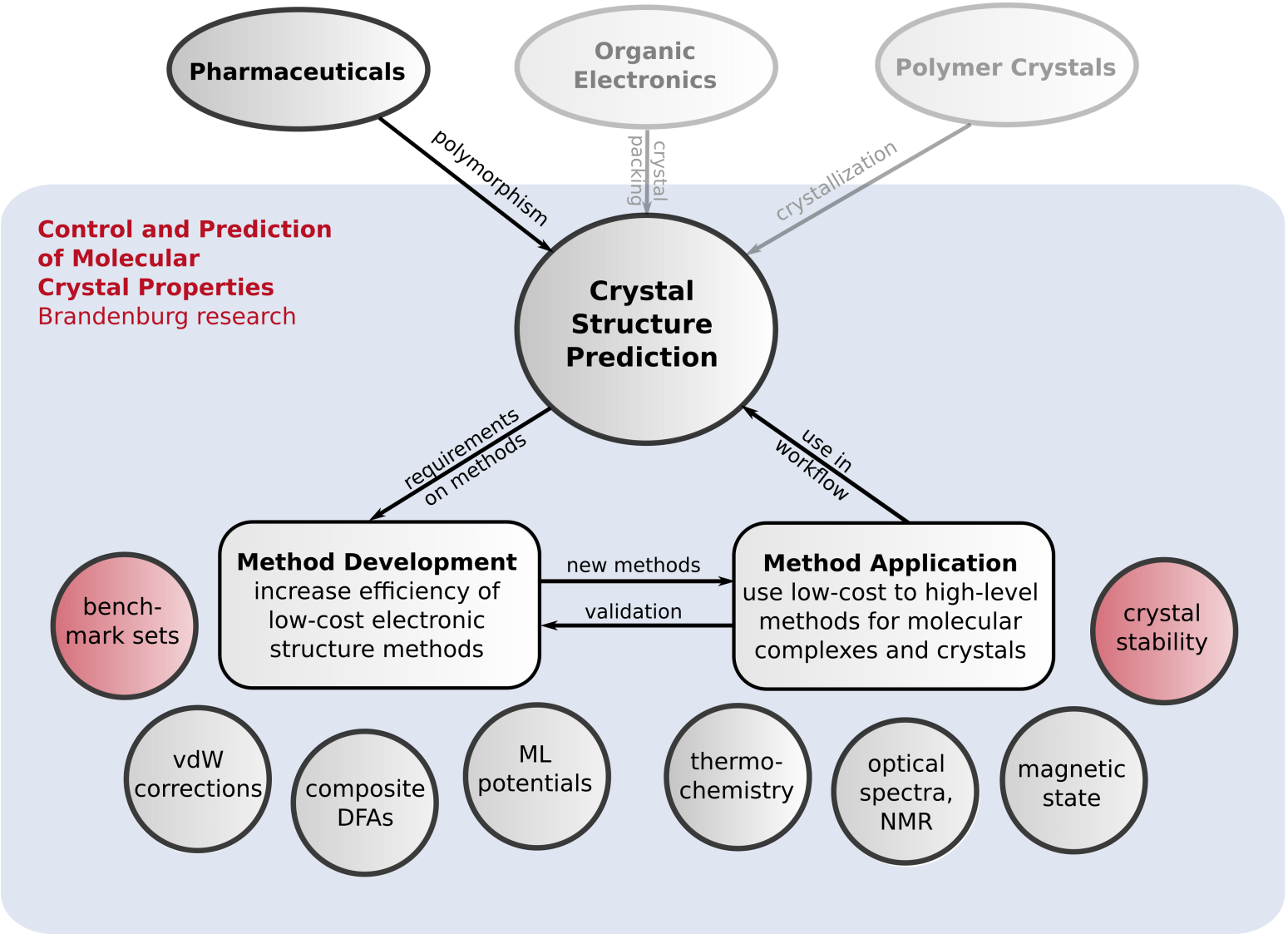
"The underlying physical laws (...) of a large part of physics and the ***whole of chemistry*** *are thus completely known,*

and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble"[1]



[3] P. A. M. Dirac, *Proc. Roy. Soc.Ser. A* **123**, 714 (1929).

Brandenburg research concept



Outline of talk

1 Introduction

2 Simulation based crystal structure prediction

3 Quantum Monte-Carlo for molecular materials

4 Project outlook

Predict most stable crystal polymorphs based on the molecular diagram

The inability to predict something as simple as how a molecule would crystallize is one of the continuing scandals in the physical sciences.[4–6]

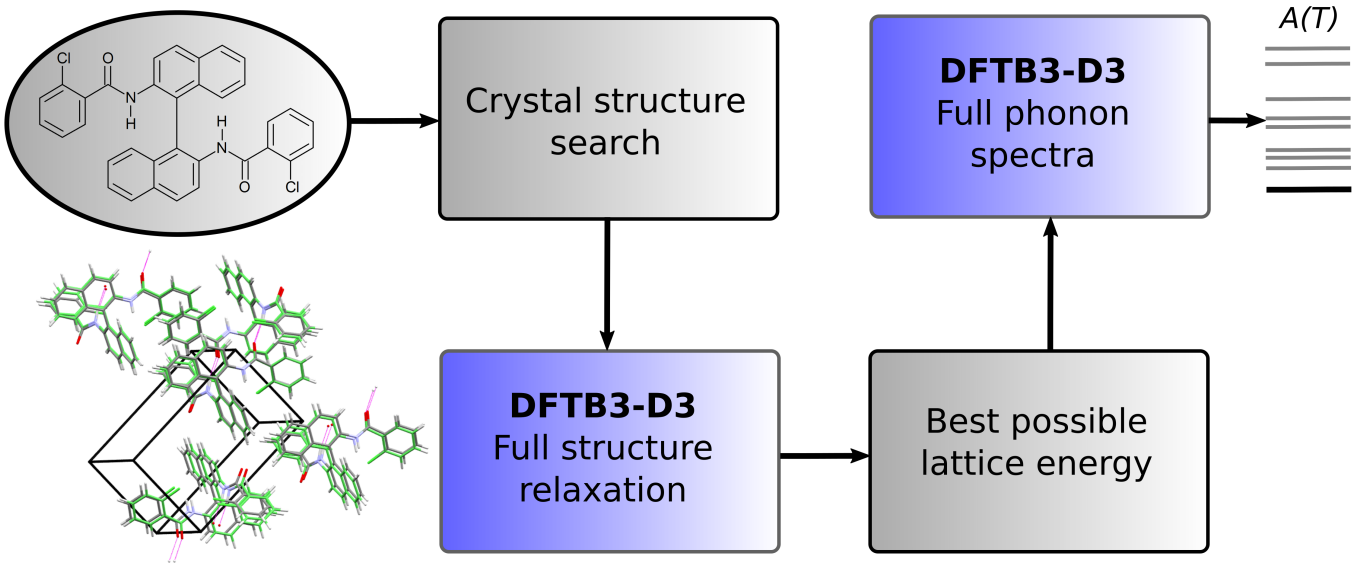
Task

- molecule is chosen due to its chemical/physical/biological properties
- based on the molecular diagram only, the most stable crystal structures should be predicted
- predict properties of interest for the most promising candidates

[4] A. Gavezzotti, *Acc. Chem. Res.* **27**, 309-314 (1994). [5] J. Maddox, *Nature* **335**, 201-201 (1988).

[6] K. N. Houk, F. Liu, *Acc. Chem. Res.*, **50**, 539 (2017).

Sampling and energetic ranking for crystal structure prediction

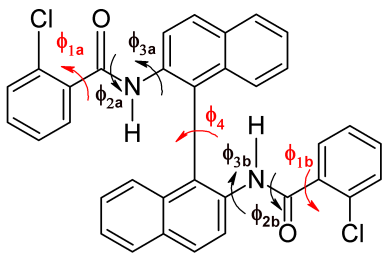


- search: use flexible torsions as search variable (0.4–2.2 million structures)
- lattice energy: established ES multipole based potential
- geometries: semi-empirical method as intermediate step

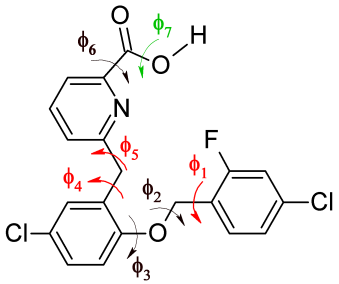
[7] S. Price, *Chem. Soc. Rev.* **43**, 2098 (2014)

[8] L. Iuzzolino, P. McCabeb, S. L. Price, JGB, *Faraday Discuss.*, *in press*, DOI: 10.1039/C8FD00010G (2018).

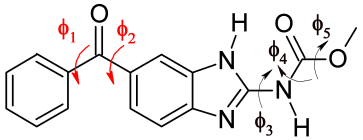
Flexible pharmaceutical-like targets



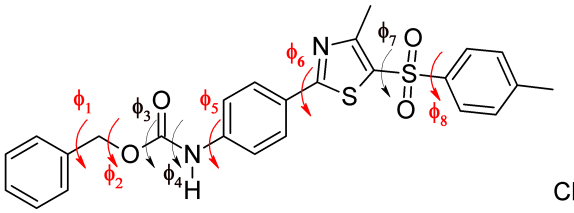
Molecule XXVI



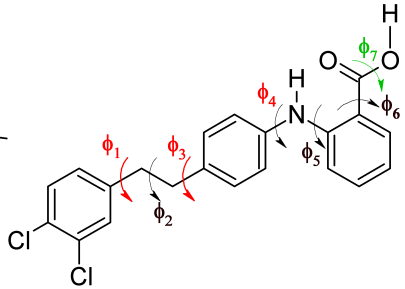
GSK269984B



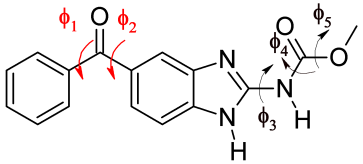
A-Mebendazole



Molecule XX



Molecule XXIII



C-Mebendazole

Possible merits of DFTB3-D3

(1) improving geometries

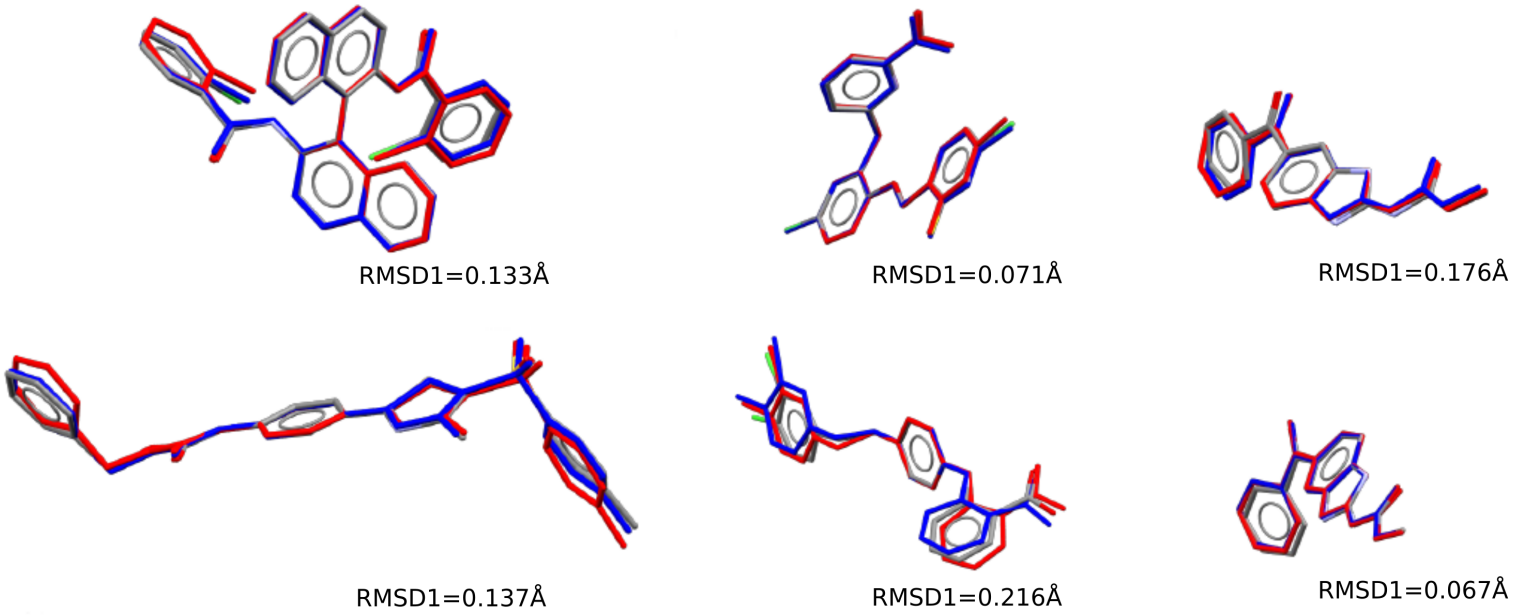
(2) improving energy ranking

(3) reducing # minima

(4) free energy contributions

[8] L. Iuzzolino, P. McCabeb, S. L. Price, JGB, *Faraday Discuss.*, *in press*, DOI: 10.1039/C8FD00010G (2018).

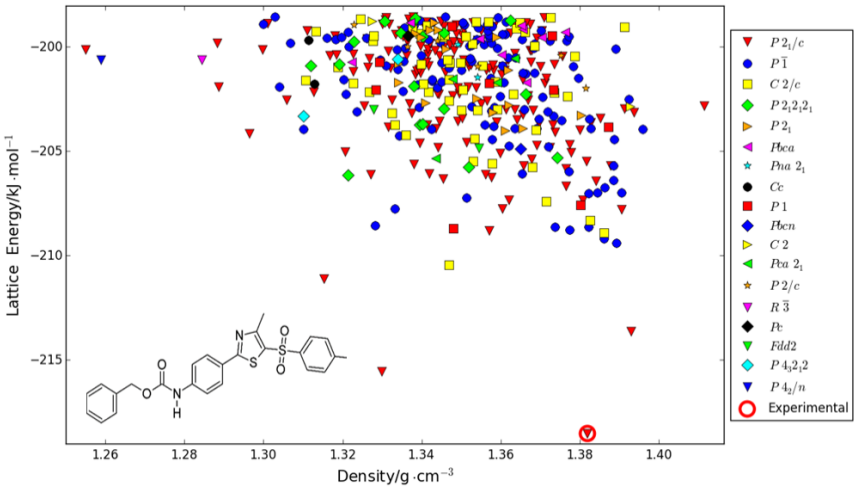
Geometries are indeed improved while energy ranking remains problematic

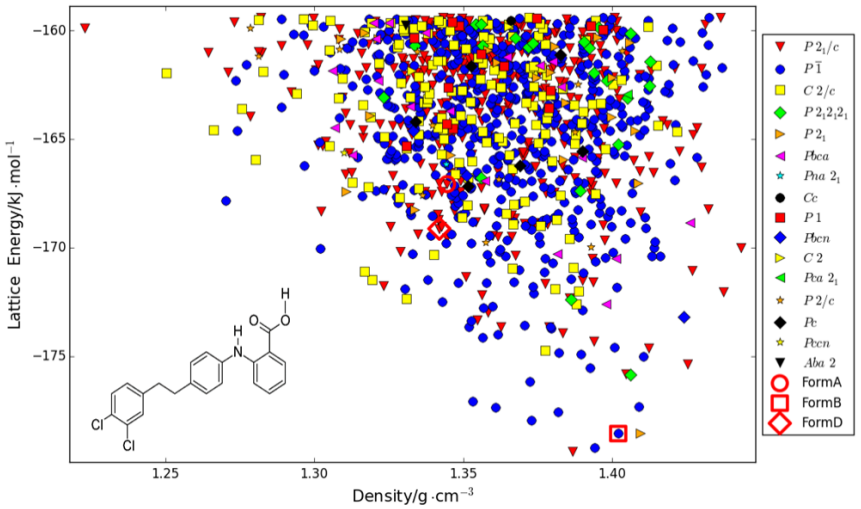


- good reproduction of conformers (RMSD1$<$ 0.3Å)
- crystal structures would be considered a match (RMSD15$<$ 0.8Å)

[8] L. Iuzzolino, P. McCabeb, S. L. Price, JGB, *Faraday Discuss.*, *in press*, DOI: 10.1039/C8FD00010G (2018).

Crystal energy landscape very molecule dependent

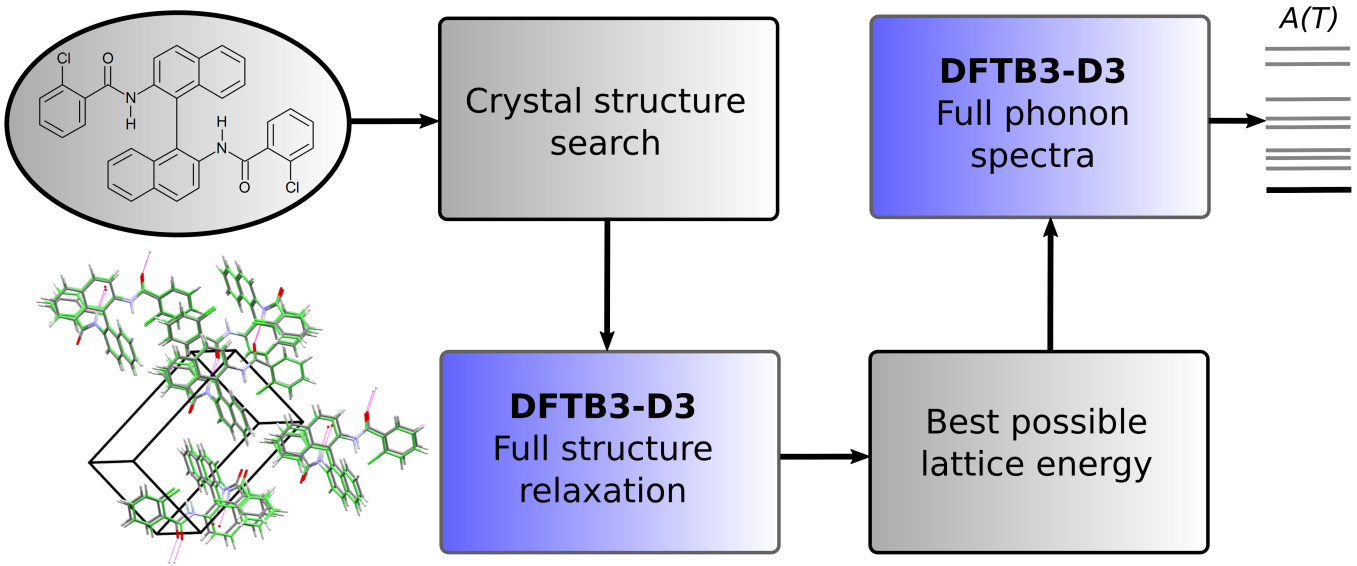




- use DFTB3-D3 geometries for improved energy evaluation
- ES multipole based potential yields good ranking
- up to 90% reduction in computational cost

[8] L. Iuzzolino, P. McCabeb, S. L. Price, JGB, *Faraday Discuss.*, *in press*, DOI: 10.1039/C8FD00010G (2018).

DFTB3-D3 intermediate optimization might be scaleable for larger molecules



Possible merits of DFTB3-D3

(1) improving geometries **yes**

(2) improving energy ranking **no**

(3) reducing # minima **partially**

(4) free energy contributions **yes**

[8] L. Iuzzolino, P. McCabeb, S. L. Price, JGB, *Faraday Discuss.*, *in press*, DOI: 10.1039/C8FD00010G (2018).

Outline of talk

1 Introduction

2 Simulation based crystal structure prediction

3 Quantum Monte-Carlo for molecular materials

4 Project outlook

Multilevel methodologies: Finding the right compromise

	task/property	example method
accurate QM	single-point energy	many-body WFT (DMC[9])
cheap QM	optimization	DFT[10−13]
very cheap QM	optimization/Hessians conformations	semi-empirical[14−15]
force field	dynamics conformational sampling	transferable or molecule specific (QM derived) FF

[9] A. Zen, JGB, J. Klimeš, A. Tkatchenko, D. Alfè, and A. Michaelides, *Proc. Natl. Acad. Sci. U.S.A*, **115**, 1724 (2018).

[10] JGB, E. Caldeweyher, S. Grimme, *Phys. Chem. Chem. Phys.*, **18**, 15519 (2016).

[11] JGB, C. Bannwarth, A. Hansen, S. Grimme, *JCP*, **148**, 064104 (2018). [12] M. Cutini, et al., *JCTC*, **12**, 3340 (2016).

[13] JGB, J. E. Bates, J. Sun, J. P. Perdew *Phys. Rev. B*, **94**, 115144 (2016).

[14] JGB, S. Grimme, *JPCL*, **5**, 1785 (2014). [15] M. Mortazavi, JGB, R. J. Maurer, A. Tkatchenko, *JPCL*, **9**, 399 (2018)

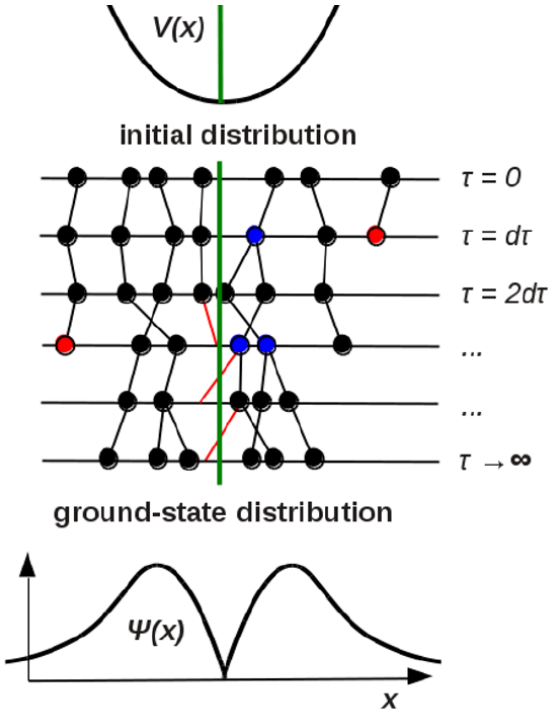
Quantum Monte-Carlo in a nutshell: A scalable high-level method

How to solve Schrödinger equation for large systems?

FN-diffusion Monte-Carlo

$$|\Psi_0(\mathbf{R})\rangle = \lim_{\tau\to\infty} \exp\left[-\tau(\hat{H} - E_T)\right] |\Psi_T(\mathbf{R},\tau)\rangle$$

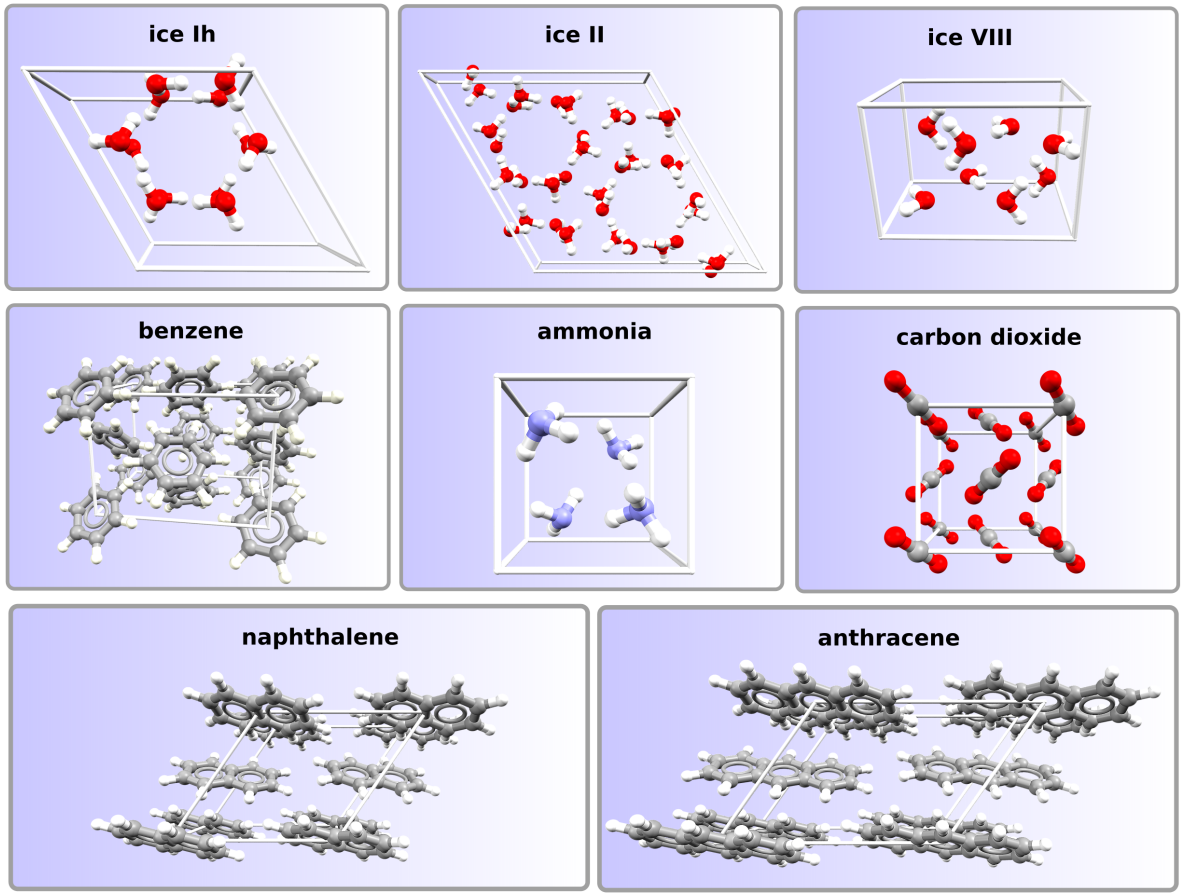
- low-scaling (N^3) with system size
- scaleable to HPCs[16]
- new size-consistent algorithm leads to substantial speed up[17]



[16] M. Ďubecký, L. Mitas, P. Jurečka, *Chem. Rev.* **116**, 5188 (2016).

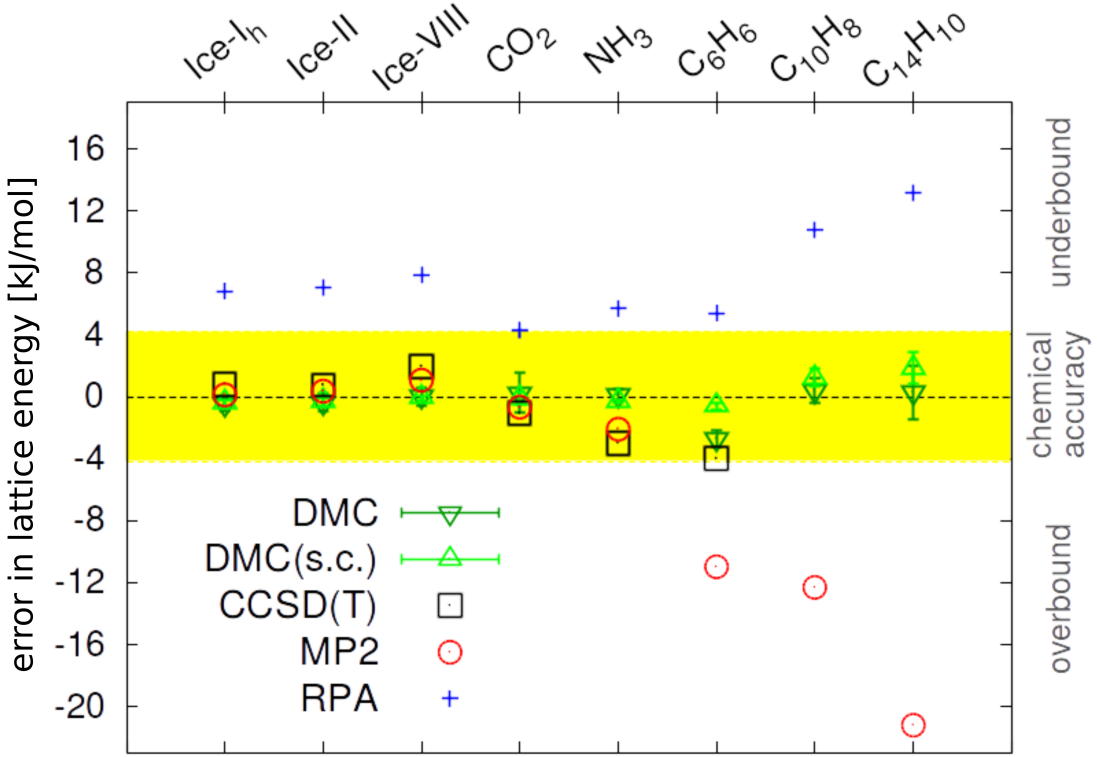
[17] A. Zen, S. Sorella, M. J. Gillan, A. Michaelides, D. Alfé, *Phys. Rev. B* **93**, 241118(R) (2016).

Diverse interactions in test crystals



- strong H-bonds, vdW of saturated and unsaturated molecules
- problematic for all readily applicable methods (DFT-D, MP2)

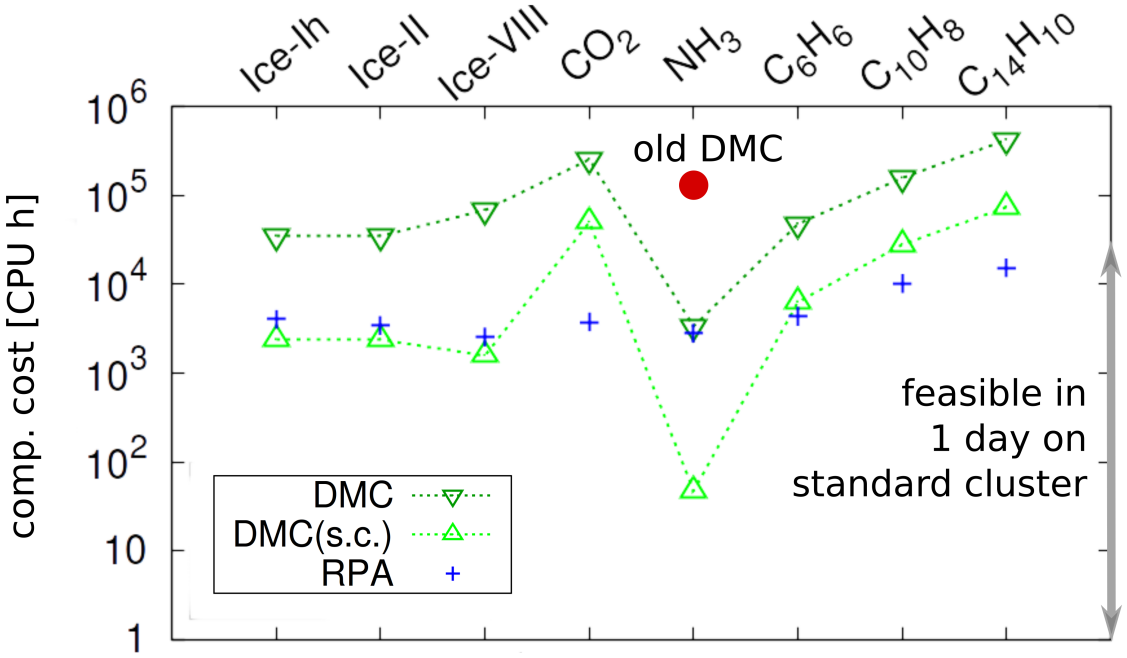
QMC delivers (sub-) chemical accuracy for all tested systems



- excellent agreement with experiment and CCSD(T)[18]
- uncertainty in $H_{\text{sub}}^{\text{exp}}$ probably larger than DMC errors

[18] Y. S. Al-Hamdani, M. Rossi, D. Alfè, T. Tsatsoulis, B. Ramberger, JGB, A. Zen, G. Kresse, A. Grüneis, A. Tkatchenko, A. Michaelides *J. Chem. Phys.* **147**, 044710 (2017).

QMC is feasible within one day on standard computer cluster



- up to three orders of magnitude speed-up compared to best DMC practice two years ago
- significance will extend to all classes of systems[16]

[16] A. Zen, JGB, J. Klimeš, A. Tkatchenko, D. Alfè, A. Michaelides, *Proc. Natl. Acad. Sci. U.S.A* , **115**, 1724 (2018).

Outline of talk

1 Introduction

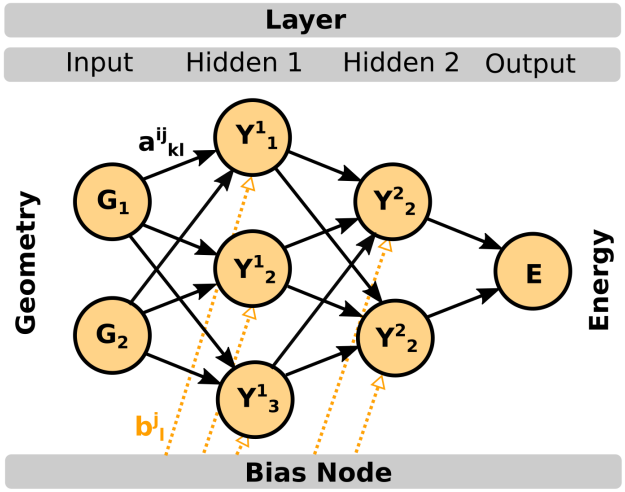
2 Simulation based crystal structure prediction

3 Quantum Monte-Carlo for molecular materials

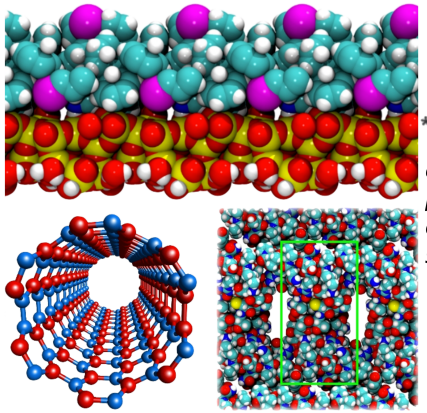
4 Project outlook

Use machine learning potentials within the CSP framework

- high dimensional neural networks for potential interpolation[19–20]
- reduce the number of required *ab initio* calculations within CSP



- major contributions to CRYTAL17
- combine efficient program handling with modern density functionals



CRYSTAL17

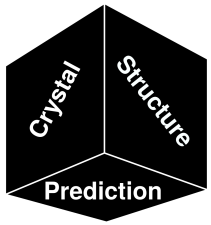
enables consistent studies of the physical and chemical properties of molecules, polymers, nanotubes, surfaces and crystalline solids

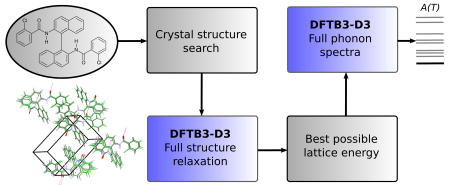
Summary

Two challenges of CSP have been discussed

A: semiempirical method help in optimizing huge number of crystal structures

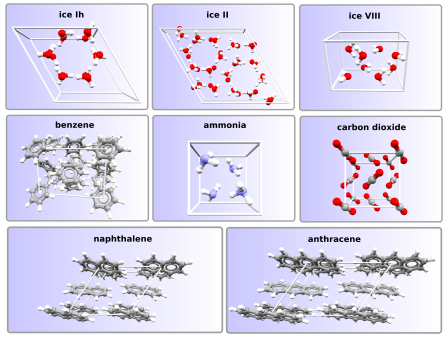
B: many-body methods can compute lattice energies within chemical accuracy





Outlook

- thermodynamic contributions have to be analyzed further
- beyond haarmonic approximation?



Acknowledgements

PhD / Postdocs

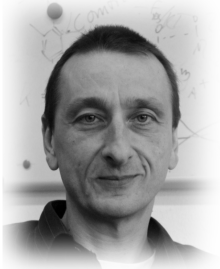

Luca Iuzzolino


Price (UCL)


Andrea Zen


Michaelides (UCL)


Eike Caldeweyer


Grimme (Bonn)

Principal investigator


McCabe (CCDC)

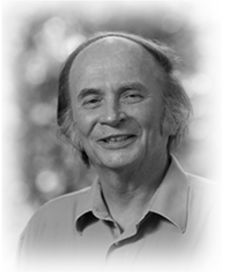

Hall (Bristol)


Behler (Göttingen)


Alfé (UCL)


Tkatchenko (Luxembourg)


Klimes (Prague)


Perdew (Temple)


Civalleri (Turin)


Maschio (Turin)

Industry


Rebecca Sure (BASF)


Antreas Afantitis (NovaMechanics)

Alexander von Humboldt Stiftung/Foundation

Acknowledgements

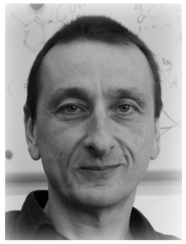
PhD / Postdocs

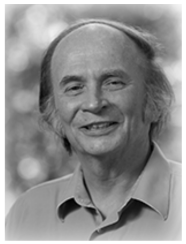

Eike Caldeweyer


Andrea Zen


Luca Iuzzolino

Principal investigator


Grimme (Bonn)


Perdew (Temple)


Civalleri (Turin)


Maschio (Turin)


Michaelides (UCL)


Alfé (UCL)

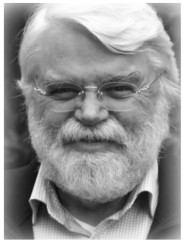

Tkatchenko (Luxembourg)


Behler (Göttingen)


Price (UCL)


Hall (Bristol)


Fernandez-Alonso (ISIS)


Erker (Münster)

Industry


Rebecca Sure (BASF)


Antreas Afantitis (NovaMechanics)

Thanks

Key references

- Crystal structure prediction:

  S. L. Price, JGB, *Molecular Crystal Structure Prediction*, G. DiLabio, A. Otero-de-la-Roza, Eds., Elsevier Australia, *ISBN: 9780128098356* (2017).

  L. Iuzzolino, P. McCabeb, S. L. Price, JGB, *Faraday Discuss.*, *in press* DOI: 10.1039/C8FD00010G (2018).

- QMC for molecular crystals:

  A. Zen, JGB, J. Klimeš, A. Tkatchenko, D. Alfè, A. Michaelides, *Proc. Natl. Acad. Sci. U.S.A*, **115**, 1724 (2018).

- DFT development

  JGB, J. E. Bates, J. Sun, J. P. Perdew, *Phys. Rev. B*, **94**, 115144 (2016).

  JGB, C. Bannwarth, A. Hansen, S. Grimme *J. Chem. Phys.*, **148**, 64104 (2018).

- website: **gerit-brandenburg.de**