Data-Driven Many-Body Models for Predictive Molecular Simulations

Francesco Paesani

Department of Chemistry and Biochemistry
Materials Science and Engineering
San Diego Supercomputer Center
University of California, San Diego

@PaesaniLab

Graduates

Colin Egan
Kartik Rallapalli
Yaoguang Zhai
Kelly Hunter
Teri Lambros
Chanin Tangtartharakul
Alessandro Caruso
Hanjun Kim

Postdocs

Ongun Ozcelik
Marc Riera
Raja Ghosh

Undergraduates

Eric Yeh
Ethan Bull-Vulpe
Xuanyu Zhu
Yuqing Shen
Kaushik Ganapathy
Alan Hirales
Eric Kang
Perry Chu
Isaac Douglas

High Schoolers

Kimmy Fogliatti

Former Members

Greg Medders
Wei Lin
Pushp Bajaj
Debbie Zhuang
Brandon Bizzarro
Volodymyr Babin
Jordi Cirera
Sandeep Reddy
Huy Pham
Dan Moberg
Ayo Ajala
Sandra Brown

Collaborations

S.M. Cohen (UCSD)
A.F. Tezcan (UCSD)
A.W. Götz (UCSD)
V. Molinero (Utah)

N.C. Gianneschi (Northwestern)
F.M. Geiger (Northwestern)
G. Galli (Chicago)
T. Heine (TU Dresden)

M. Ceriotti (EPFL)
G. Csányi (Cambridge)
J. Behler (Göttingen)
J.O. Richardson (ETH)

G.K. Schenter (PNNL)
J.L. Fulton (PNNL)
T. Carrington (Queen's)
A.C. Simmonett (NIH)

M. Dinča (MIT)
M. Bonn (MPI Mainz)
Y. Nagata (MPI Mainz)
E. Tyrode (KTH)

NIH
National Institutes of Health

NERSC

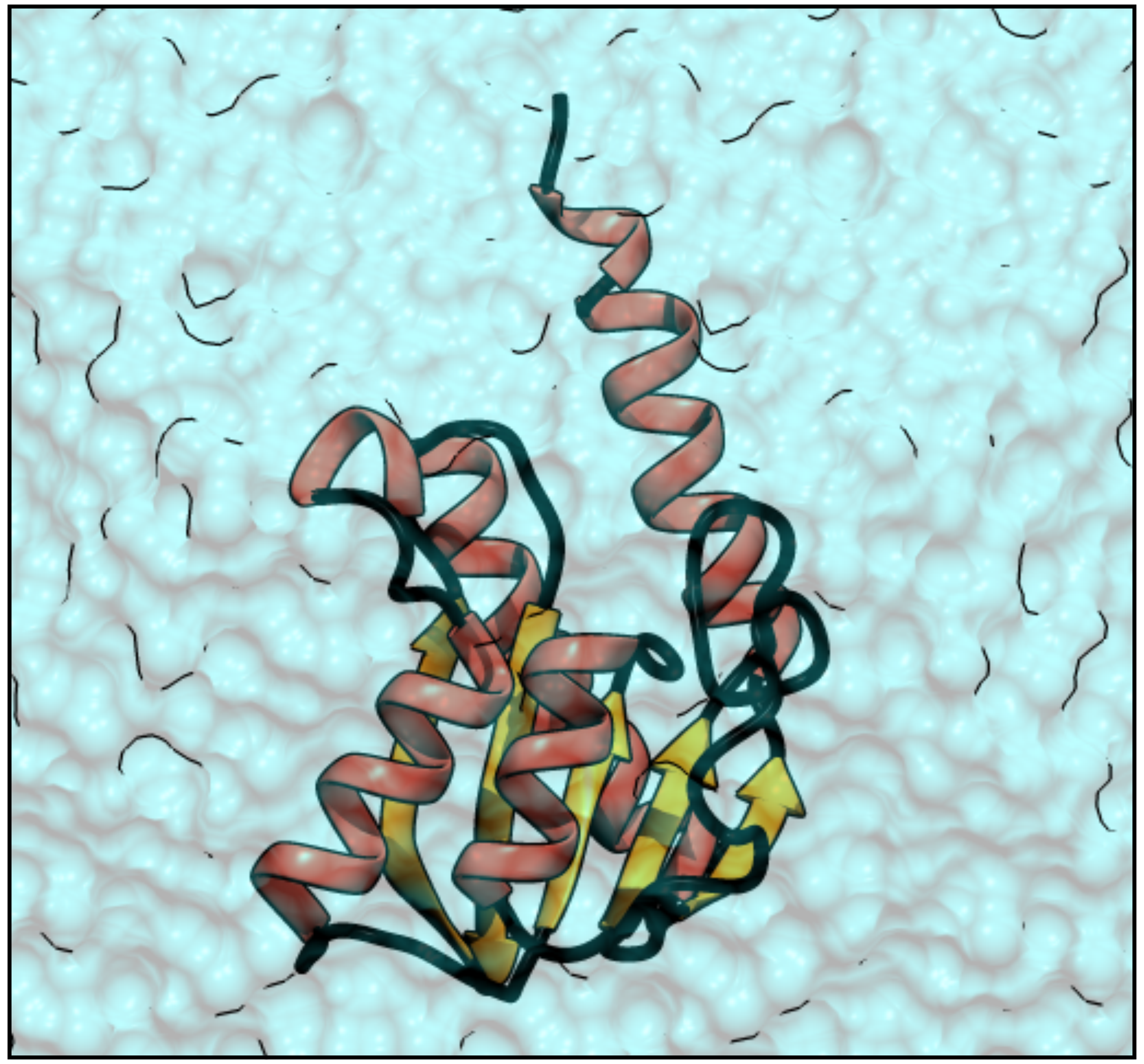
XSEDE
Extreme Science and Engineering Discovery Environment

Triton Resource

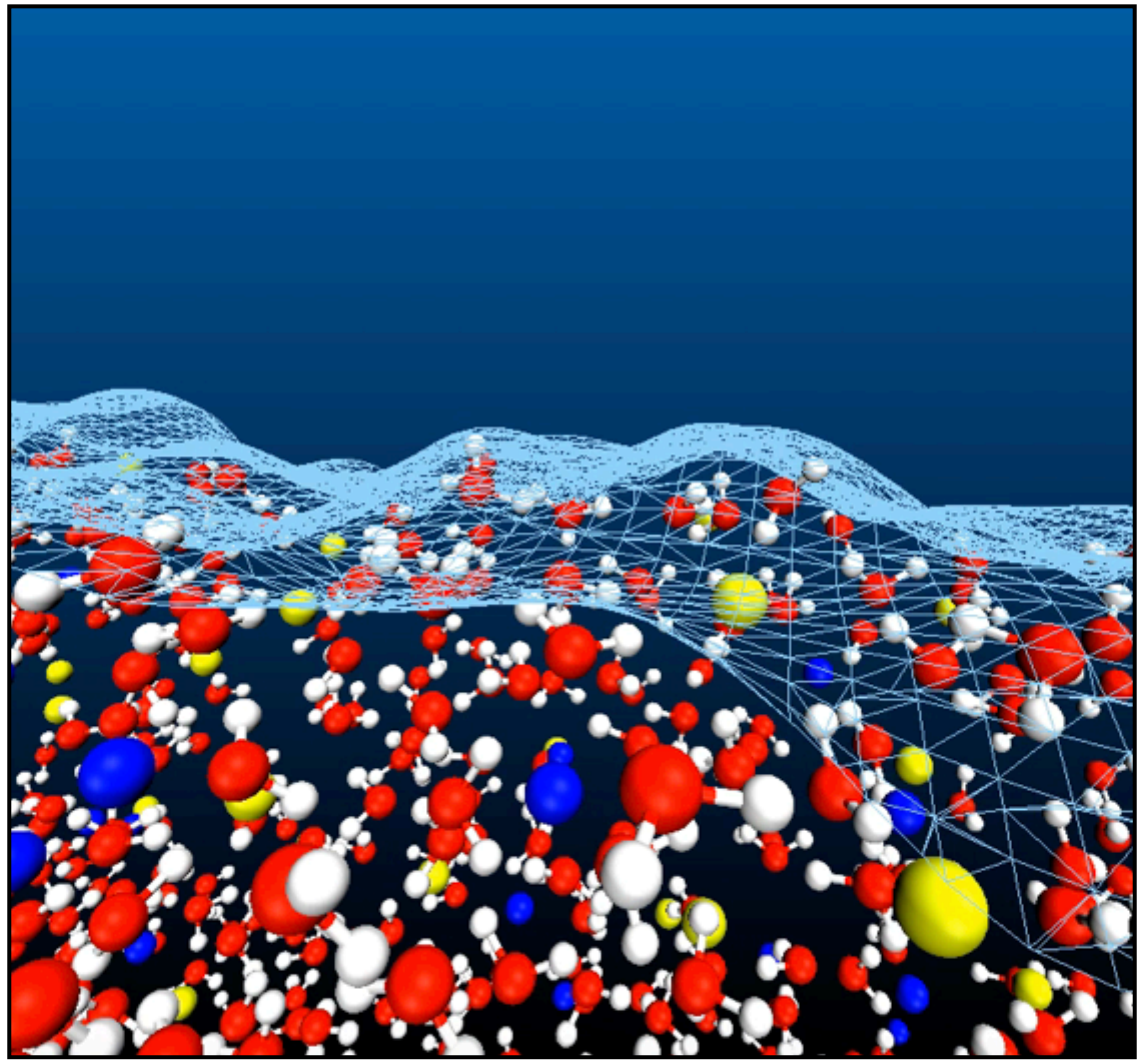
http://paesanigroup.ucsd.edu

The Importance of Water Cannot Be Overemphasized…

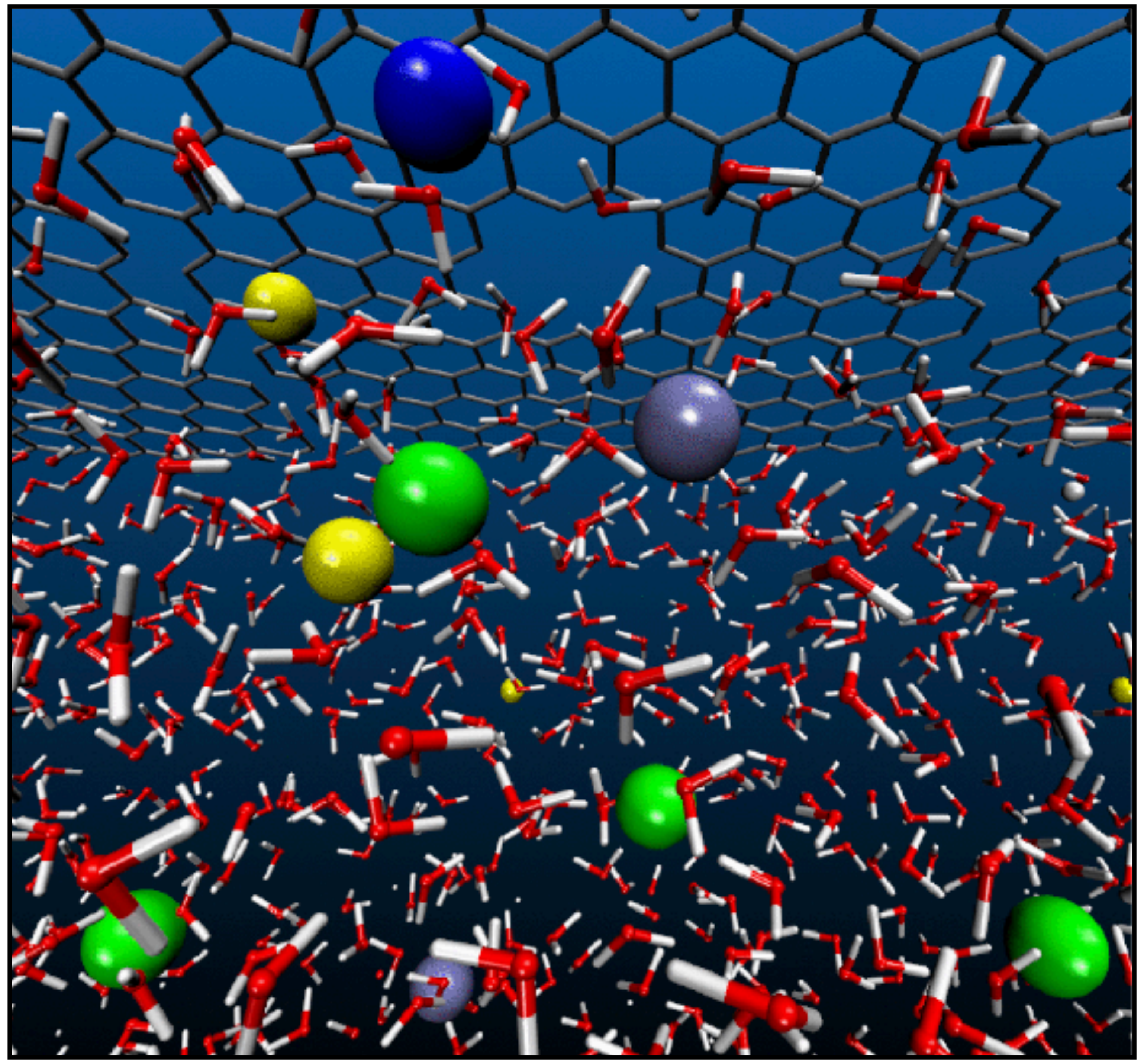
hydration of biomolecules



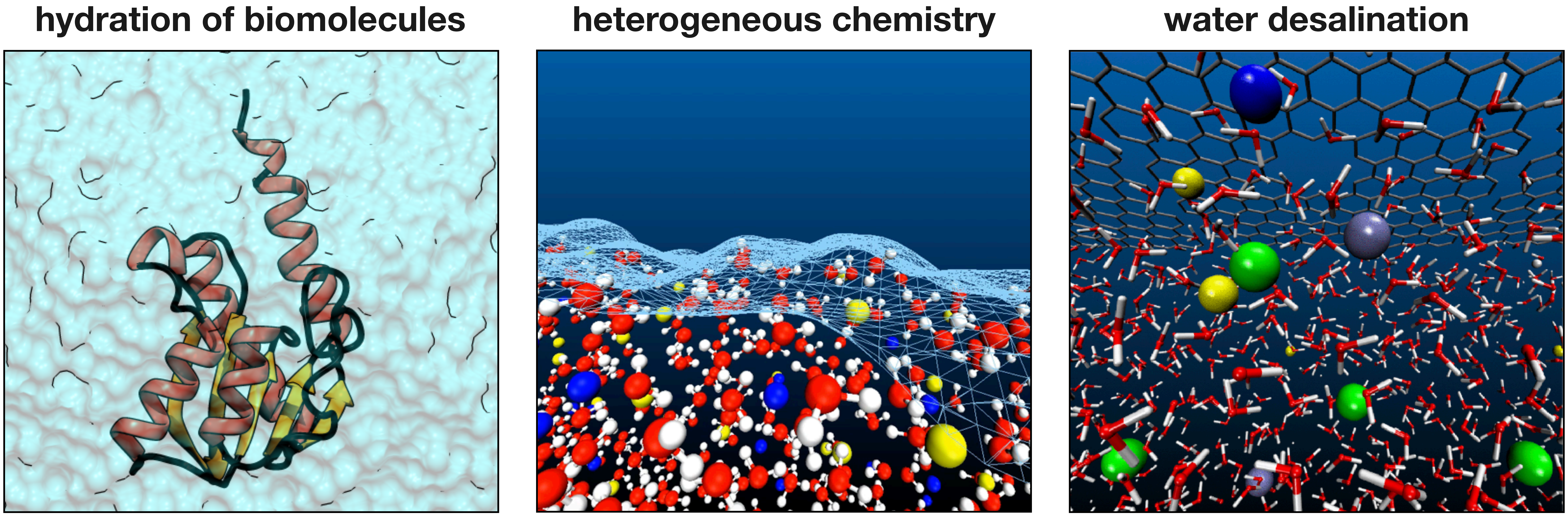
heterogeneous chemistry



water desalination



The Importance of Water Cannot Be Overemphasized…


hydration of biomolecules

heterogeneous chemistry

water desalination

Molecular simulations: Fiction or reality?

Water: The Perfect Example!

Water models

The following is an incomplete list of models used in computer simulations of water:

A

- AMOEBA
- ASP-S
- ASP-W

B

- Baranyai
- BBL
- Bell-Lavis
- Ben-Naim models
- BF
- BJH
- BK3
- BKd1
- BKd2
- BKd3
- BNS
- Bol

C

- CC-pol
- CF
- CI
- COS
- COS/B2
- COS/D2
- COS/G2
- COS/G3
- CMP
- CKL

D

- Dahl and Andersen
- Dang97
- DCF
- DEC
- DPP
- DPP2

E

- E3B
- E3B2
- E3B3
- ELBA
- ENCS
- EPEN/2

F

- F3C

G

- Gaussian model
- GCPM

H

- HBB
- HBB2-pol

K

- KJ
- KKY

M

- MB-pol
- MCDHO
- MCHO
- MCY
- MCYL
- Mercedes-Benz model
- mW

N

- NCC
- NCCvib
- NCF
- NEMO
- NSPCE
- NvdE

O

- OPC3

P

- PE
- PMW
- POL3
- POL4D
- POL5
- POLIR
- PPC
- PTIP4P
- PSRWK
- PW

Q

- QCT
- q-SPC/Fw
- q-TIP4P/F

R

- R
- refPOL
- RPOL
- RSL2
- RWFF
- RWK

S

- SAPT-5s
- SAPT-pp
- SCPDP
- SDFT-5s

Simple Point Charge models

- SPC
- SPC/ε
- SPC/A
- SPC/E
- SPC/F
- SPC/F2
- SPC/FP
- SPC/FQ
- SPC/Fw
- SPC/HW
- SPC/L
- SPCP
- SPC-pol
- SSD
- ST2
- ST2RF
- SWFLEX
- SWM4-DP
- SWM4-NDP
- SWM6

T

- TCPE

Transferable Intermolecular Potential models

- TIP3P
- TIP3P/Fs
- TIP4F
- TIP4P
- TIP4P/2005
- TIP4P/2005f
- TIP4P/ε
- TIP4P-Ew
- TIP4P/FQ
- TIP4P-HB
- TIP4P-I
- TIP4P/Ice
- TIP4P-pol
- TIP4PQ/2005
- TIP4PQ_D2O
- TIP4PQ_T2O
- TIP4P-QDP
- TIP4P-QDP-LJ
- TIP4Q
- TIP5P
- TIP5P-E
- TIPS
- TIPS2
- TTM2-F
- TTM2-R
- TTM2.1
- TTM2.1-F
- TTM3-F

V

- VRT(MCY-5f)
- VRT(ASP-W)III

W

- WK

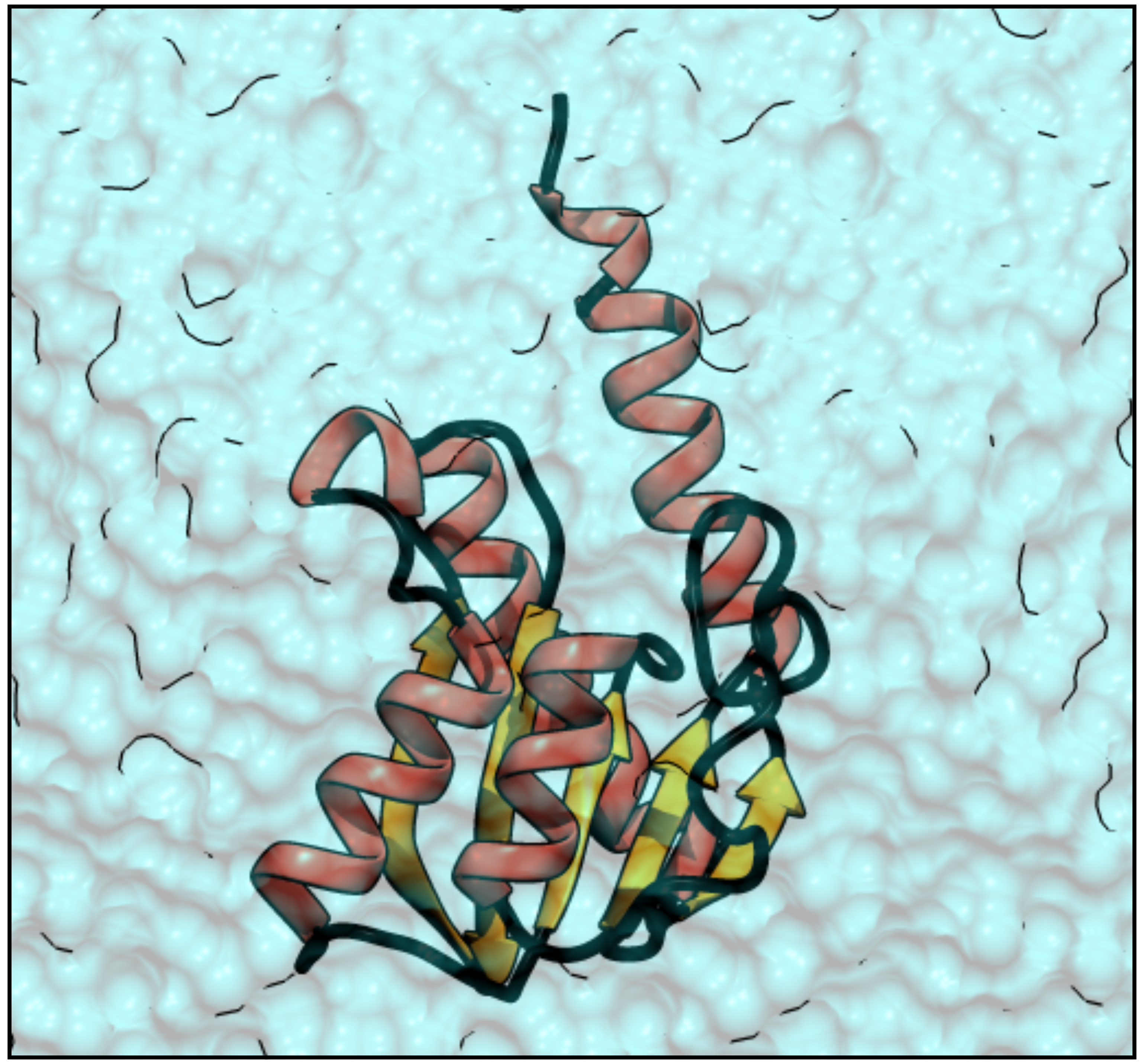
http://www.sklogwiki.org/SklogWiki/index.php/Water_models

Learning How to Survive in the “Water Community”

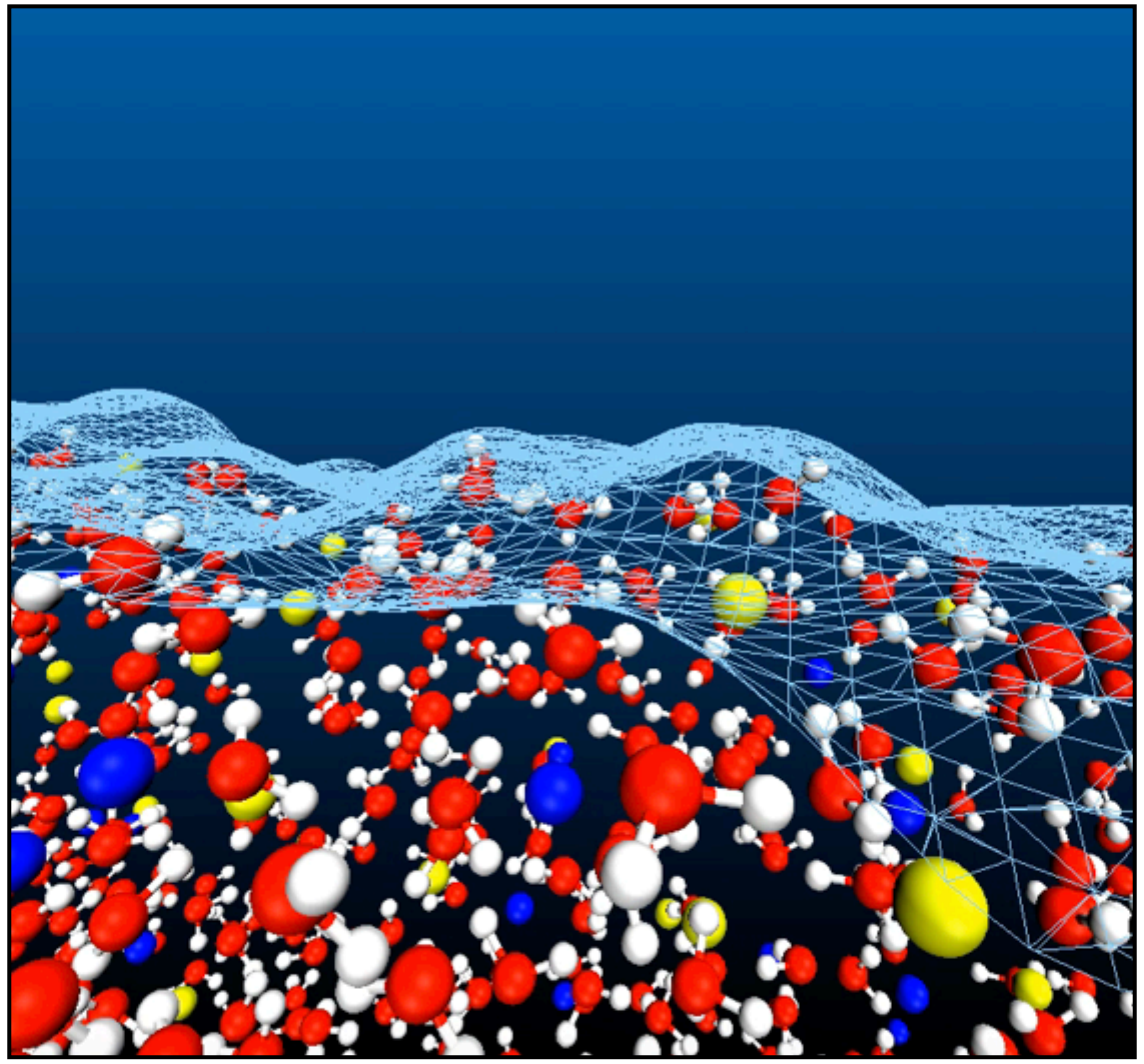
Learning How to Survive in the "Water Community"

The Importance of Water Cannot Be Overemphasized…

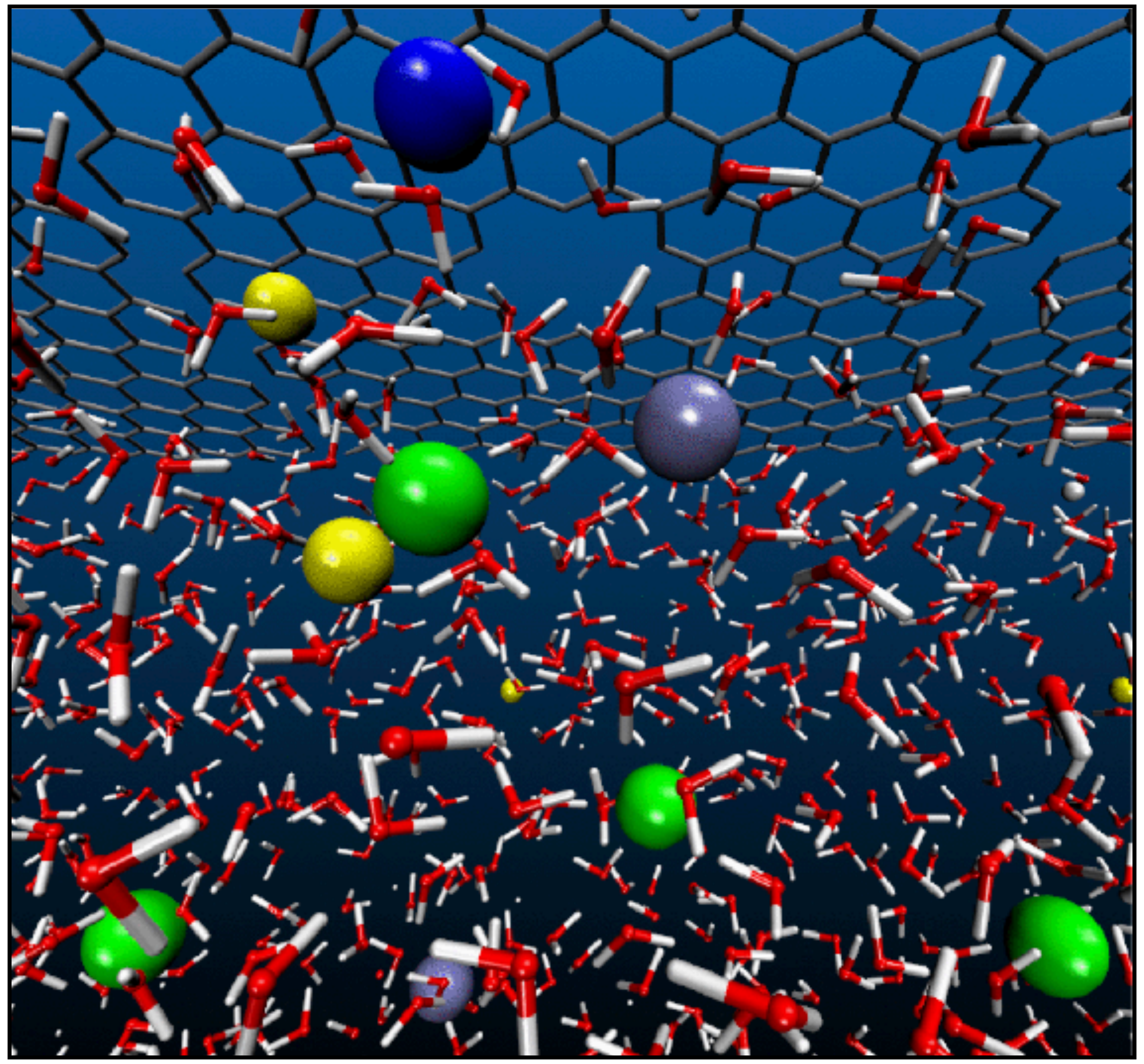
hydration of biomolecules



heterogeneous chemistry



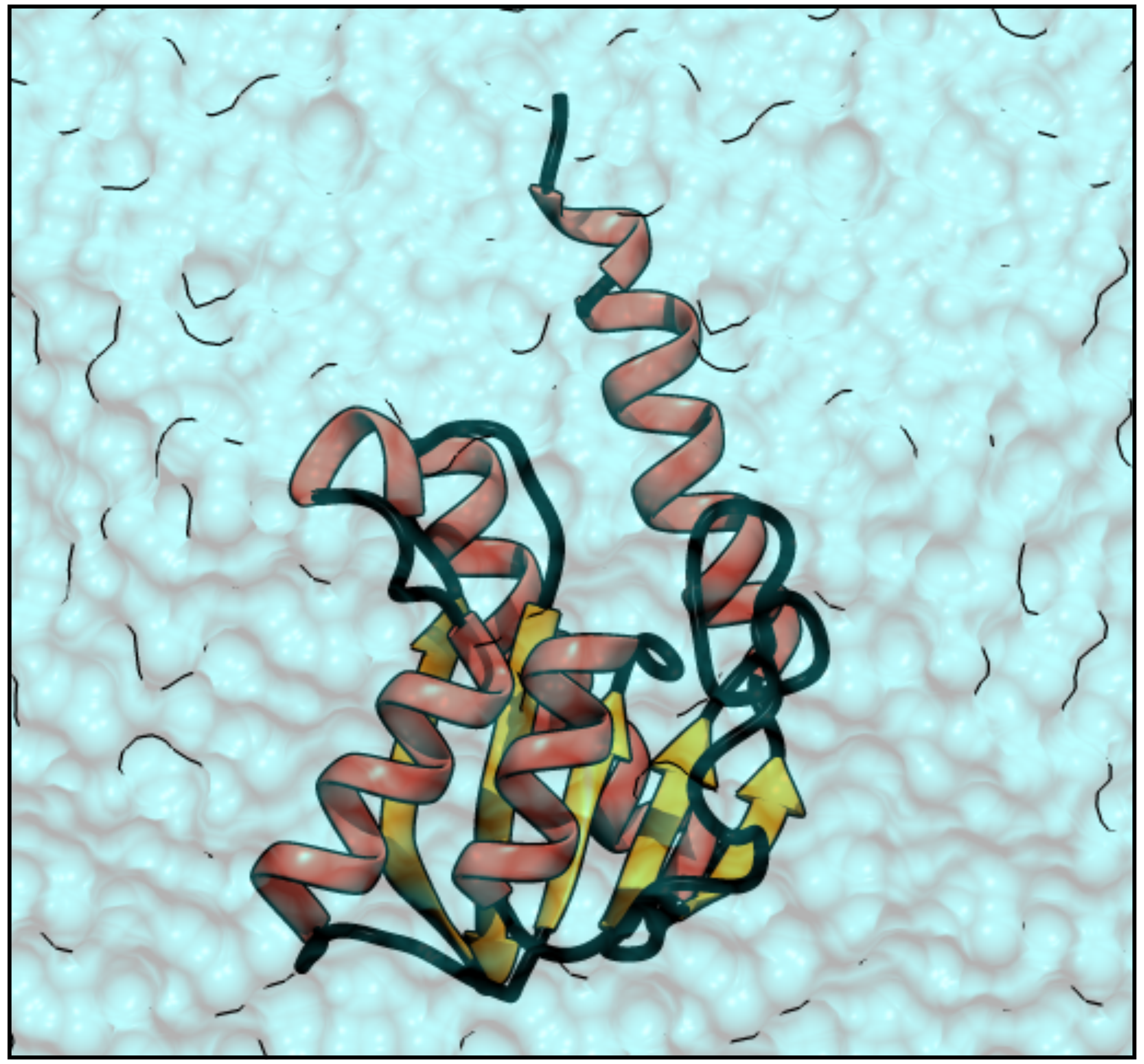
water desalination



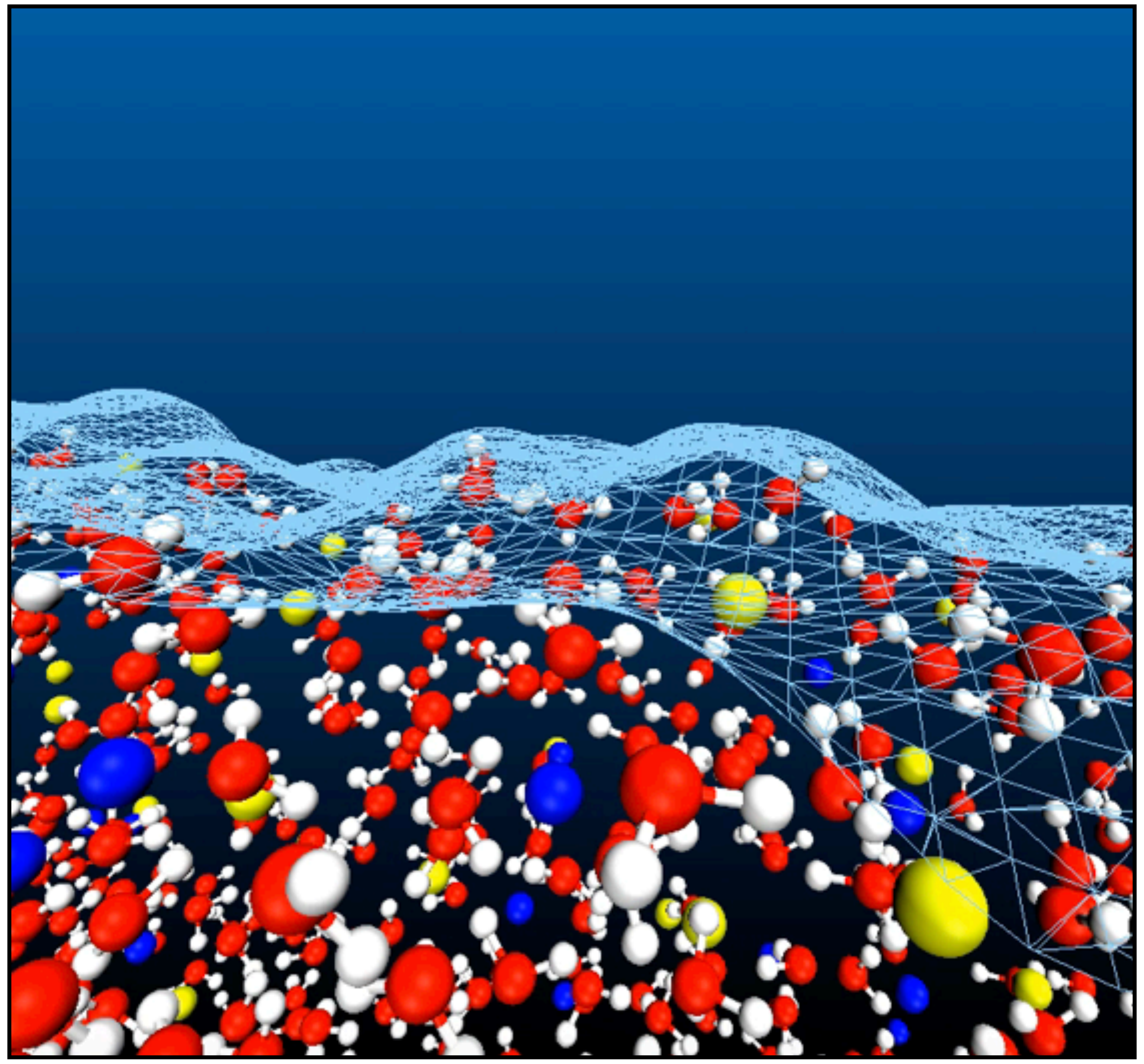
Molecular simulations: Fiction or reality?

The Importance of Water Cannot Be Overemphasized…

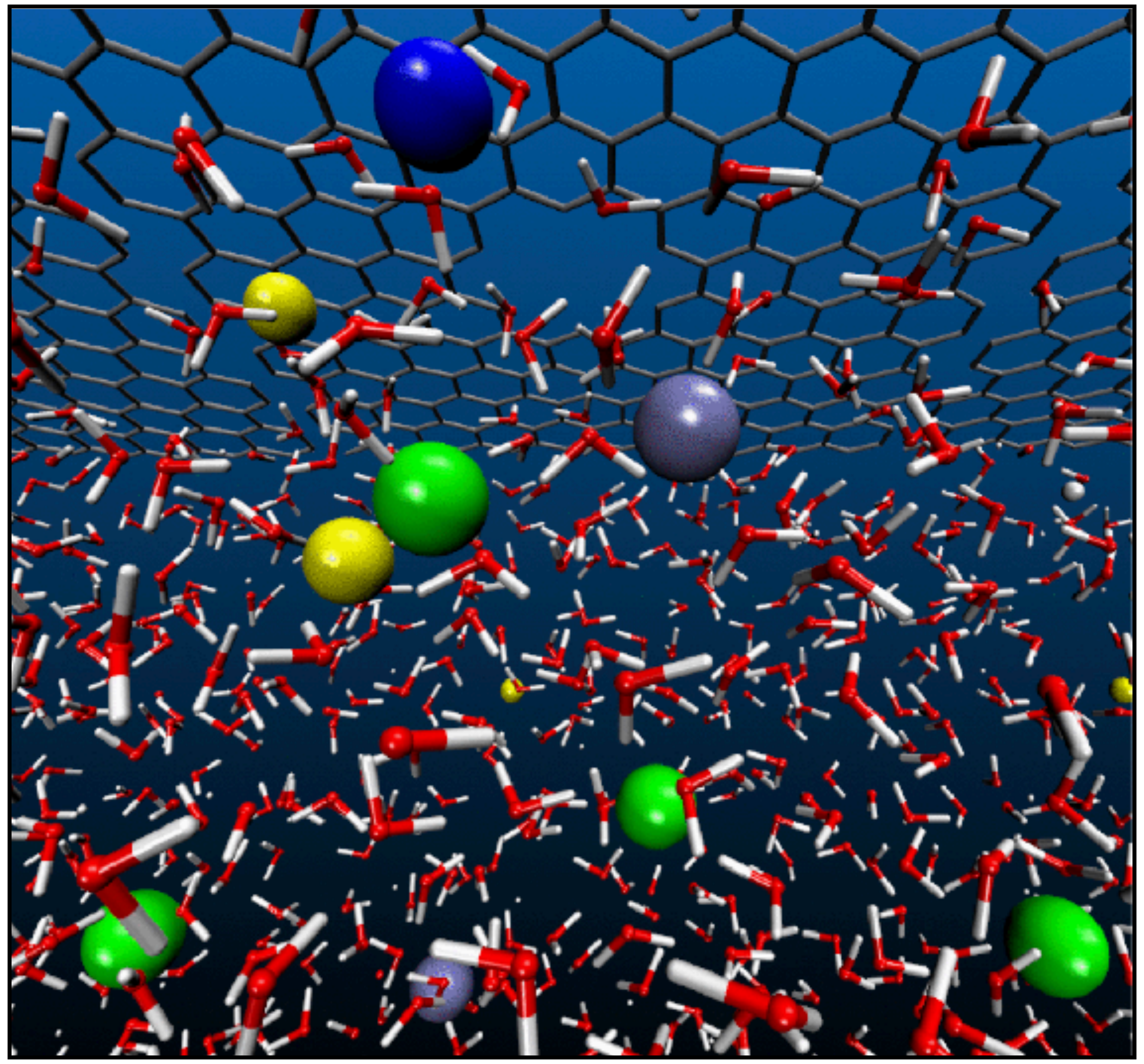
hydration of biomolecules



heterogeneous chemistry



water desalination

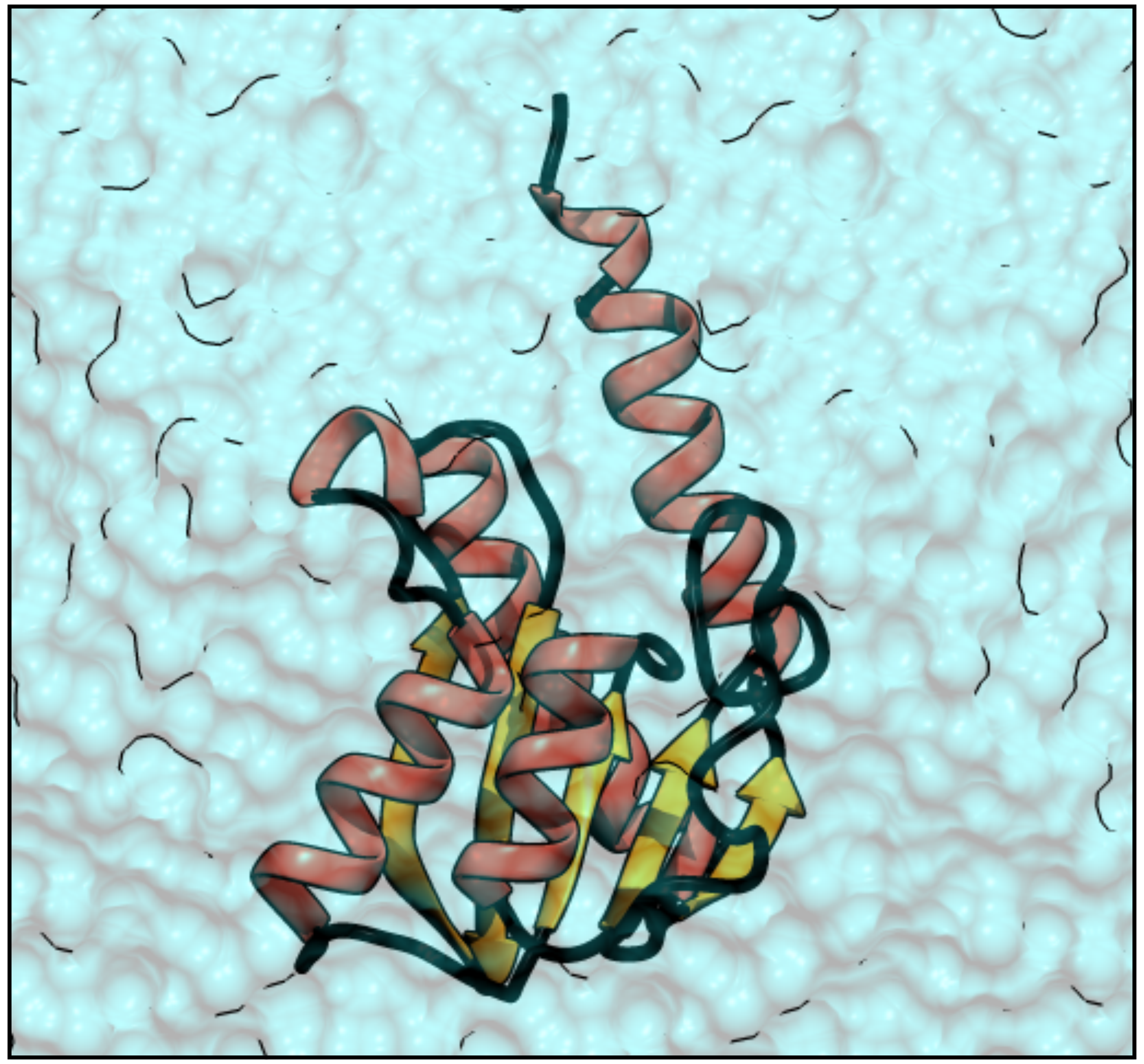


Molecular simulations: Fiction or reality?

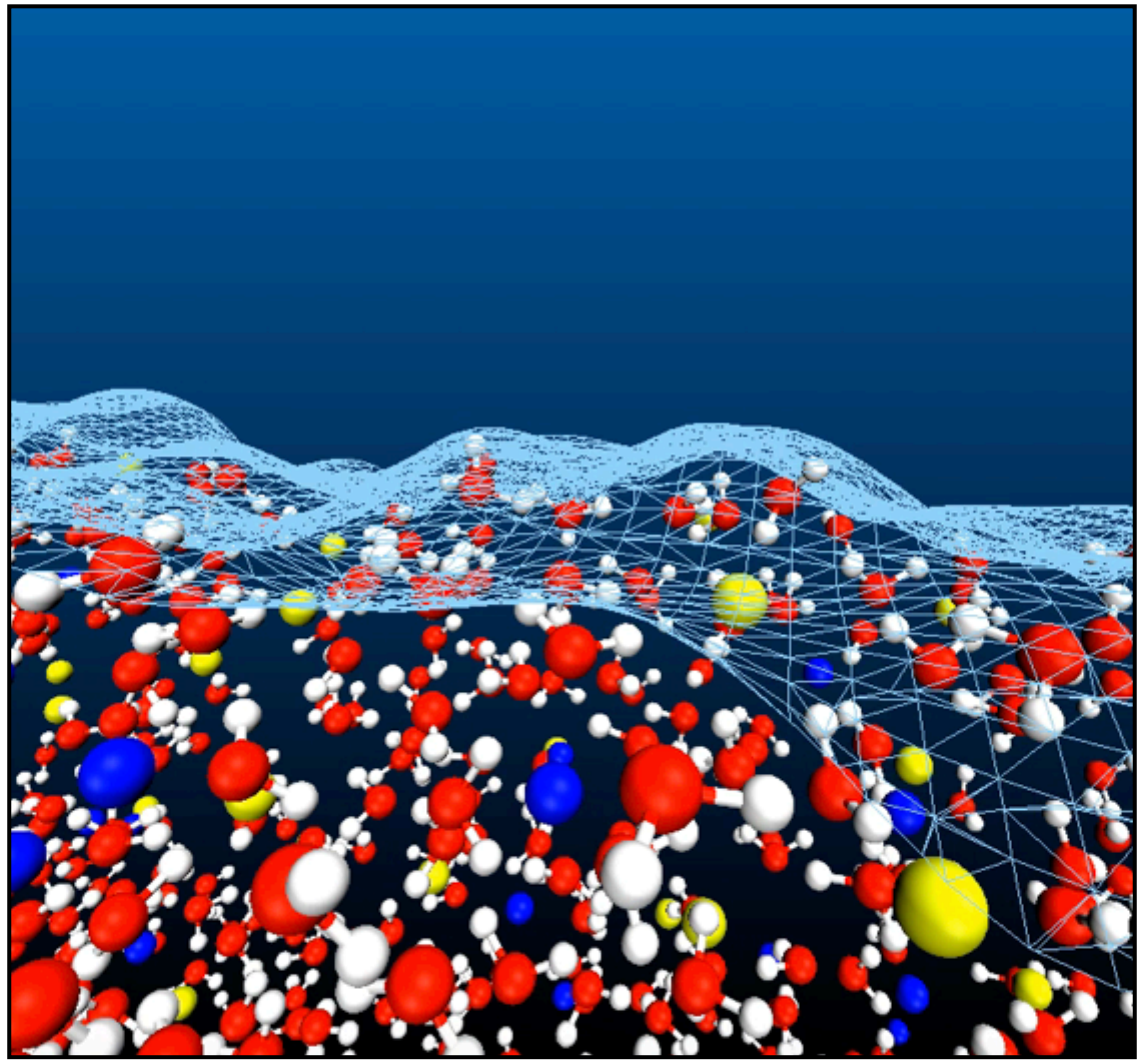
1) Accuracy of the reference data used to build the model

The Importance of Water Cannot Be Overemphasized…

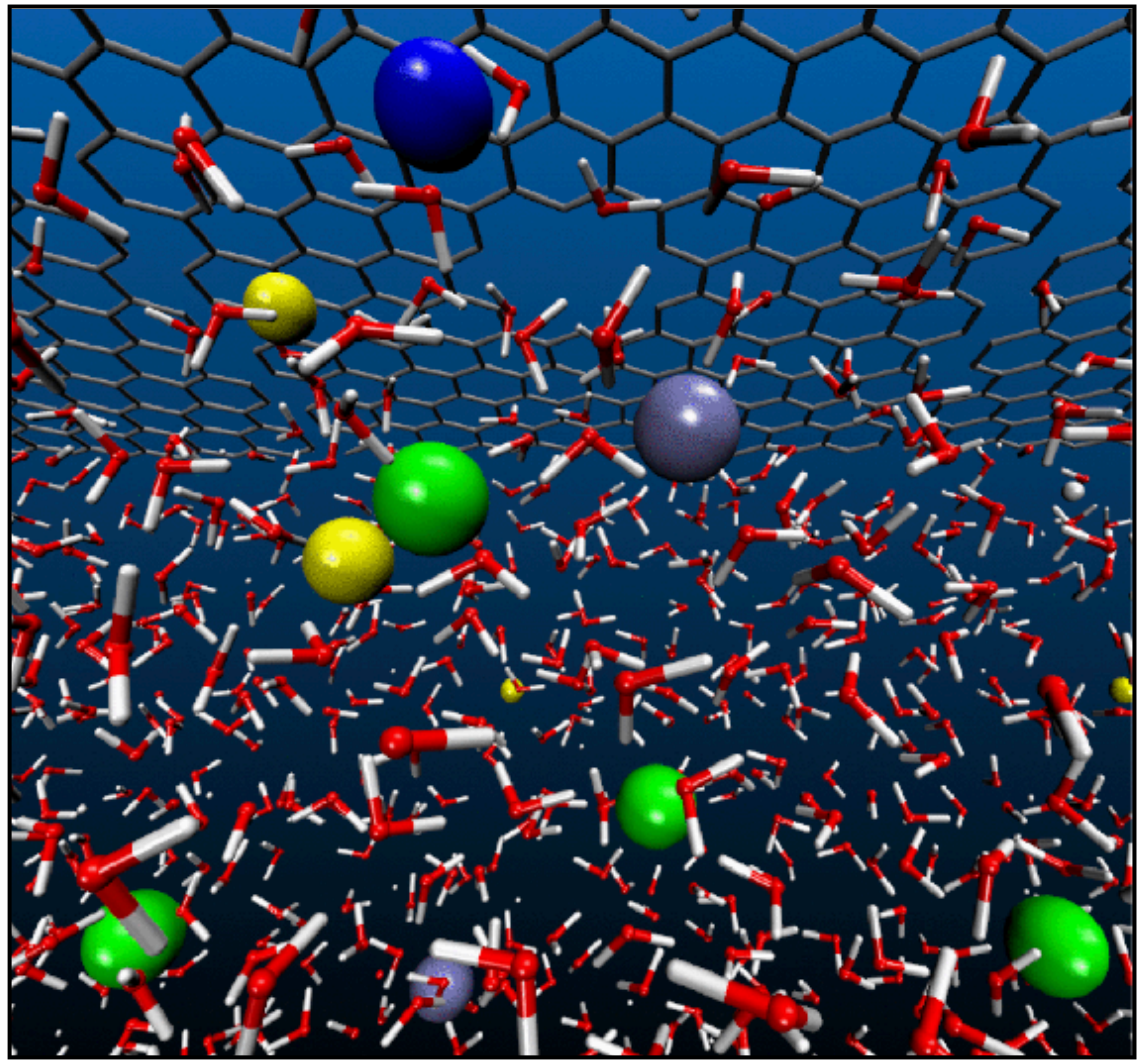
hydration of biomolecules



heterogeneous chemistry



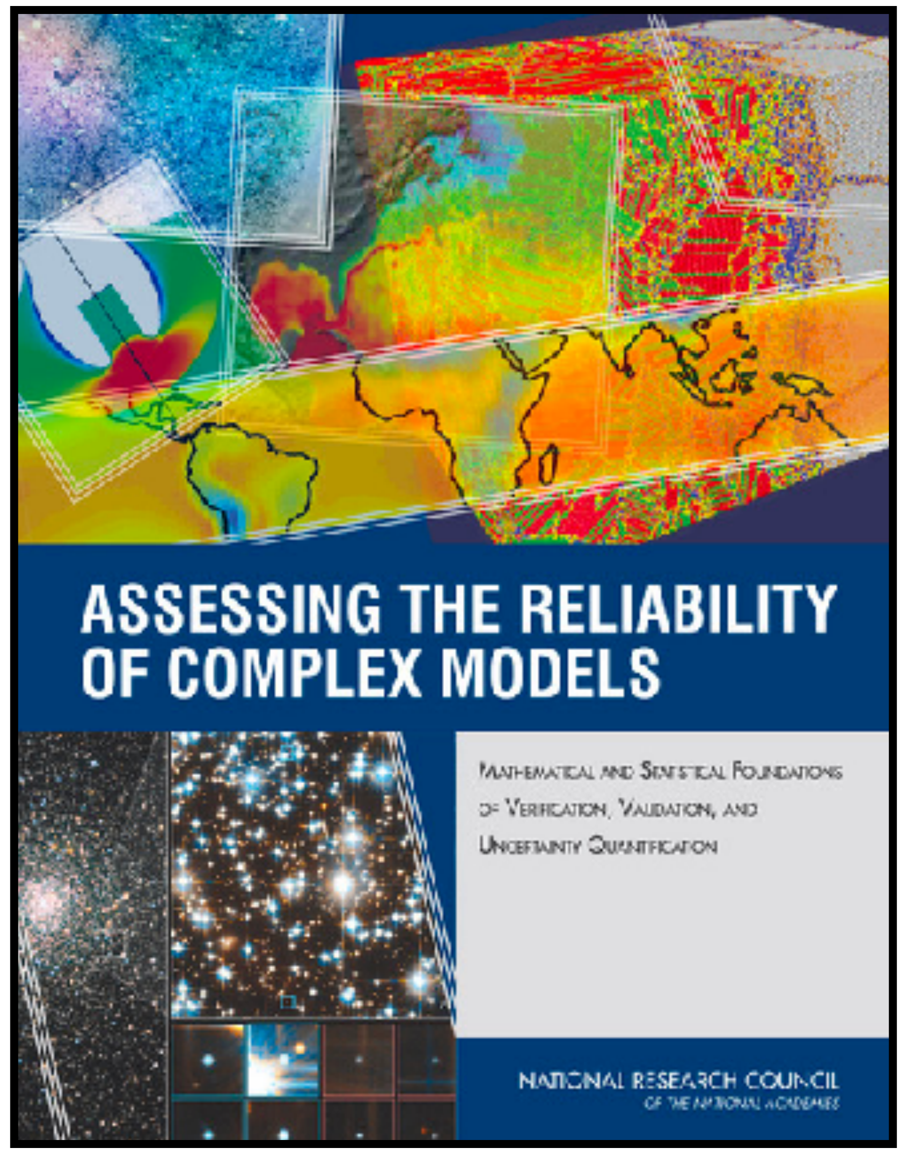
water desalination



Molecular simulations: Fiction or reality?

1) Accuracy of the reference data used to build the model
2) Ability to represent the reference data accurately

Validation and Verification


ASSESSING THE RELIABILITY OF COMPLEX MODELS

Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification

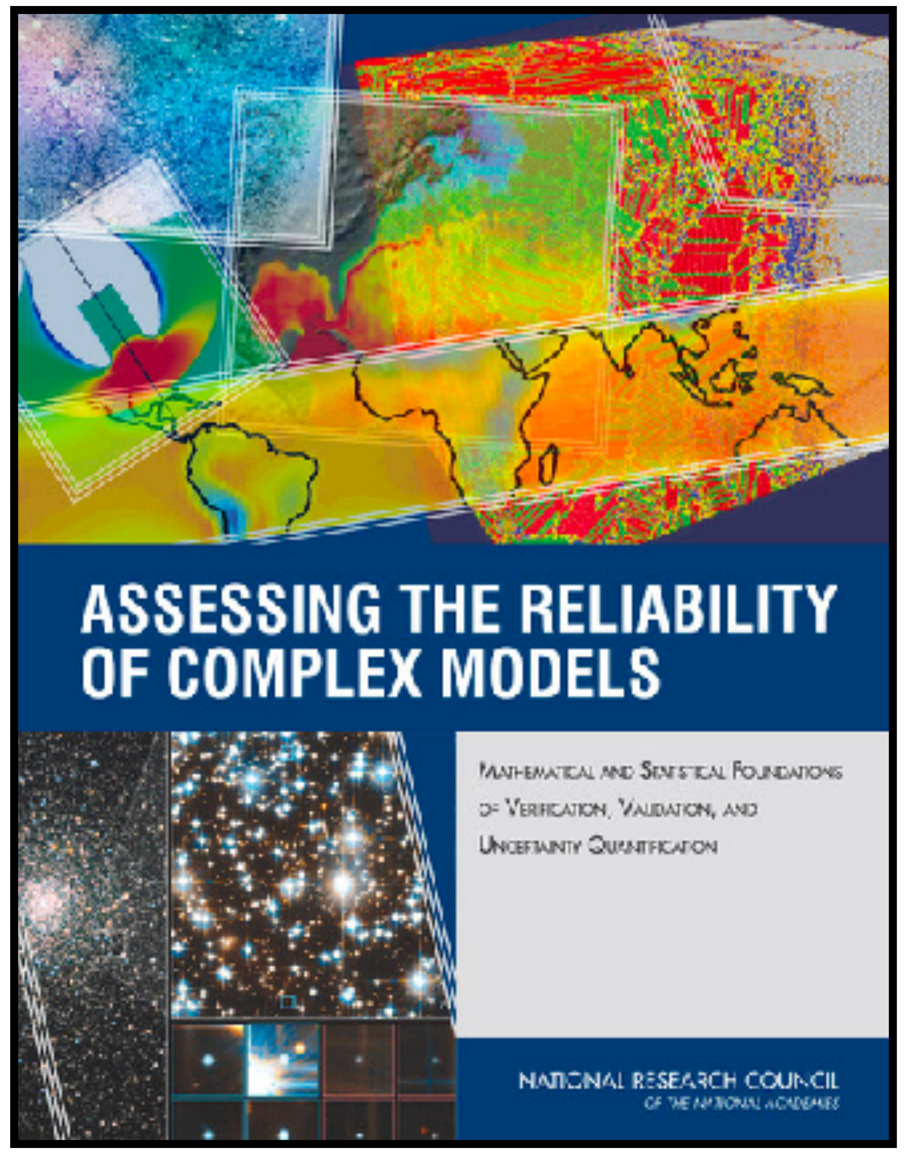
NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

Description

Advances in computing hardware and algorithms have dramatically improved the ability to simulate complex processes computationally. Today's simulation capabilities offer the prospect of addressing questions that in the past could be addressed only by resource-intensive experimentation, if at all.

Assessing the Reliability of Complex Models recognizes the ubiquity of uncertainty in computational estimates of reality and the necessity for its quantification.

Validation and Verification



Description

Advances in computing hardware and algorithms have dramatically improved the ability to simulate complex processes computationally. Today's simulation capabilities offer the prospect of addressing questions that in the past could be addressed only by resource-intensive experimentation, if at all.

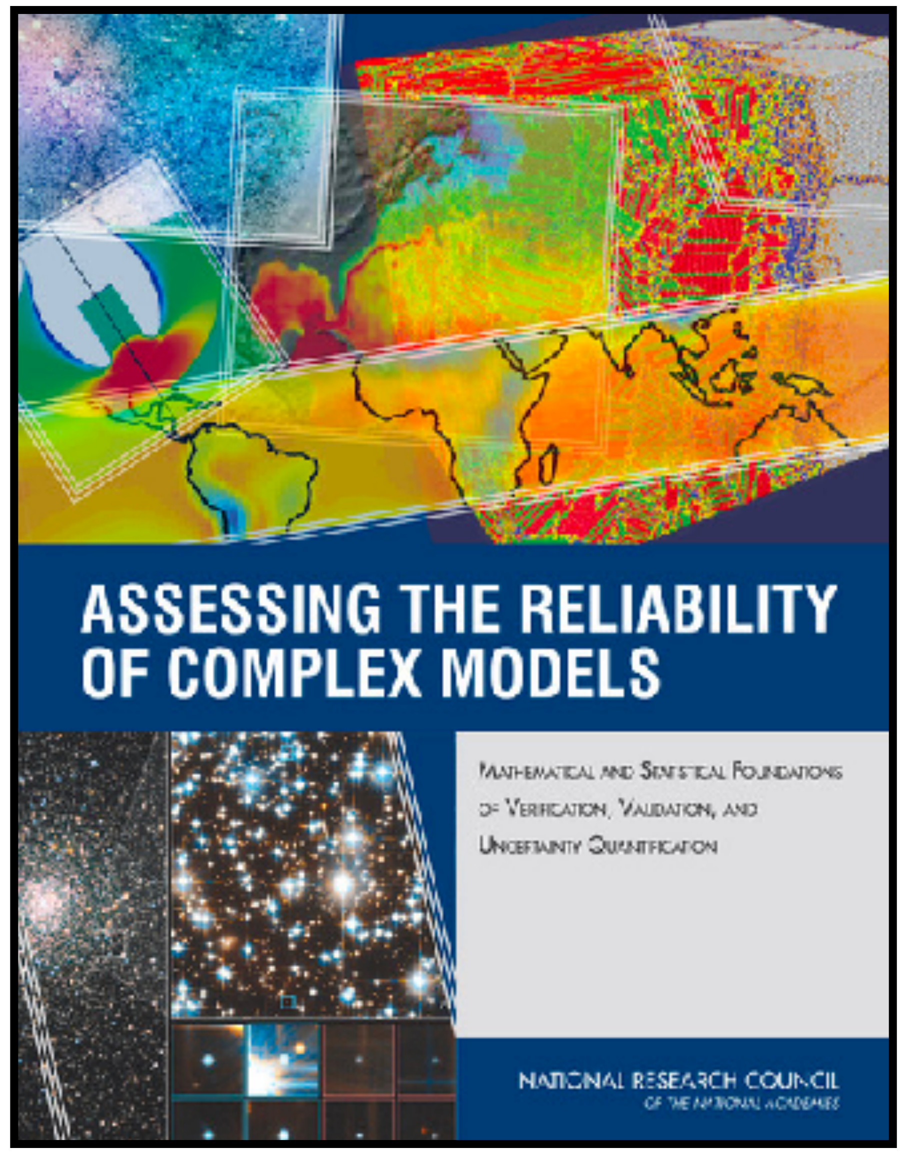
Assessing the Reliability of Complex Models recognizes the ubiquity of uncertainty in computational estimates of reality and the necessity for its quantification.

Molecular simulations

$$A = -k_B T \ln Z$$

free energy

Validation and Verification



ASSESSING THE RELIABILITY OF COMPLEX MODELS

Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

Description

Advances in computing hardware and algorithms have dramatically improved the ability to simulate complex processes computationally. Today's simulation capabilities offer the prospect of addressing questions that in the past could be addressed only by resource-intensive experimentation, if at all.

Assessing the Reliability of Complex Models recognizes the ubiquity of uncertainty in computational estimates of reality and the necessity for its quantification.

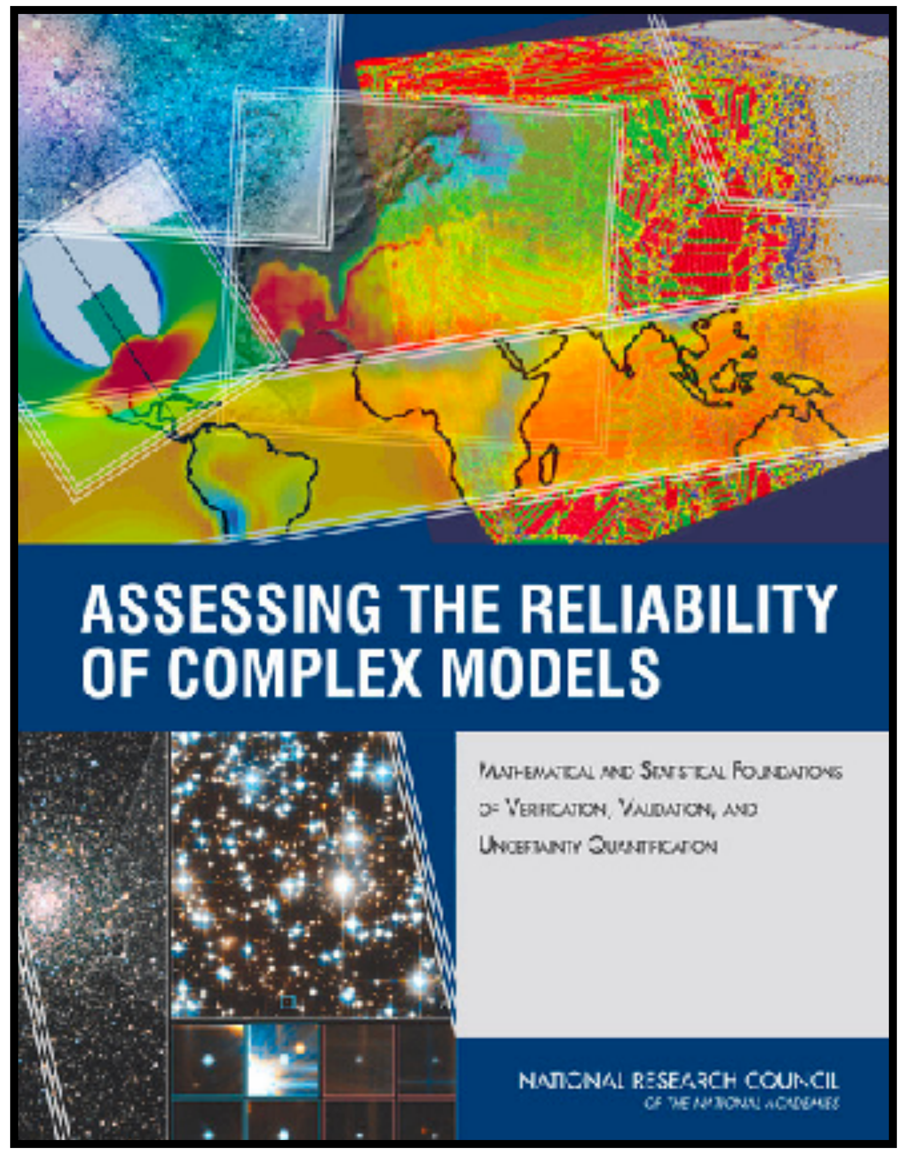
Molecular simulations

$$A = -k_B T \ln Z \quad \Longleftrightarrow \quad Z = \frac{1}{N} \int e^{-\beta H(\mathbf{q}^n, \mathbf{p}^n)} d\mathbf{q}^n d\mathbf{p}^n$$

free energy

classical partition function

Validation and Verification



ASSESSING THE RELIABILITY OF COMPLEX MODELS

Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

Description

Advances in computing hardware and algorithms have dramatically improved the ability to simulate complex processes computationally. Today's simulation capabilities offer the prospect of addressing questions that in the past could be addressed only by resource-intensive experimentation, if at all.

Assessing the Reliability of Complex Models recognizes the ubiquity of uncertainty in computational estimates of reality and the necessity for its quantification.

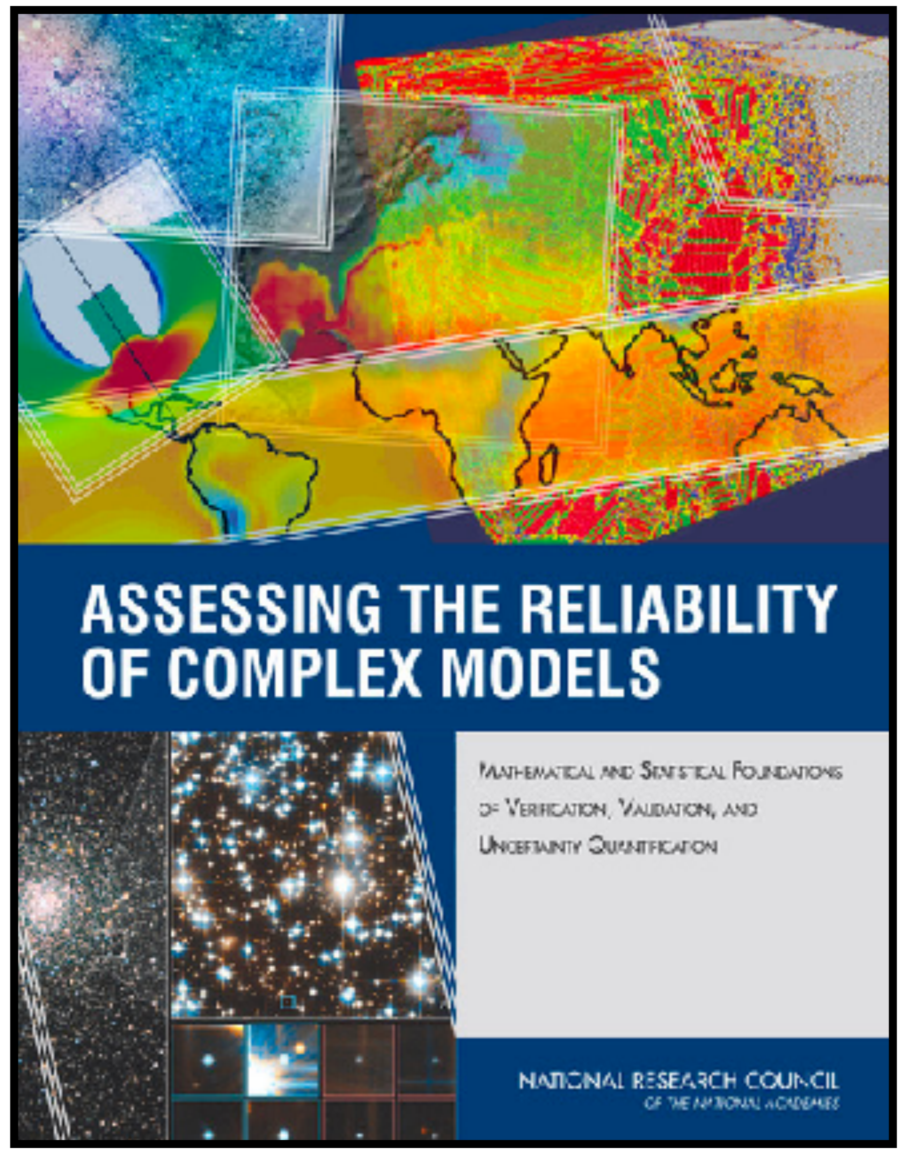
Molecular simulations

$$A = -k_B T \ln Z \quad \Longleftrightarrow \quad Z = \frac{1}{N}\int \langle \mathbf{p}^n, \mathbf{q}^n \,|\, e^{-\beta H(\mathbf{q}^n,\mathbf{p}^n)} \,|\, \mathbf{p}^n, \mathbf{q}^n \rangle$$

free energy

quantum partition function

Validation and Verification



ASSESSING THE RELIABILITY OF COMPLEX MODELS

Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

Description

Advances in computing hardware and algorithms have dramatically improved the ability to simulate complex processes computationally. Today's simulation capabilities offer the prospect of addressing questions that in the past could be addressed only by resource-intensive experimentation, if at all.

Assessing the Reliability of Complex Models recognizes the ubiquity of uncertainty in computational estimates of reality and the necessity for its quantification.

Molecular simulations

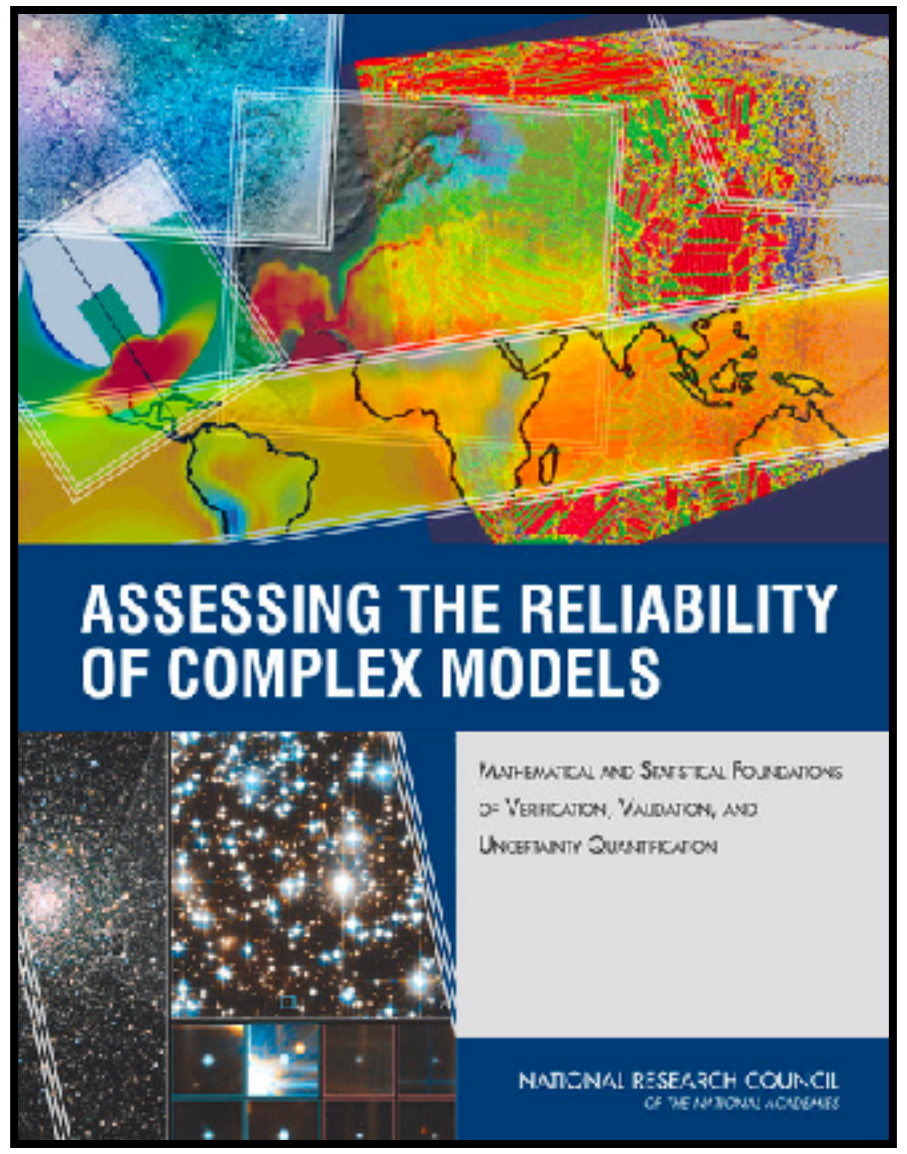
$$A = -k_B T \ln Z \quad \Longleftrightarrow \quad Z = \frac{1}{N}\int \langle \mathbf{p}^n, \mathbf{q}^n \,|\, e^{-\beta H(\mathbf{q}^n,\mathbf{p}^n)} \,|\, \mathbf{p}^n, \mathbf{q}^n \rangle \quad \Longrightarrow$$

Molecular properties & mechanisms

free energy

quantum partition function

Validation and Verification



Description

Advances in computing hardware and algorithms have dramatically improved the ability to simulate complex processes computationally. Today's simulation capabilities offer the prospect of addressing questions that in the past could be addressed only by resource-intensive experimentation, if at all.

Assessing the Reliability of Complex Models recognizes the ubiquity of uncertainty in computational estimates of reality and the necessity for its quantification.

Molecular simulations

$$A = -k_B T \ln Z \quad \Leftrightarrow \quad Z = \frac{1}{N}\int \langle \mathbf{p}^n, \mathbf{q}^n \,|\, e^{-\beta \overset{\text{molecular Hamiltonian}}{\boxed{H(\mathbf{q}^n,\mathbf{p}^n)}}} \,|\, \mathbf{p}^n, \mathbf{q}^n \rangle \quad \Rightarrow$$

Molecular properties & mechanisms

free energy

quantum partition function

Molecular Hamiltonian: Molecular Interactions

FP, Acc. Chem. Res. 49*, 1844 (2016)*

Molecular Interactions: Molecular Mechanics

Force fields

$$V(\vec{R}^N) = \sum_{bonds} \frac{k_b}{2}(r_b - r_{b,o})^2 + \sum_{angles} \frac{k_a}{2}(\theta_a - \theta_{a,o})^2 + \sum_{torsion} \left[\sum_n \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \right]$$

$$+ \sum_{i<j} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}} \right\}$$

1948: Hill, J. Chem. Phys. 14, 465 (1946) / Westheimer & Mayer, J. Chem. Phys. 14, 733 (1946)

1968: Lifson & Warshel, J. Chem. Phys. 49, 5116 (1968)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: Molecular Mechanics

Force fields

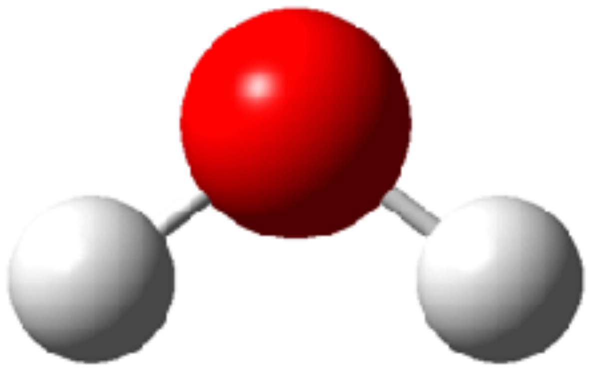
$$V(\vec{R}^N) = \boxed{\sum_{bonds} \frac{k_b}{2}(r_b - r_{b,o})^2} + \sum_{angles} \frac{k_a}{2}(\theta_a - \theta_{a,o})^2 + \sum_{torsion} \left[\sum_n \frac{V_n}{2}(1 + \cos(n\omega - \gamma)) \right]$$

$$+ \sum_{i<j} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}} \right\}$$

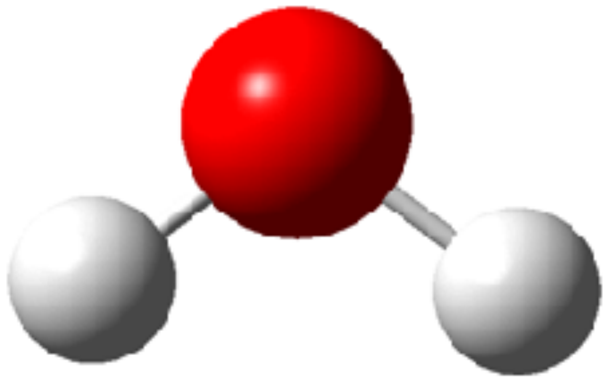
1948: Hill, J. Chem. Phys. 14, 465 (1946) / Westheimer & Mayer, J. Chem. Phys. 14, 733 (1946)

1968: Lifson & Warshel, J. Chem. Phys. 49, 5116 (1968)

symmetric stretch



asymmetric stretch



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: Molecular Mechanics

Force fields

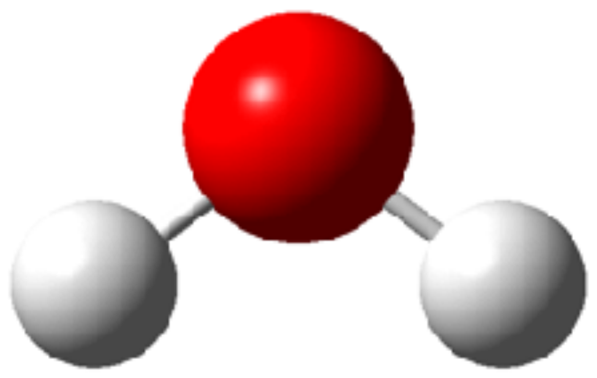
$$V(\vec{R}^N) = \boxed{\sum_{bonds} \frac{k_b}{2}(r_b - r_{b,o})^2} + \boxed{\sum_{angles} \frac{k_a}{2}(\theta_a - \theta_{a,o})^2} + \sum_{torsion} \left[\sum_n \frac{V_n}{2}\left(1 + \cos\left(n\omega - \gamma\right)\right) \right]$$

$$+ \sum_{i<j} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}} \right\}$$

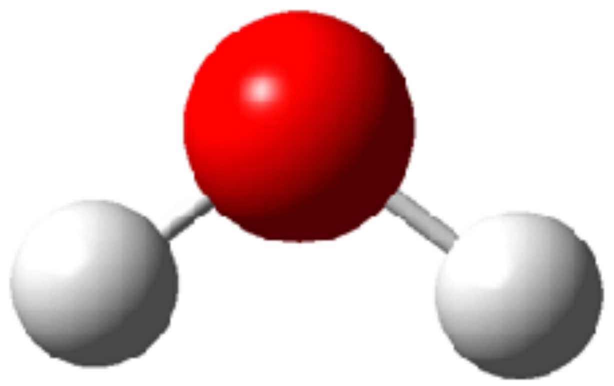
1948: Hill, J. Chem. Phys. 14, 465 (1946) / Westheimer & Mayer, J. Chem. Phys. 14, 733 (1946)

1968: Lifson & Warshel, J. Chem. Phys. 49, 5116 (1968)

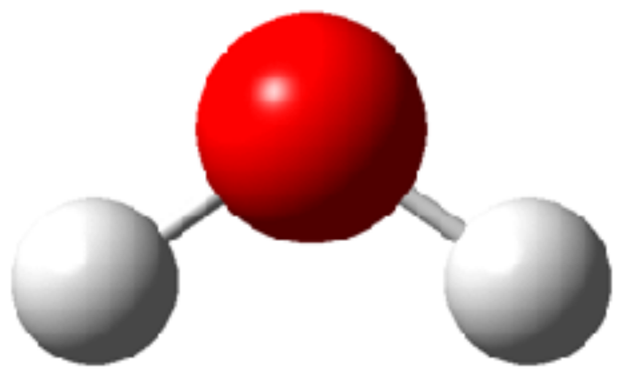
symmetric stretch



asymmetric stretch



bend



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: Molecular Mechanics

Force fields

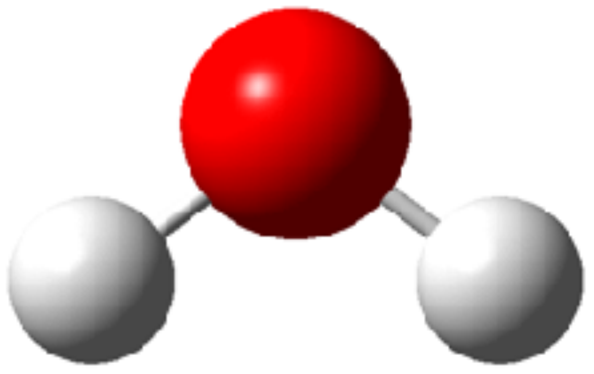
$$V(\vec{R}^N) = \boxed{\sum_{bonds} \frac{k_b}{2}(r_b - r_{b,o})^2} + \boxed{\sum_{angles} \frac{k_a}{2}(\theta_a - \theta_{a,o})^2} + \sum_{torsion}\left[\sum_n \frac{V_n}{2}(1+\cos(n\omega-\gamma))\right]$$

$$+\boxed{\sum_{i<j}\left\{4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}}\right\}}$$

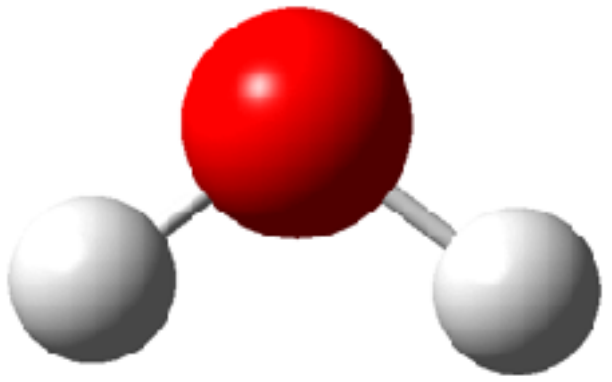
1948: Hill, J. Chem. Phys. 14, 465 (1946) / Westheimer & Mayer, J. Chem. Phys. 14, 733 (1946)

1968: Lifson & Warshel, J. Chem. Phys. 49, 5116 (1968)

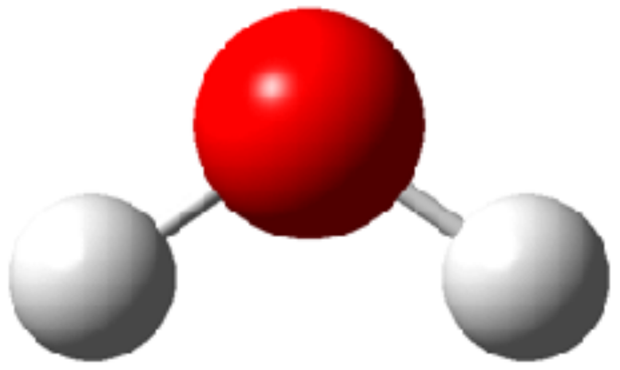
symmetric stretch



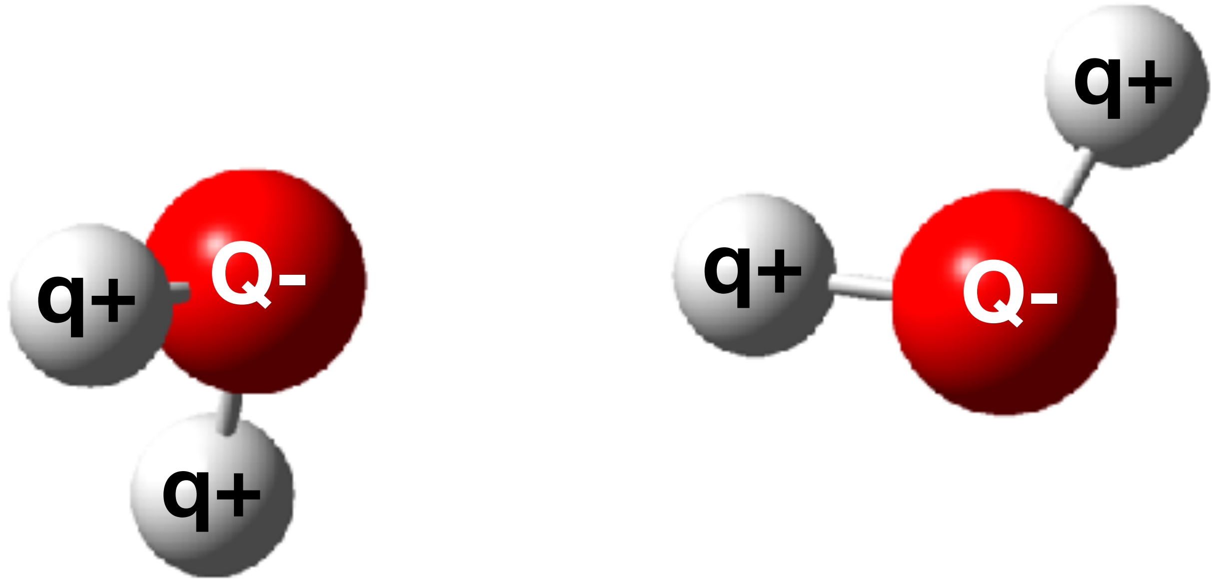
asymmetric stretch



bend



intermolecular interactions



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: Quantum Mechanics

Electronic structure (ab initio) methods

$$\hat{H}\Psi = E\Psi$$

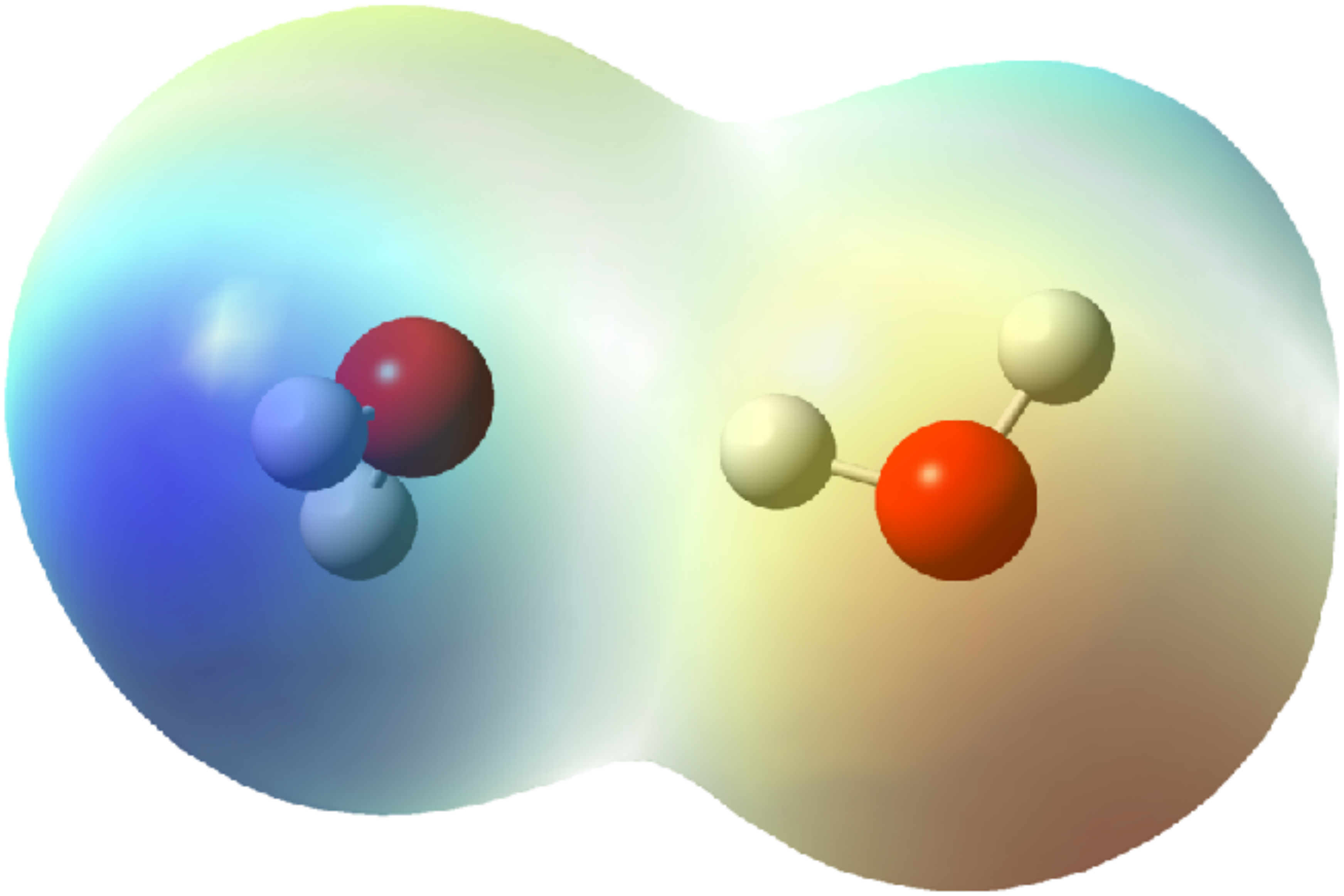
$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{xc}[\rho]$$

1930: Hartree-Fock equations

1951: Hartree-Fock-Roothan equations

1965: Kohn & Sham equations

1966: Coupled-Cluster equations



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: Quantum Mechanics

Electronic structure (ab initio) methods

$$\hat{H}\Psi = E\Psi$$

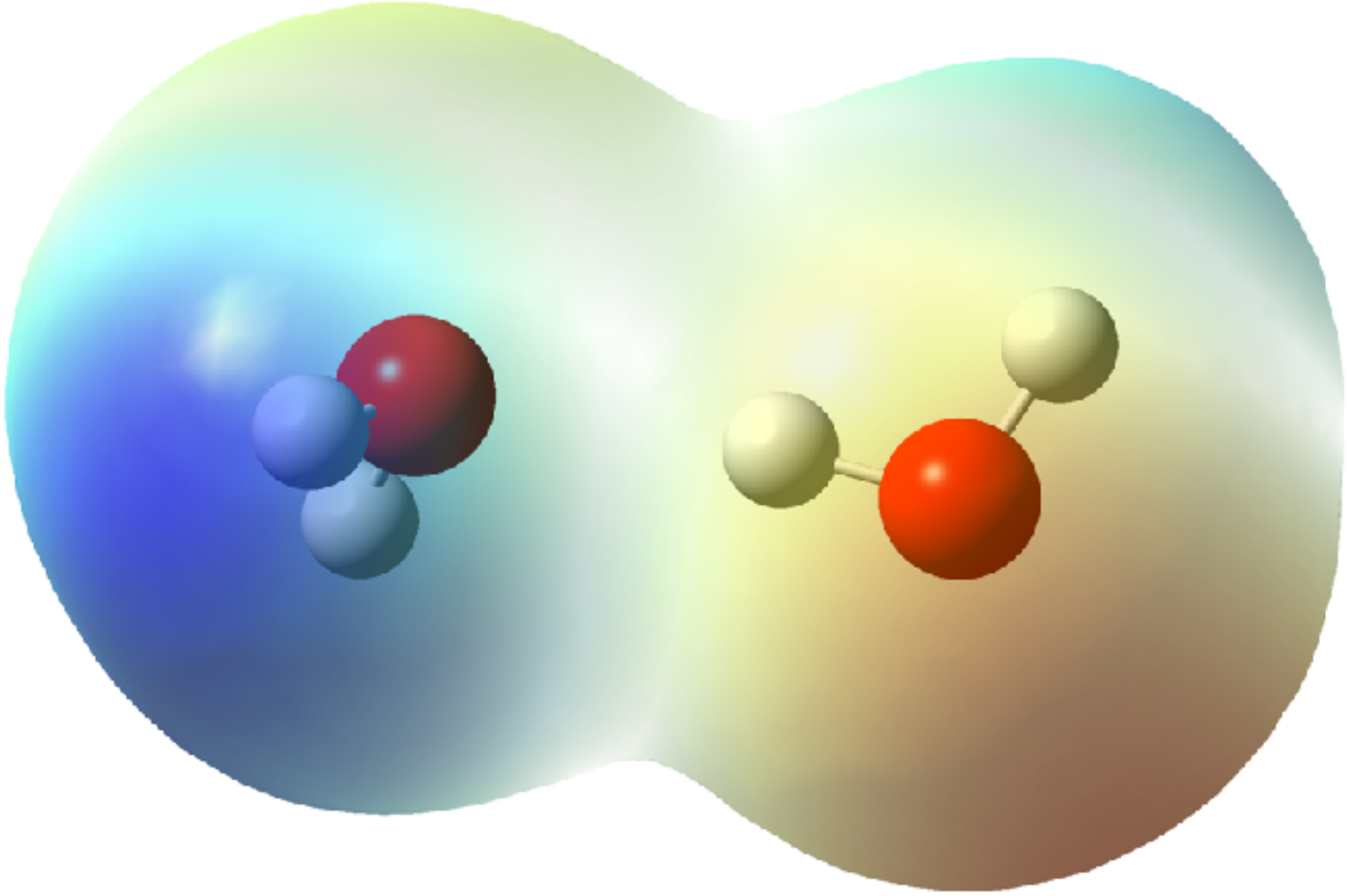
$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{xc}[\rho]$$

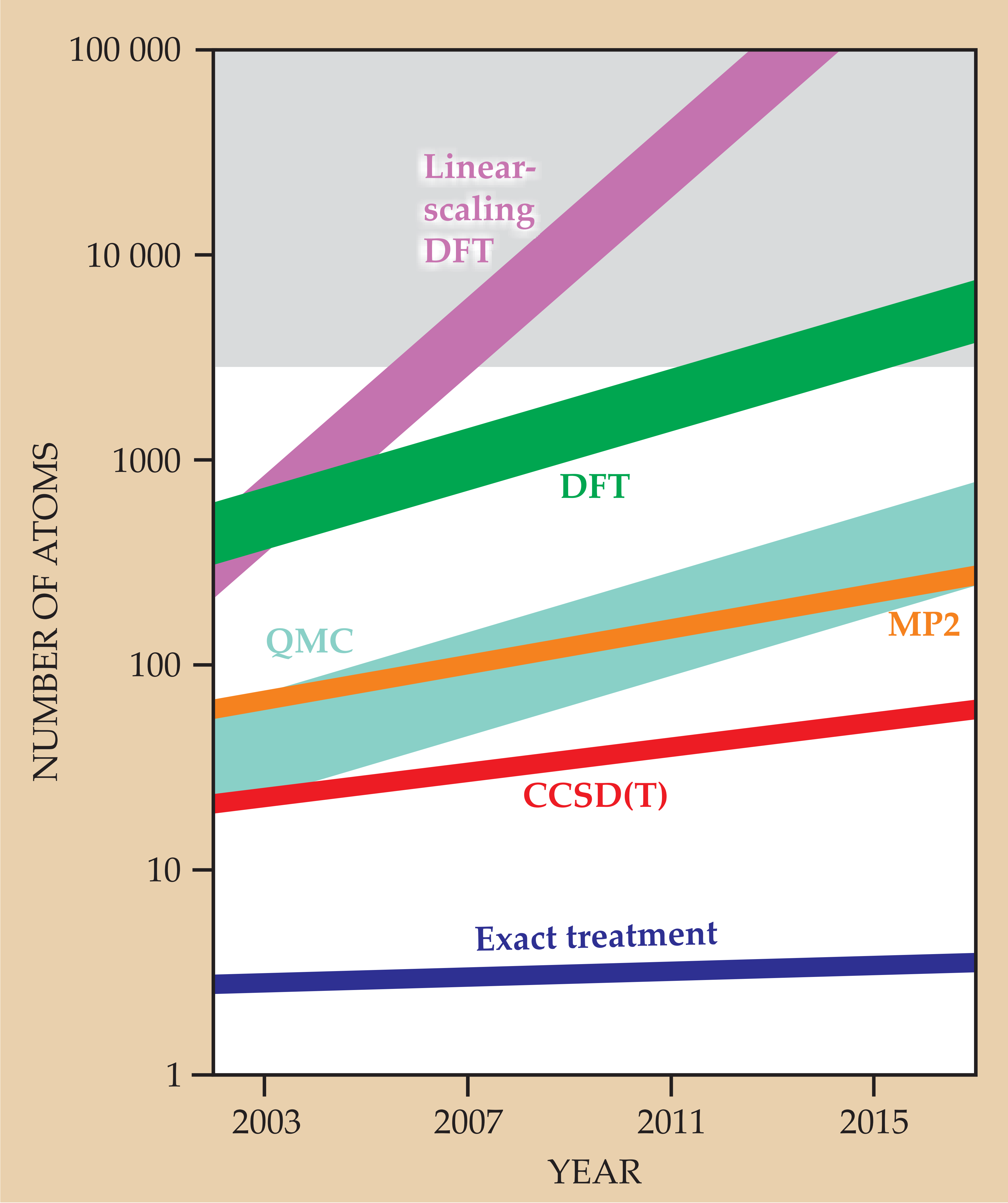
1930: Hartree-Fock equations

1951: Hartree-Fock-Roothan equations

1965: Kohn & Sham equations

1966: Coupled-Cluster equations





Head-Gordon & Artacho, Phys. Today <u>*61*</u>*, 58 (2008)*

FP, Acc. Chem. Res. <u>*49*</u>*, 1844 (2016)*

Molecular Interactions: Quantum Mechanics

Electronic structure (ab initio) methods

$$\hat{H}\Psi = E\Psi$$

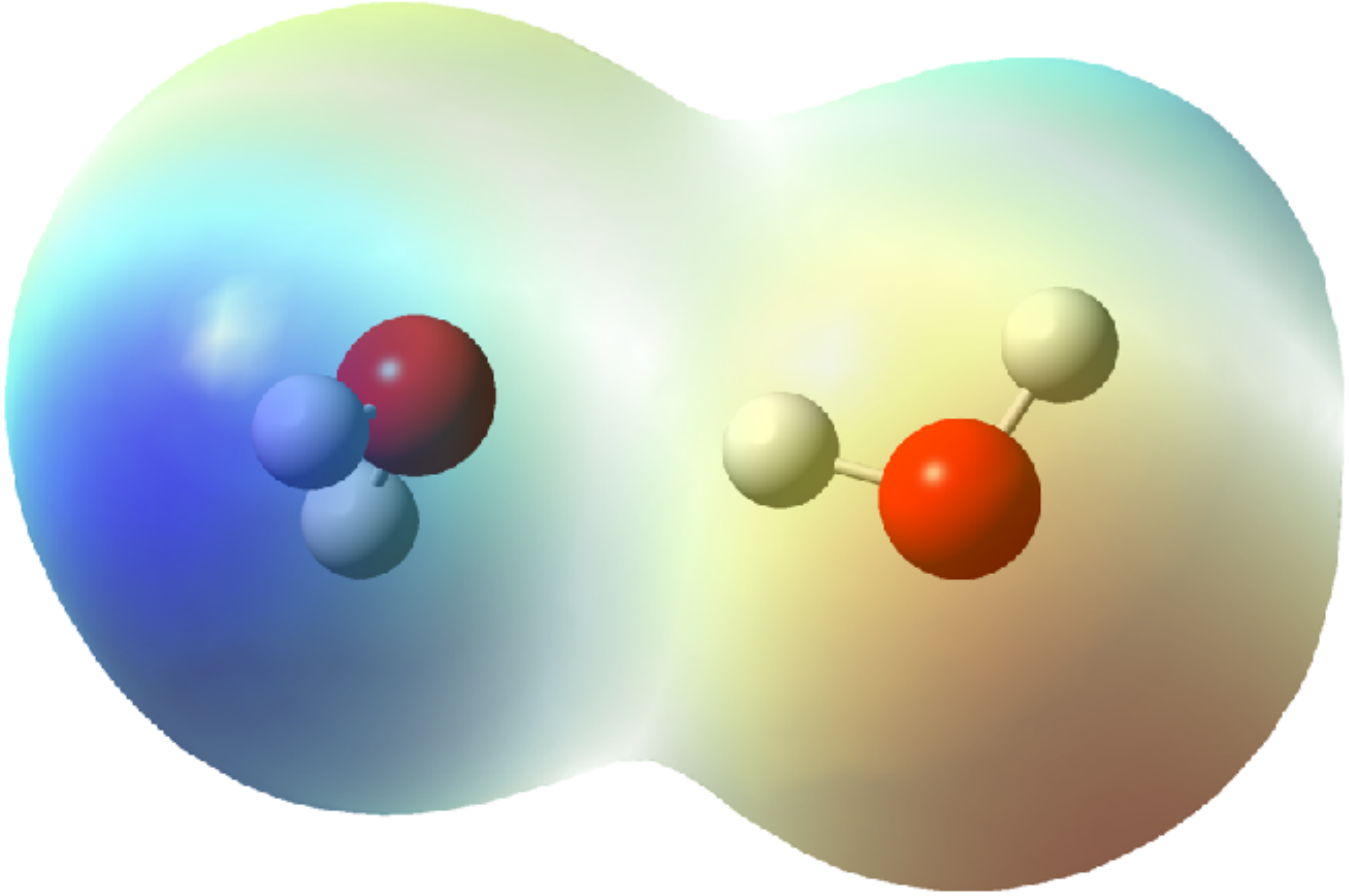
$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{xc}[\rho]$$

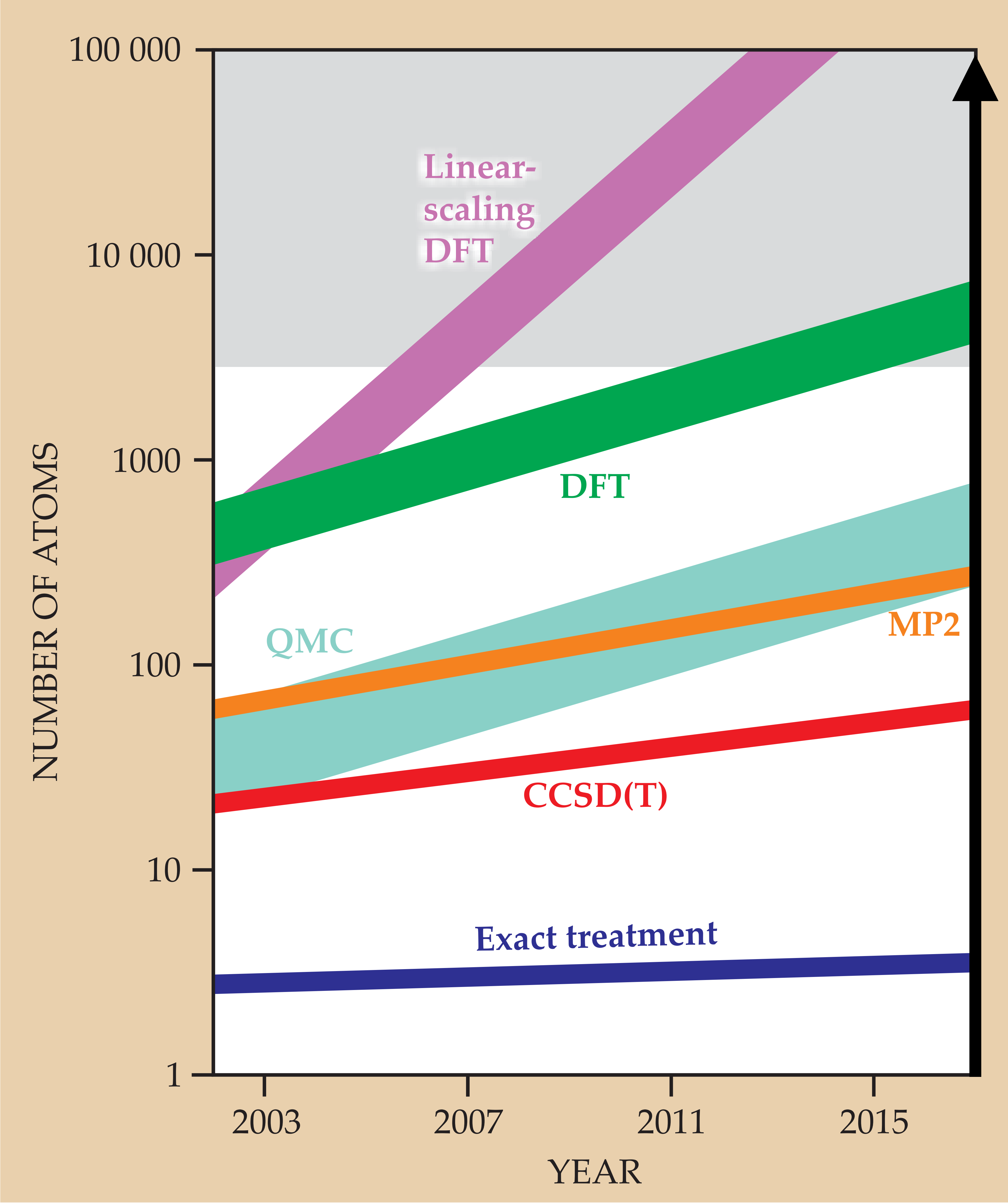
1930: Hartree-Fock equations

1951: Hartree-Fock-Roothan equations

1965: Kohn & Sham equations

1966: Coupled-Cluster equations





Head-Gordon & Artacho, Phys. Today 61, 58 (2008)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions

Molecular mechanics = Force fields

$$V(\vec{R}^N) = \sum_{bonds} \frac{k_b}{2}(r_b - r_{b,o})^2 + \sum_{angles} \frac{k_a}{2}(\theta_a - \theta_{a,o})^2 + \sum_{torsion} \left[\sum_n \frac{V_n}{2}(1 + \cos(n\omega - \gamma)) \right]$$

$$+ \sum_{i<j} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}} \right\}$$

1948: Hill, J. Chem. Phys. 14, 465 (1946) / Westheimer & Mayer, J. Chem. Phys. 14, 733 (1946)

1968: Lifson & Warshel, J. Chem. Phys. 49, 5116 (1968)

Quantum mechanics = Ab initio

$$\hat{H}\Psi = E\Psi$$

$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{xc}[\rho]$$

1930: Hartree-Fock equations

1951: Hartree-Fock-Roothan equations

1965: Kohn & Sham equations

1966: Coupled-Cluster equations

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions

Molecular mechanics = Force fields: Not accurate, but cheap

$$V(\vec{R}^N) = \sum_{bonds} \frac{k_b}{2}(r_b - r_{b,o})^2 + \sum_{angles} \frac{k_a}{2}(\theta_a - \theta_{a,o})^2 + \sum_{torsion} \left[\sum_n \frac{V_n}{2}(1 + \cos(n\omega - \gamma)) \right]$$

$$+ \sum_{i<j} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}} \right\}$$

1948: Hill, J. Chem. Phys. 14, 465 (1946) / Westheimer & Mayer, J. Chem. Phys. 14, 733 (1946)

1968: Lifson & Warshel, J. Chem. Phys. 49, 5116 (1968)

Quantum mechanics = Ab initio

$$\hat{H}\Psi = E\Psi$$

$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{xc}[\rho]$$

1930: Hartree-Fock equations

1951: Hartree-Fock-Roothan equations

1965: Kohn & Sham equations

1966: Coupled-Cluster equations

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions

Molecular mechanics = Force fields: Not accurate, but cheap

$$V(\vec{R}^N) = \sum_{bonds} \frac{k_b}{2}(r_b - r_{b,o})^2 + \sum_{angles} \frac{k_a}{2}(\theta_a - \theta_{a,o})^2 + \sum_{torsion} \left[\sum_n \frac{V_n}{2}(1 + \cos(n\omega - \gamma)) \right]$$

$$+ \sum_{i<j} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}} \right\}$$

1948: Hill, J. Chem. Phys. 14, 465 (1946) / Westheimer & Mayer, J. Chem. Phys. 14, 733 (1946)

1968: Lifson & Warshel, J. Chem. Phys. 49, 5116 (1968)

Quantum mechanics = Ab initio: Accurate, but expensive

$$\hat{H}\Psi = E\Psi$$

$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{xc}[\rho]$$

1930: Hartree-Fock equations

1951: Hartree-Fock-Roothan equations

1965: Kohn & Sham equations

1966: Coupled-Cluster equations

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions

Molecular mechanics = Force fields: Not accurate, but cheap

$$V(\vec{R}^N) = \sum_{bonds} \frac{k_b}{2}(r_b - r_{b,o})^2 + \sum_{angles} \frac{k_a}{2}(\theta_a - \theta_a$$

$$\frac{V_n}{2}(1 + \cos(n\omega - \gamma))\Bigg]$$

$$+\sum_{i<j}\left\{4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] +$$

$$R_{ij}\Bigg\}$$

1948: Hill, J. Chem. Phys. 14, 465 (1946) / Westheimer & Mayer, J. Ch

. 14, 733 (1946)

1968: Lifson & Warshel, J. Chem. Phys. 49, 5116 (1968)

?

Quantum mechanics = Ab initio: Acc , but expensive

$$\hat{H}\Psi = E\Psi$$

$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{xc}[\rho]$$

1930: Hartree-Fock equations

1951: Hartree-Fock-Roothan equations

1965: Kohn & Sham equations

1966: Coupled-Cluster equations

FP, Acc. Chem. Res. 49, 1844 (2016)

Machine Learning

Machine Learning

Sumpter & Noid, Chem. Phys. Lett.192, 455 (1992)

Lorenz, Gross & Scheffler, Chem. Phys. Lett. 395, 210 (2004)

Manzhos & Carrington, J. Chem. Phys. 125, 194105 (2006)

Behler & Parrinello, Phys. Rev. Lett. 98, 146401 (2007)

Bartók, Payne, Kondor, Csányi, Phys. Rev. Lett. 104, 136403 (2010)

Rupp, Tkatchenko, K.-R. Müller & von Lilienfeld, Phys. Rev. Lett. 108, 058301 (2012)

Hansen, Montavon, Biegler, Fazli, Rupp, Scheffler, von Lilienfeld, Tkatchenko & Müller, J. Chem. Theory Comput. 9, 3404 (2013)

Smith, Isayev & Roitberg, Chem. Sci. 8, 3192 (2017)

Chmiela, Tkatchenko, Sauceda, Poltavsky, Schütt & Müller, Sci. Adv. 3, e1603015 (2017)

Yao, Herr & Parkhill, J. Chem. Phys. 146, 014106 (2017)

Schütt, Arbabzadah, Chmiela, Müller &. Tkatchenko, Nat. Commun. 8, 13890 (2017)

Chmiela, Sauceda, Müller &. Tkatchenko, Nat. Commun. 9, 3887 (2018)

Faber, Christensen, Huang & von Lilienfeld, J. Chem. Phys. 148, 241717 (2018)

Schütt, Sauceda, Kindermans, Tkatchenko & Müller, J. Chem. Phys. 148, 241722 (2018)

Grisafi, Wilkins, Csányi & Ceriotti, Phys. Rev. Lett. 120, 036002 (2018)

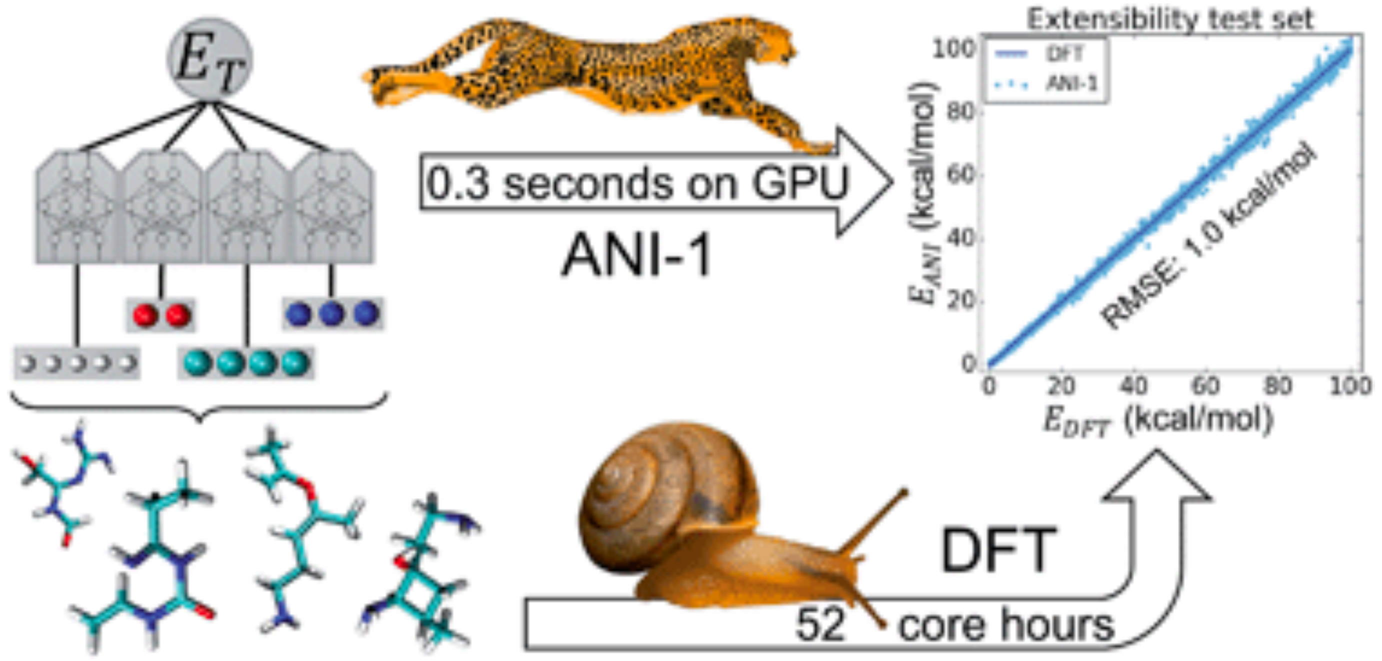
Zhang, Han, Wang, Car & E, Phys. Rev. Lett. 120, 143001 (2018)

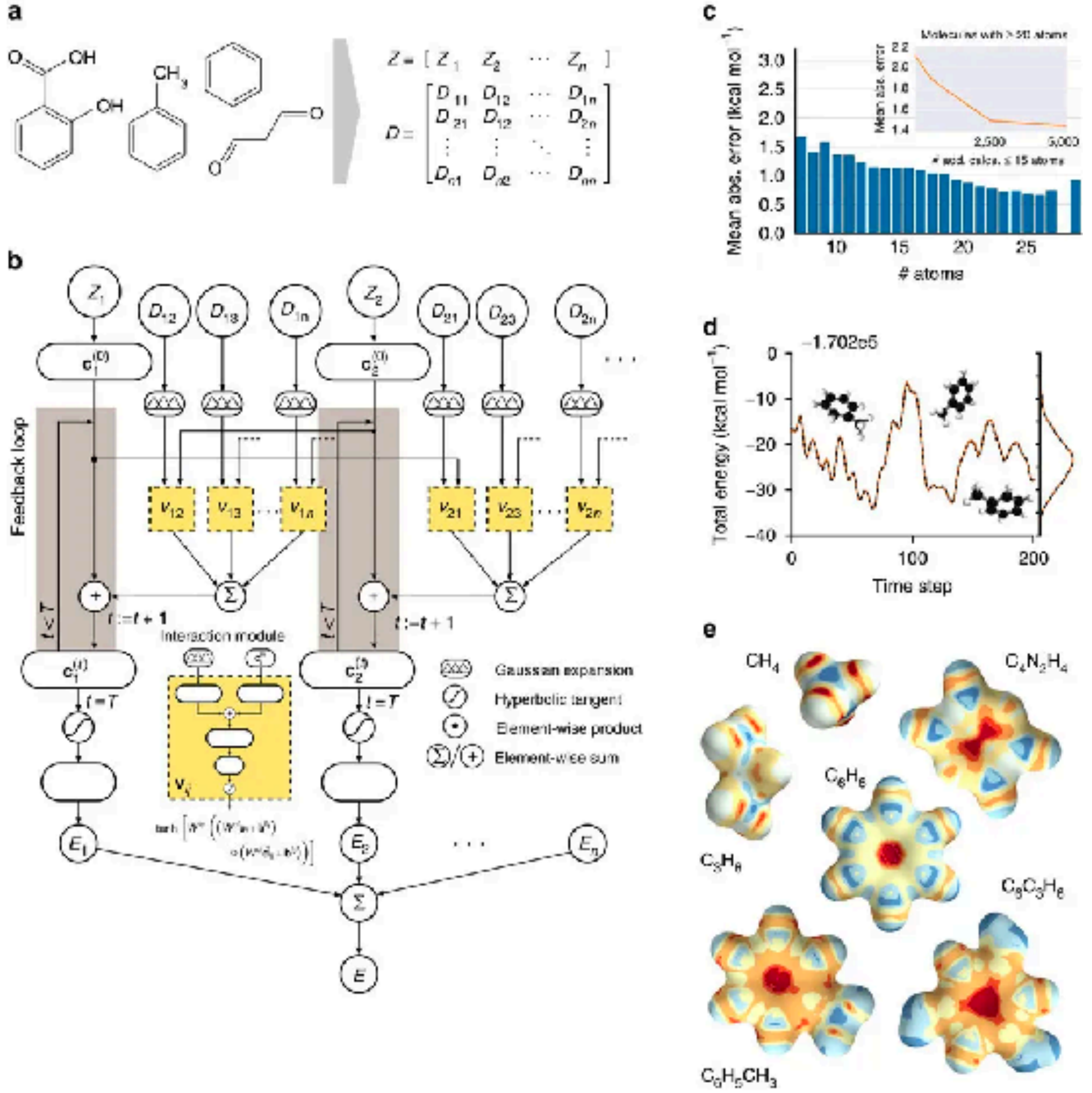
Collins, Gordon, von Lilienfeld & D. J. Yaron, J. Chem. Phys. 148, 241718 (2018)

Schütt, Kessel, Gastegger, Nicoli, Tkatchenko & Müller, J. Chem. Theory Comput. 15, 448 (2019)

Unke & Meuwly, J. Chem. Theory Comput. 15, 3678 (2019)

…





Molecular Interactions: A Many-Body Perspective

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

1 molecule at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

1 molecule at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

1 molecule at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

1 molecule at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

1 molecule at a time

FP, *Acc. Chem. Res.* 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

2 molecules at a time

FP, *Acc. Chem. Res.* 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

2 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

2 molecules at a time

FP, *Acc. Chem. Res.* 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

2 molecules at a time

FP, Acc. Chem. Res. 49*, 1844 (2016)*

Molecular Interactions: A Many-Body Perspective

2 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

3 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

3 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

3 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

3 molecules at a time

FP, *Acc. Chem. Res.* 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

3 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

4 molecules at a time

FP, *Acc. Chem. Res.* 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

4 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

5 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

6 molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

N molecules at a time

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$\begin{aligned} E_N = & \sum_{i}^{N} V^{1B}(i) \\ & + \sum_{i<j}^{N} V^{2B}(i,j) \\ & + \sum_{i<j<k}^{N} V^{3B}(i,j,k) \\ & + \dots \\ & + V^{NB}(1,\dots,N) \end{aligned}$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

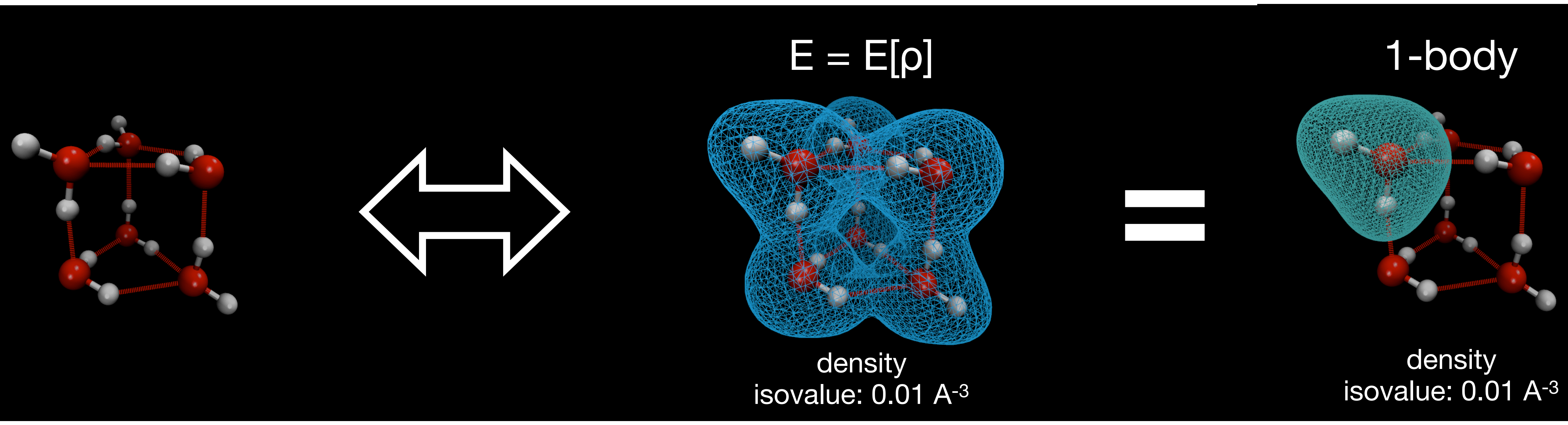
Molecular Interactions: A Many-Body Perspective

$$\begin{aligned} E_N = & \sum_{i}^{N} V^{1B}(i) \\ & + \sum_{i<j}^{N} V^{2B}(i,j) \\ & + \sum_{i<j<k}^{N} V^{3B}(i,j,k) \\ & + \dots \\ & + V^{NB}(1,\dots,N) \end{aligned}$$

1-body

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$
\begin{aligned}
E_N = &\sum_{i}^{N} V^{1B}(i) \\
&+ \sum_{i<j}^{N} V^{2B}(i,j) \\
&+ \sum_{i<j<k}^{N} V^{3B}(i,j,k) \\
&+ \dots \\
&+ V^{NB}(1,\dots,N)
\end{aligned}
$$

1-body

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

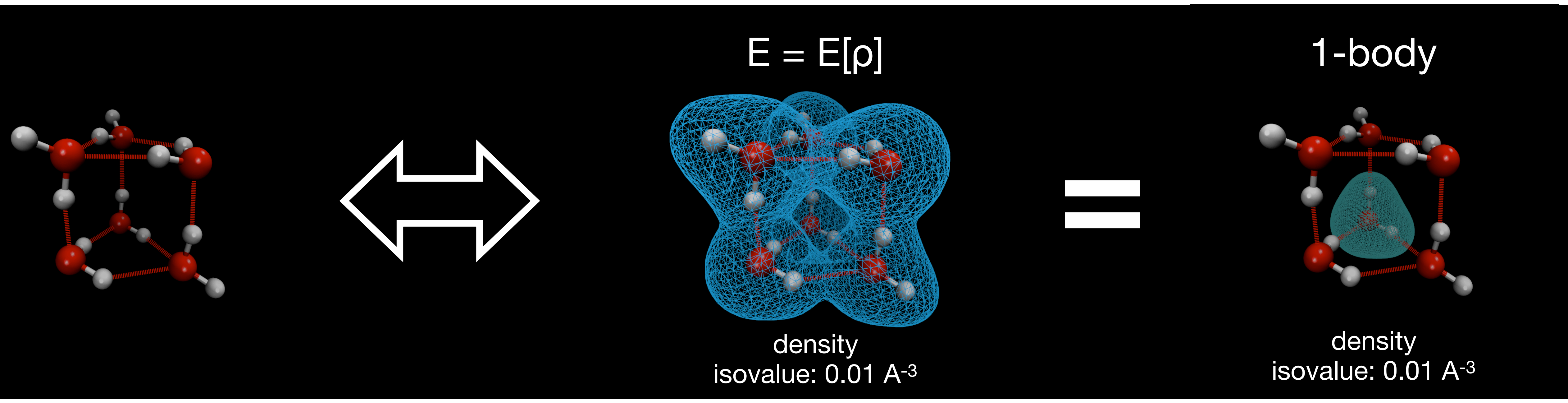
Molecular Interactions: A Many-Body Perspective

$$
\begin{aligned}
E_N = & \sum_{i}^{N} V^{1B}(i) \\
& + \sum_{i<j}^{N} V^{2B}(i,j) \\
& + \sum_{i<j<k}^{N} V^{3B}(i,j,k) \\
& + \ldots \\
& + V^{NB}(1,\ldots,N)
\end{aligned}
$$

1-body

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)



FP, Acc. Chem. Res. 49, 1844 (2016)

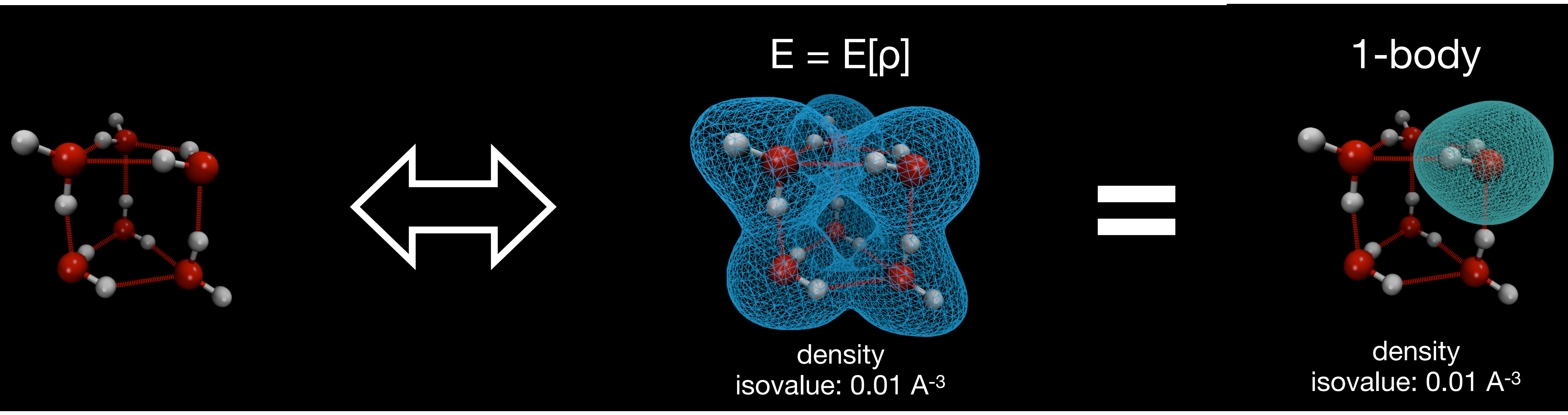
Molecular Interactions: A Many-Body Perspective

$$
\begin{aligned}
E_N = & \sum_{i}^{N} V^{1B}(i) \\
& + \sum_{i<j}^{N} V^{2B}(i,j) \\
& + \sum_{i<j<k}^{N} V^{3B}(i,j,k) \\
& + \dots \\
& + V^{NB}(1,\dots,N)
\end{aligned}
$$

1-body

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$
\begin{aligned}
E_N = & \sum_{i}^{N} V^{1B}(i) \\
& + \sum_{i<j}^{N} V^{2B}(i,j) \\
& + \sum_{i<j<k}^{N} V^{3B}(i,j,k) \\
& + \dots \\
& + V^{NB}(1,\dots,N)
\end{aligned}
$$

1-body

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$\begin{aligned} E_N = & \sum_{i}^{N} V^{1B}(i) \\ & + \sum_{i<j}^{N} V^{2B}(i,j) \\ & + \sum_{i<j<k}^{N} V^{3B}(i,j,k) \\ & + \dots \\ & + V^{NB}(1,\dots,N) \end{aligned}$$

1-body

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$E_N = \sum_{i}^{N} V^{1B}(i)$$

1-body

$$+ \sum_{i<j}^{N} V^{2B}(i,j)$$

$$+ \sum_{i<j<k}^{N} V^{3B}(i,j,k)$$

$$+ \ldots$$

$$+ V^{NB}(1,\ldots,N)$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

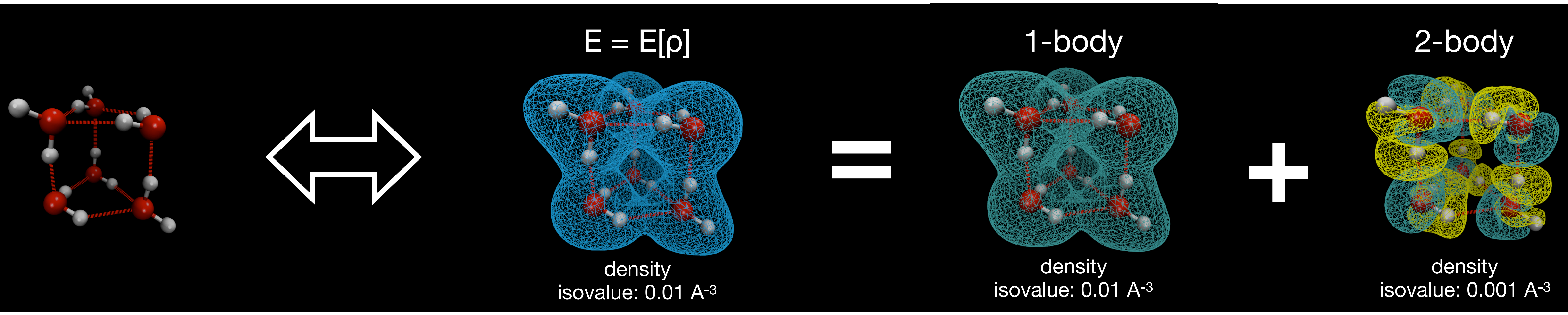
FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$
\begin{aligned}
E_N &= \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\
&+ \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\
&+ \sum_{i<j<k}^{N} V^{3B}(i,j,k) \\
&+ \ldots \\
&+ V^{NB}(1,\ldots,N)
\end{aligned}
$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$\begin{aligned} E_N = {} & \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\ & + \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\ & + \sum_{i<j<k}^{N} V^{3B}(i,j,k) && \text{3-body} \\ & + \ldots \\ & + V^{NB}(1,\ldots,N) \end{aligned}$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$
\begin{aligned}
E_N = \sum_{i}^{N} V^{1B}(i) & \qquad \text{1-body} \\
+ \sum_{i<j}^{N} V^{2B}(i,j) & \qquad \text{2-body} \\
+ \sum_{i<j<k}^{N} V^{3B}(i,j,k) & \qquad \text{3-body} \\
+ \ldots & \qquad \text{4-body} \\
& \qquad \ldots \\
+ V^{NB}(1,\ldots,N) & \qquad \text{N-body}
\end{aligned}
$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Many-Body Perspective

$$
\begin{aligned}
E_N = & \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\
& + \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\
& + \sum_{i<j<k}^{N} V^{3B}(i,j,k) && \text{3-body} \\
& + \ldots && \text{4-body} \\
& && \ldots \\
& + V^{NB}(1,\ldots,N) && \text{N-body}
\end{aligned}
$$

MB expansion converges quickly for non-metallic systems

see: Clementi, Stillinger, Xantheas, Gordon, Truhlar, Jordan, Szalewicz, Bowman, ...

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

On the Nature of Many-Body Molecular Interactions

FP, Acc. Chem. Res. *49, 1844 (2016)*

On the Nature of Many-Body Molecular Interactions

FP, Acc. Chem. Res. 49, 1844 (2016)

On the Nature of Many-Body Molecular Interactions

FP, Acc. Chem. Res. 49, 1844 (2016)

On the Nature of Many-Body Molecular Interactions

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Data-Driven Many-Body Perspective

$$\begin{aligned}
E_N = & \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\
& + \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\
& + \sum_{i<j<k}^{N} V^{3B}(i,j,k) && \text{3-body} \\
& + \ldots && \text{4-body} \\
& && \ldots \\
& + V^{NB}(1,\ldots,N) && \text{N-body}
\end{aligned}$$

MB expansion converges quickly for non-metallic systems

see: Clementi, Stillinger, Xantheas, Gordon, Truhlar, Jordan, Szalewicz, Bowman, ...

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Data-Driven Many-Body Perspective

$$\begin{aligned} E_N = & \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\ & + \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\ & + \sum_{i<j<k}^{N} V^{3B}(i,j,k) && \text{3-body} \\ & + \ldots && \text{4-body} \\ & && \ldots \\ & + V^{NB}(1,\ldots,N) && \text{N-body} \end{aligned}$$

quantum mechanics

MB expansion converges quickly for non-metallic systems

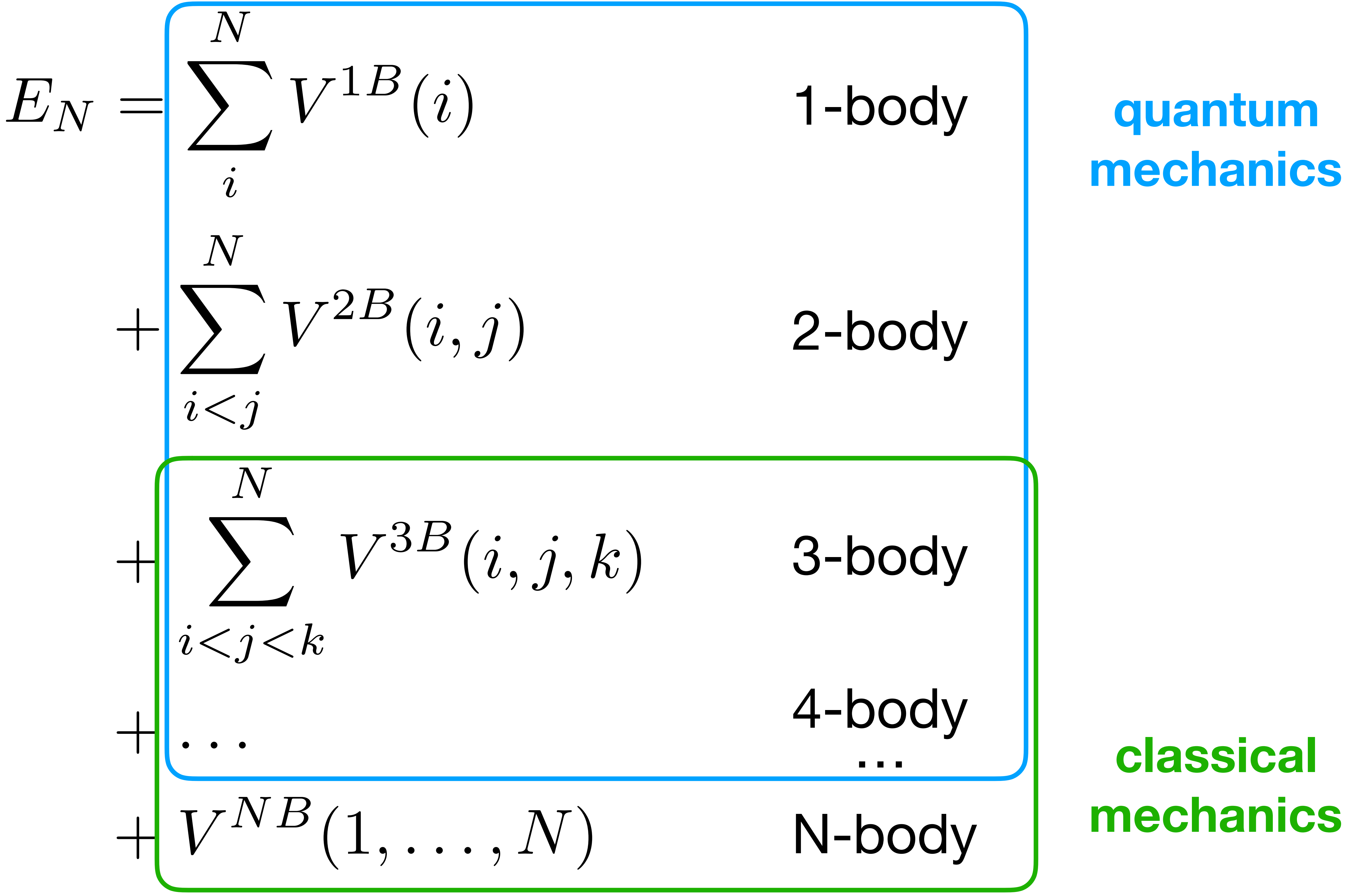
see: Clementi, Stillinger, Xantheas, Gordon, Truhlar, Jordan, Szalewicz, Bowman, ...

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Data-Driven Many-Body Perspective



$$E_N = \sum_{i}^{N} V^{1B}(i) + \sum_{i<j}^{N} V^{2B}(i,j) + \sum_{i<j<k}^{N} V^{3B}(i,j,k) + \ldots + V^{NB}(1,\ldots,N)$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. *<u>53</u>, 4544 (1970)*

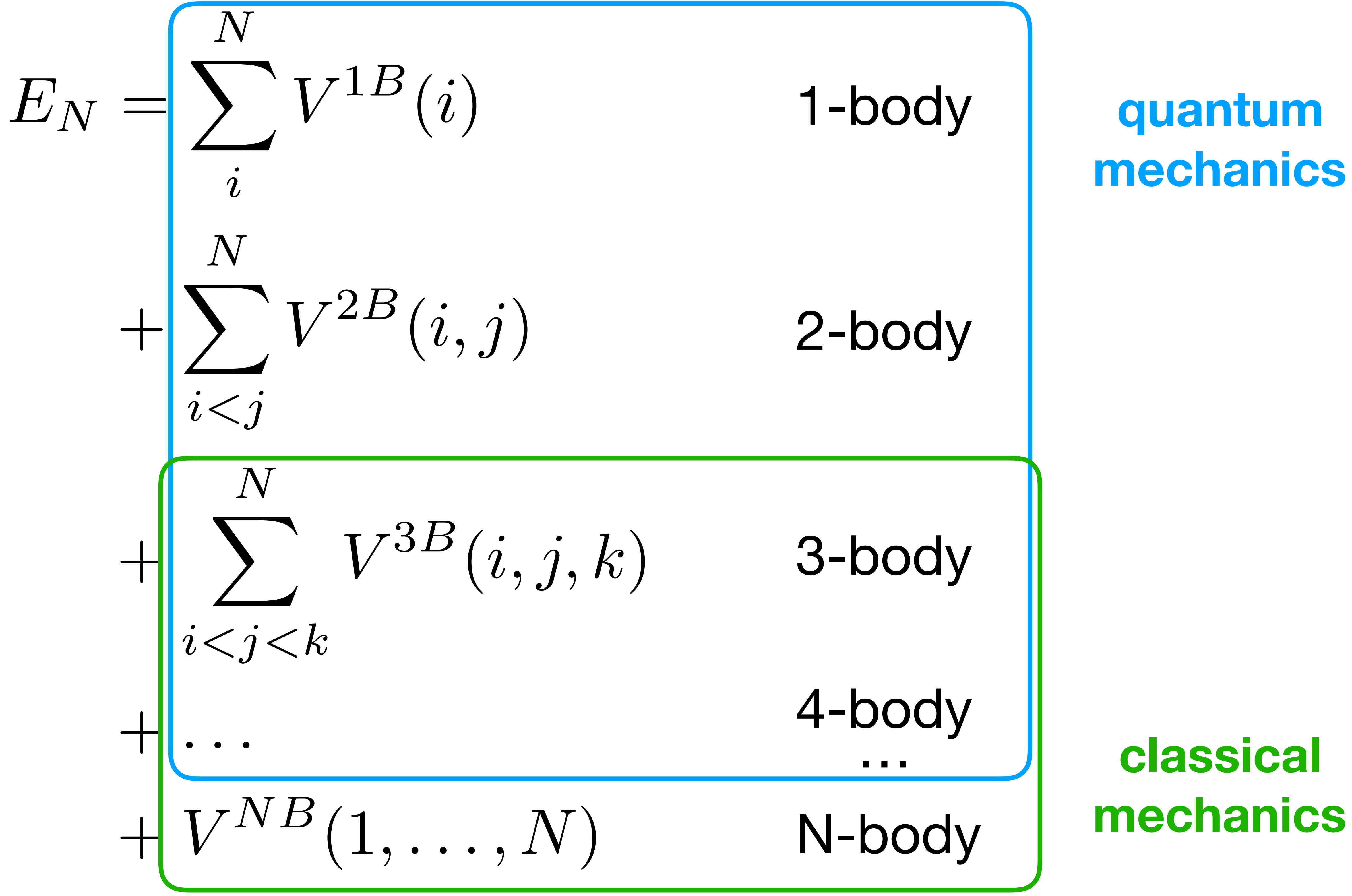
Popkie, Kistenmacher & Clementi, J. Chem. Phys. *<u>59</u>, 1325 (1973)*

MB expansion converges quickly for non-metallic systems

see: Clementi, Stillinger, Xantheas, Gordon, Truhlar, Jordan, Szalewicz, Bowman, ...

FP, Acc. Chem. Res. <u>49</u>, 1844 (2016)

Molecular Interactions: A Data-Driven Many-Body Perspective



$$E_N = \sum_{i}^{N} V^{1B}(i) \quad \text{1-body}$$

$$+ \sum_{i<j}^{N} V^{2B}(i,j) \quad \text{2-body}$$

$$+ \sum_{i<j<k}^{N} V^{3B}(i,j,k) \quad \text{3-body}$$

$$+ \ldots \quad \text{4-body} \quad \ldots$$

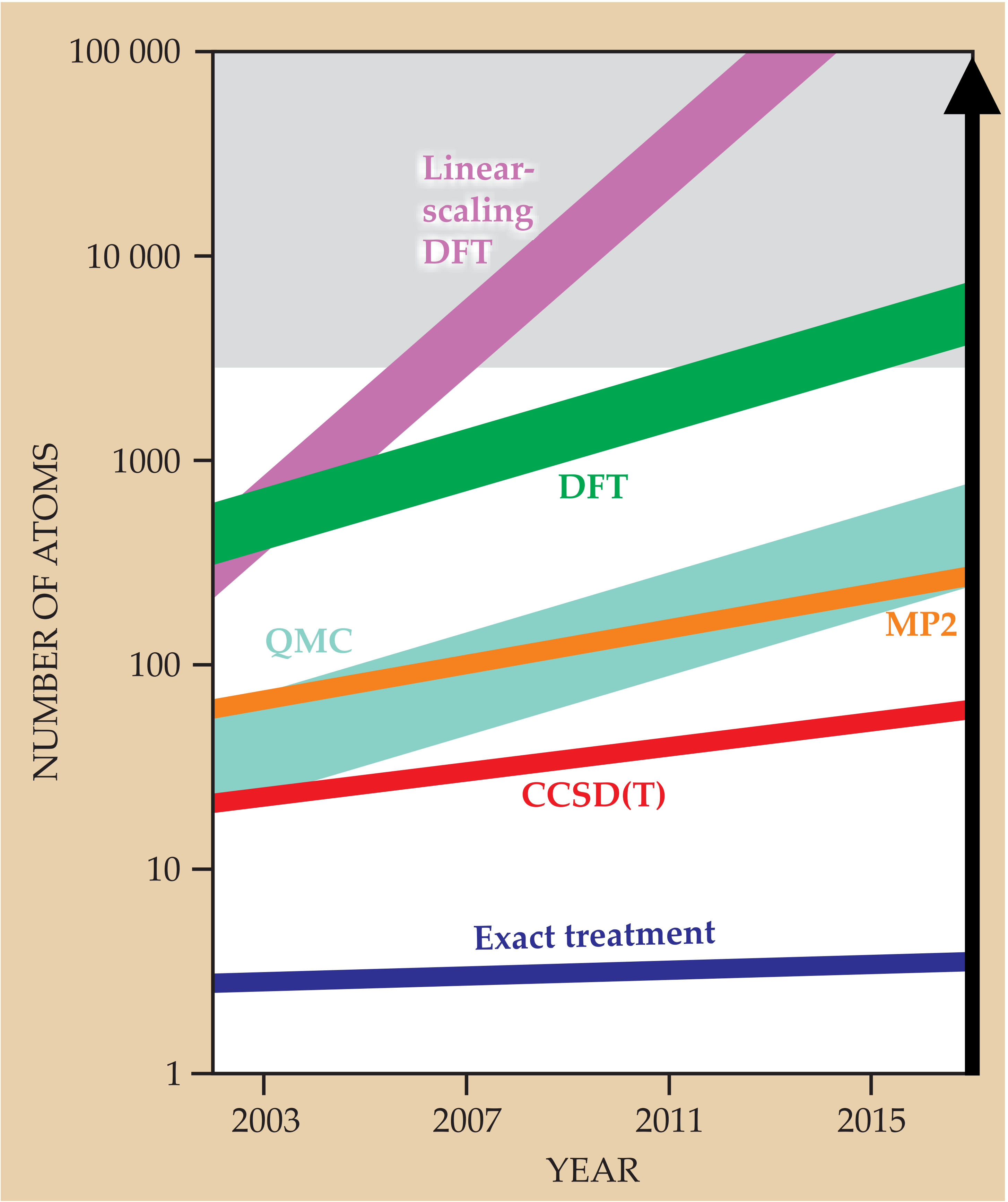
$$+ V^{NB}(1,\ldots,N) \quad \text{N-body}$$

quantum mechanics

classical mechanics

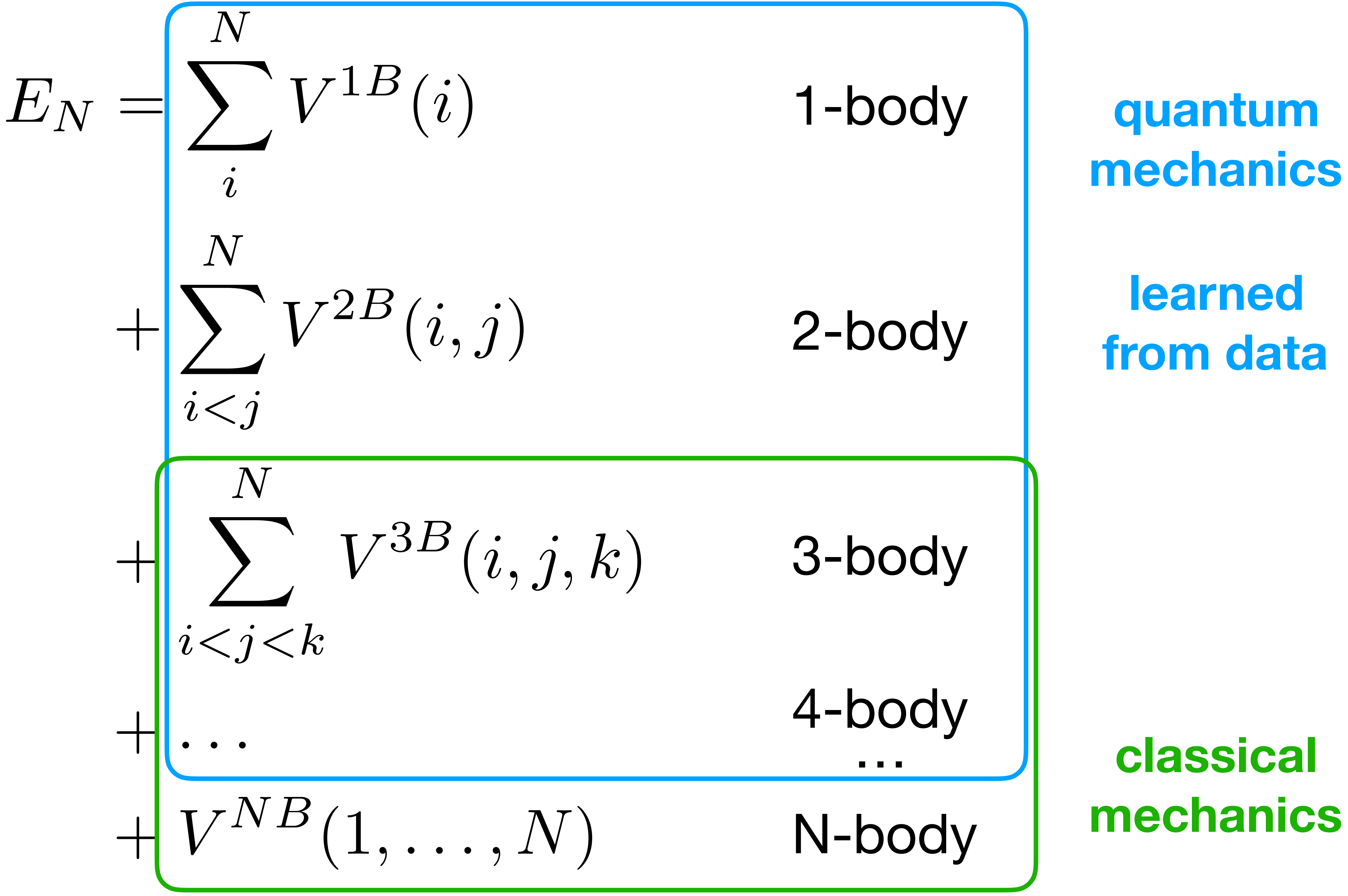
Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Data-Driven Many-Body Perspective



$$\begin{aligned} E_N = & \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\ & + \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\ & + \sum_{i<j<k}^{N} V^{3B}(i,j,k) && \text{3-body} \\ & + \dots && \text{4-body} \\ & && \dots \\ & + V^{NB}(1,\dots,N) && \text{N-body} \end{aligned}$$

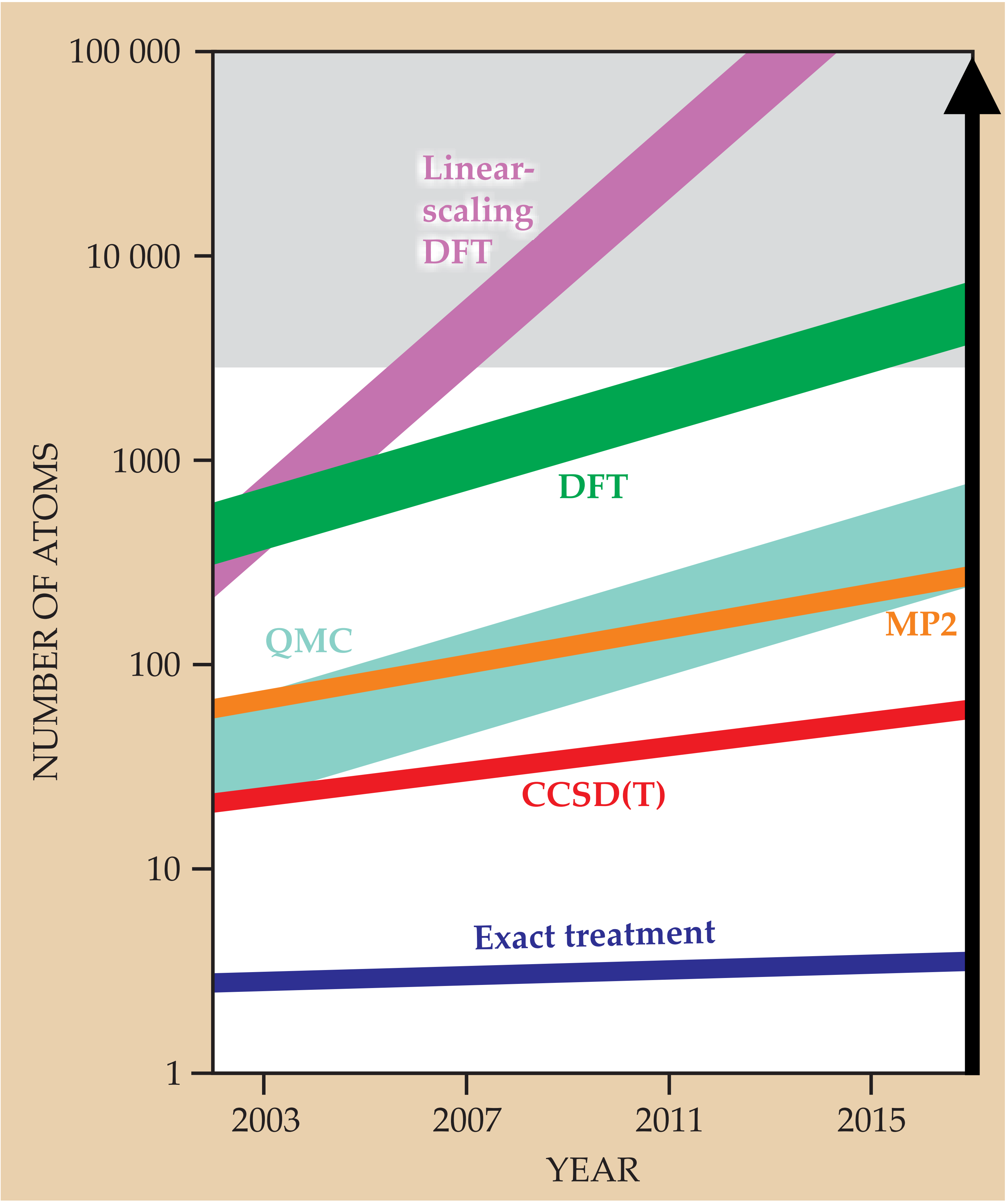
quantum mechanics

learned from data

classical mechanics

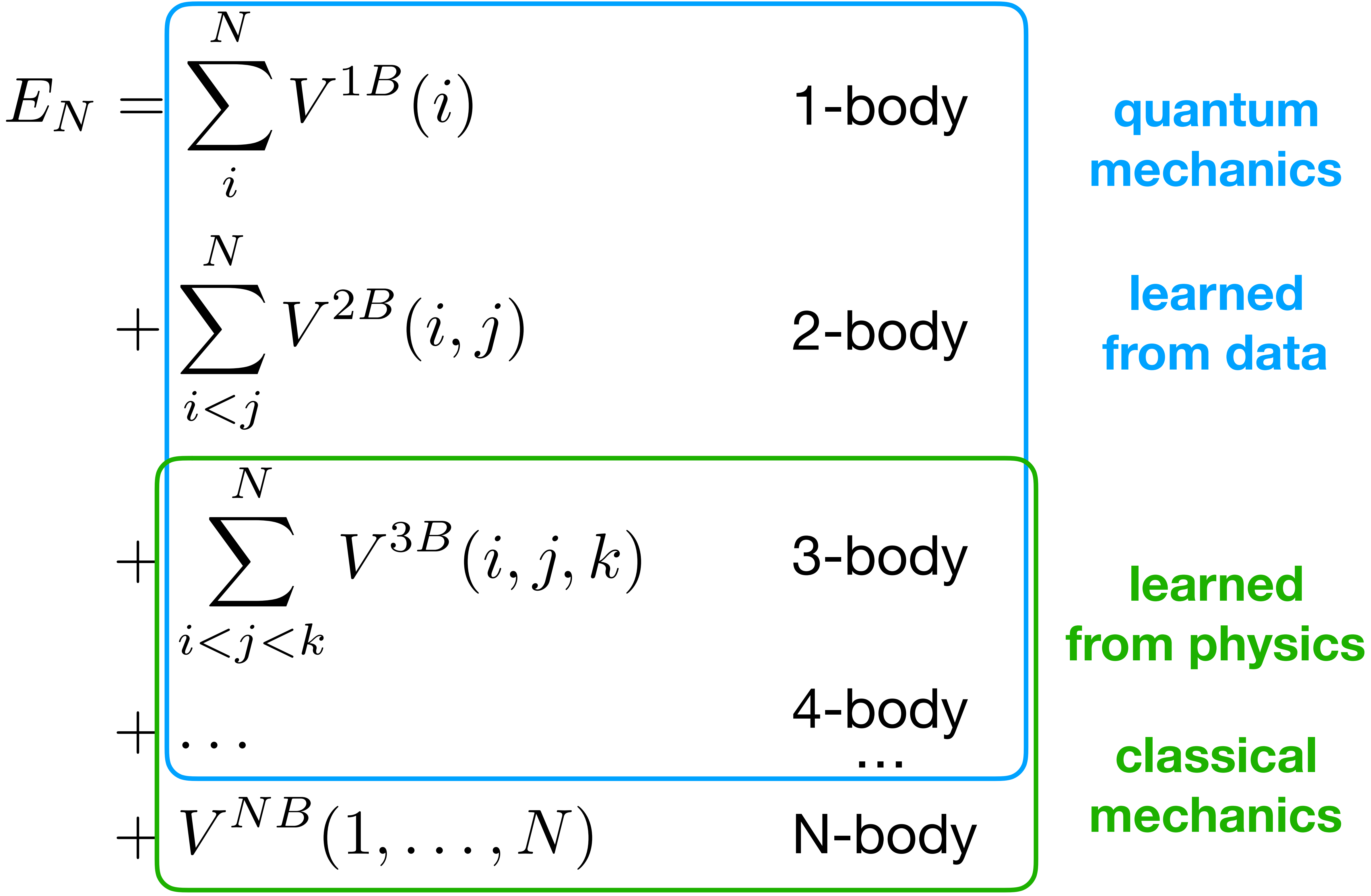
Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Data-Driven Many-Body Perspective



$$E_N = \sum_{i}^{N} V^{1B}(i) \quad \text{1-body}$$

$$+ \sum_{i<j}^{N} V^{2B}(i,j) \quad \text{2-body}$$

$$+ \sum_{i<j<k}^{N} V^{3B}(i,j,k) \quad \text{3-body}$$

$$+ \dots \quad \text{4-body} \; \dots$$

$$+ V^{NB}(1,\dots,N) \quad \text{N-body}$$

quantum mechanics

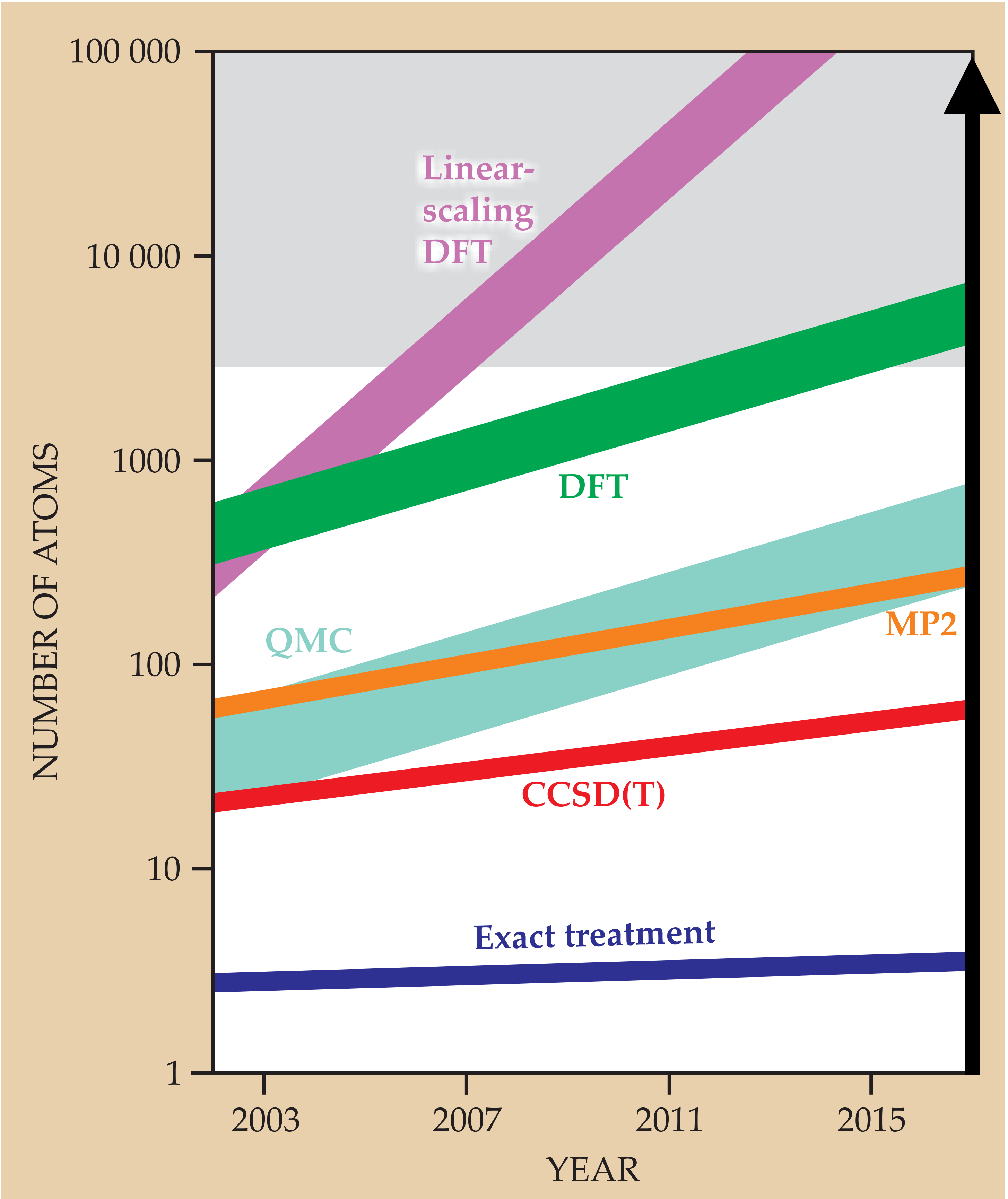
learned from data

learned from physics

classical mechanics

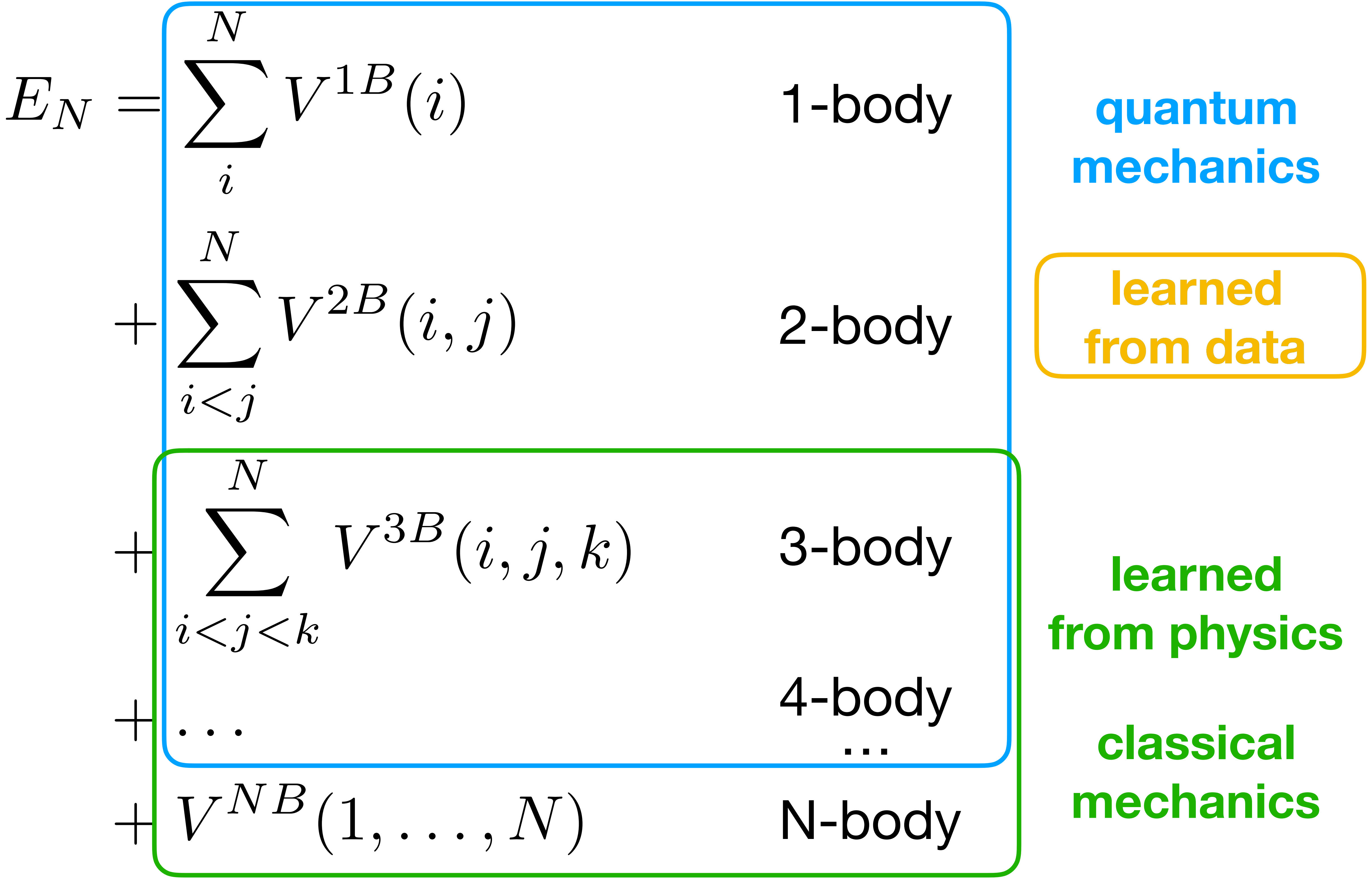
Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)



FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Data-Driven Many-Body Perspective



$$\begin{aligned} E_N = & \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\ & + \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\ & + \sum_{i<j<k}^{N} V^{3B}(i,j,k) && \text{3-body} \\ & + \ldots && \text{4-body} \\ & && \ldots \\ & + V^{NB}(1,\ldots,N) && \text{N-body} \end{aligned}$$

quantum mechanics

learned from data

learned from physics

classical mechanics

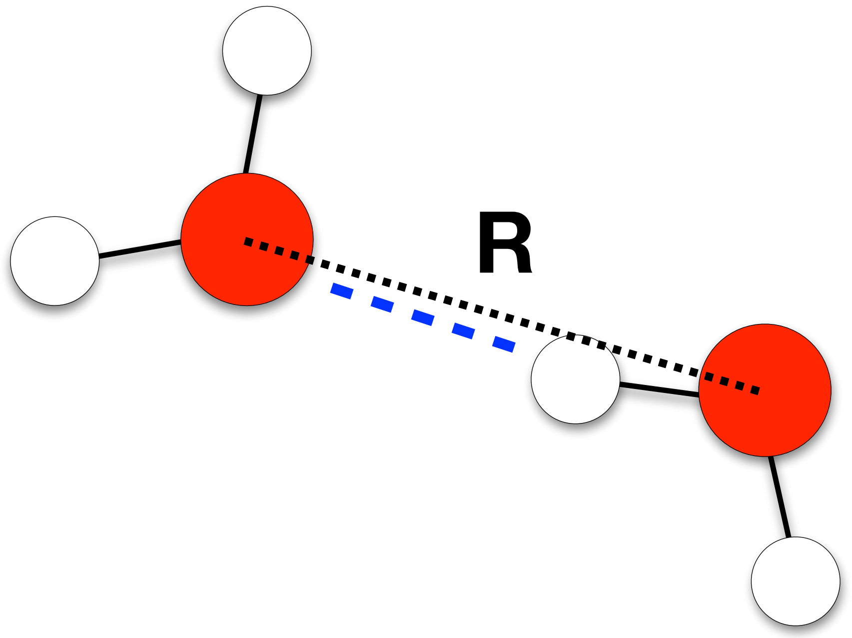
$$\boxed{\Delta^{nB} = V^{nB} - V^{nB}_{CM}}$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

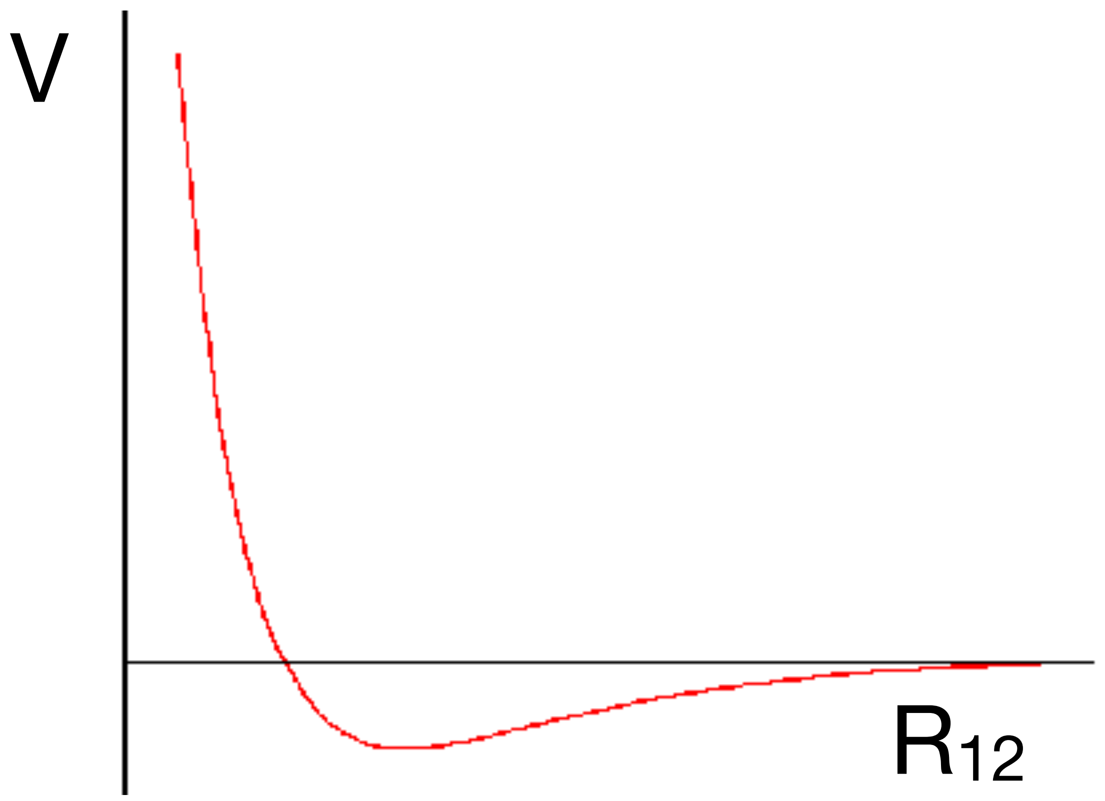
Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

FP, Acc. Chem. Res. 49, 1844 (2016)

Molecular Interactions: A Data-Driven Many-Body Perspective

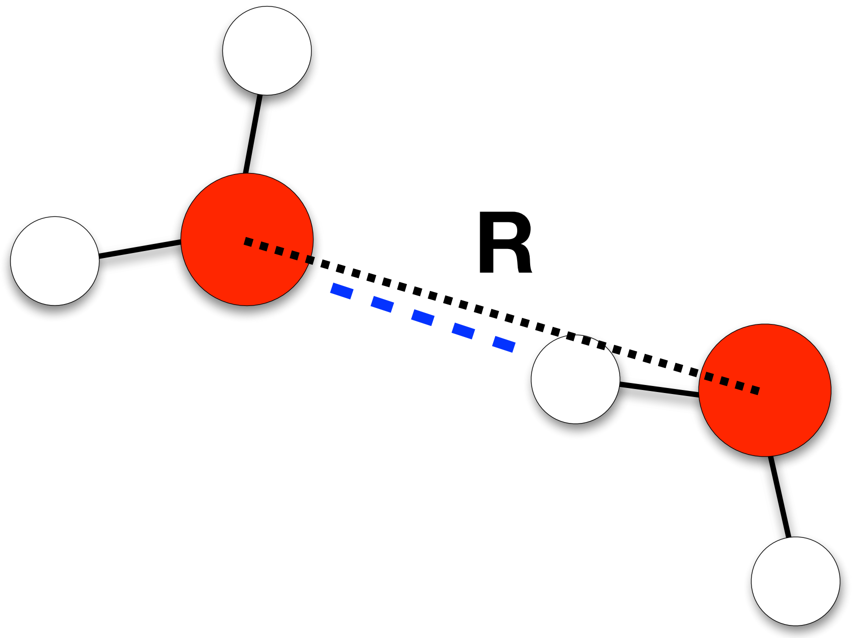


$$V^{2B} = s_2 V^{2B}_{sr} \quad + \quad V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$



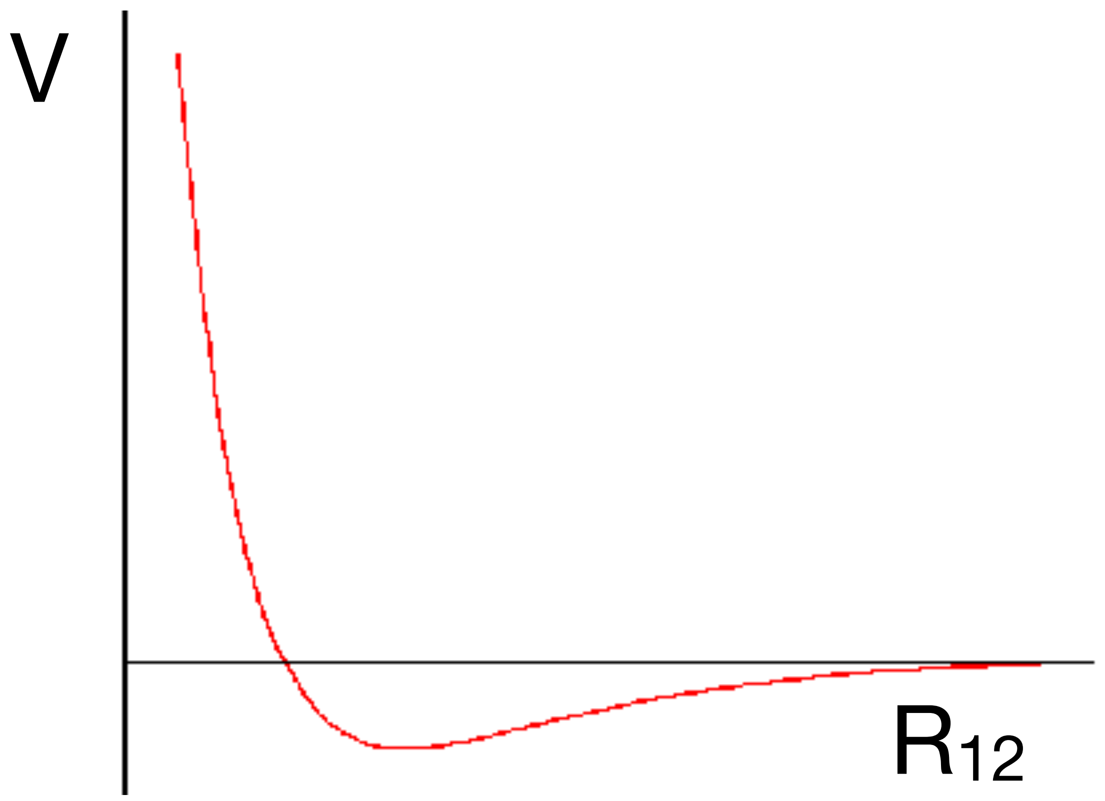
Babin, Leforestier & FP, J. Chem. Theory Comput. 9, *5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput.* 10, *1599 (2014)*

Molecular Interactions: A Data-Driven Many-Body Perspective



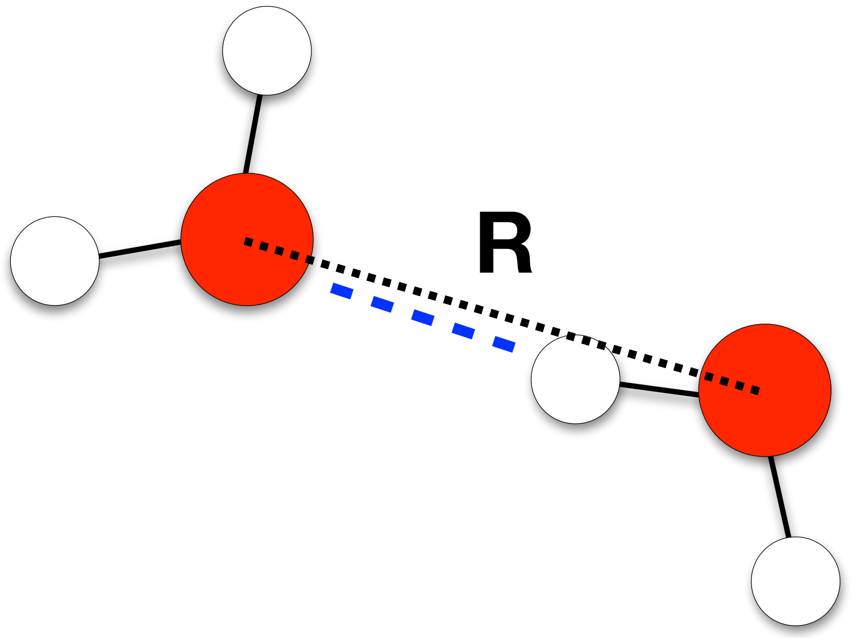
permanent electrostatics

$$V^{2B} = s_2 V^{2B}_{sr} + \boxed{V^{2B}_{elec}} + V^{2B}_{ind} + V^{2B}_{disp}$$



Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

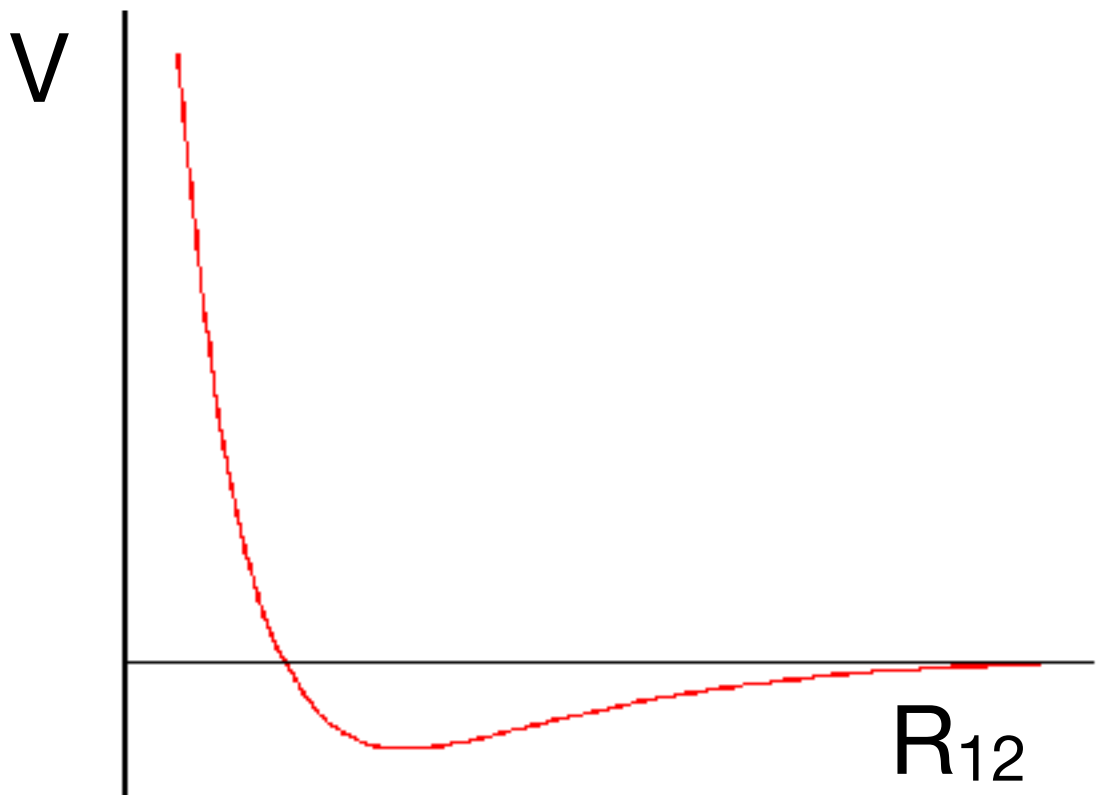
Molecular Interactions: A Data-Driven Many-Body Perspective



permanent electrostatics

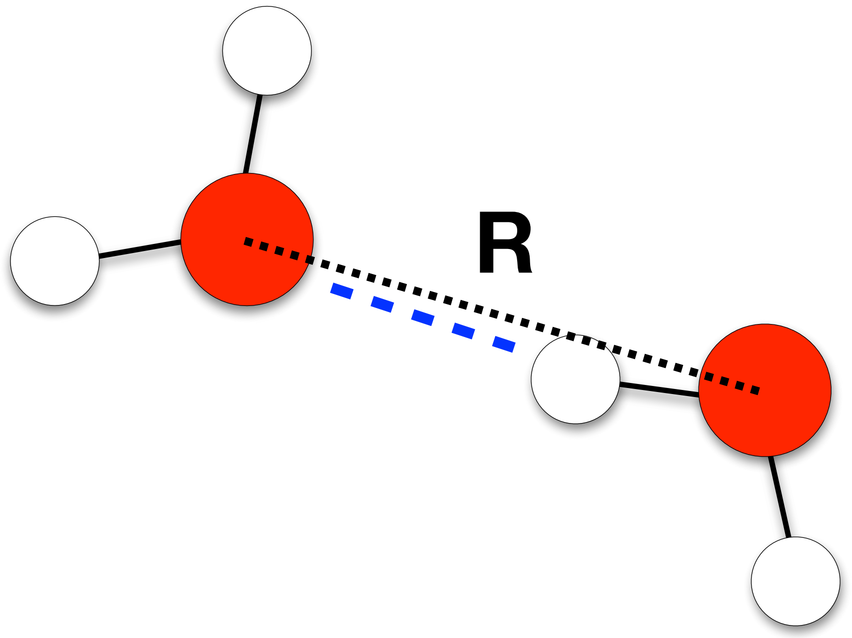
$$V^{2B} = s_2 V_{sr}^{2B} + \boxed{V_{elec}^{2B}} + \boxed{V_{ind}^{2B}} + V_{disp}^{2B}$$

2B induction



Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

Molecular Interactions: A Data-Driven Many-Body Perspective

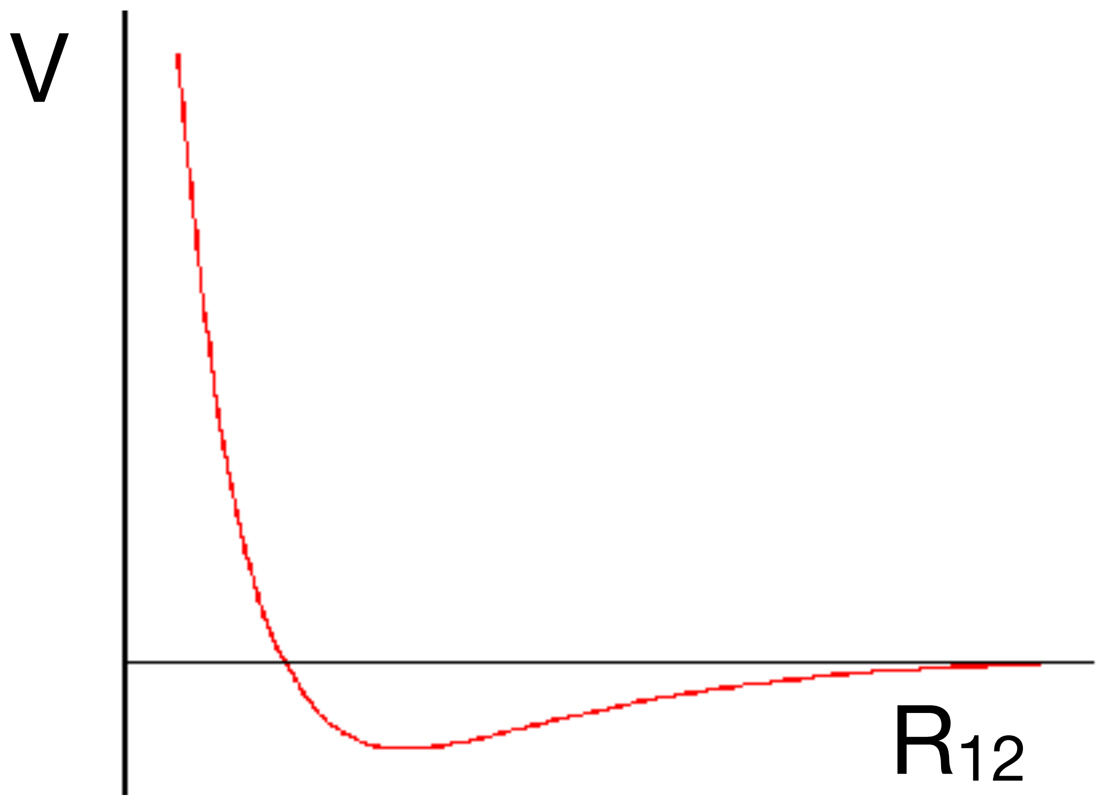


permanent electrostatics

2B dispersion

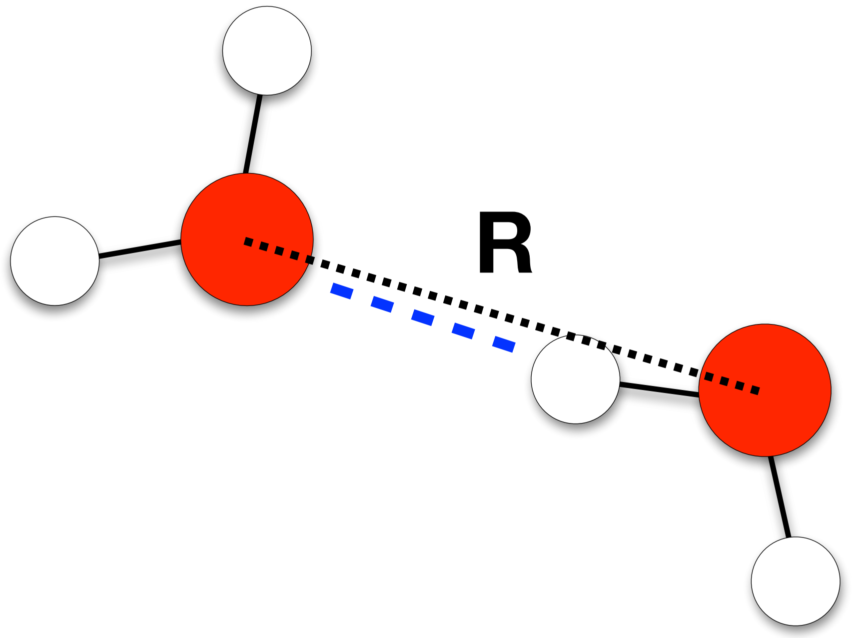
$$V^{2B} = s_2 V^{2B}_{sr} + \boxed{V^{2B}_{elec}} + \boxed{V^{2B}_{ind}} + \boxed{V^{2B}_{disp}}$$

2B induction



Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

Molecular Interactions: A Data-Driven Many-Body Perspective



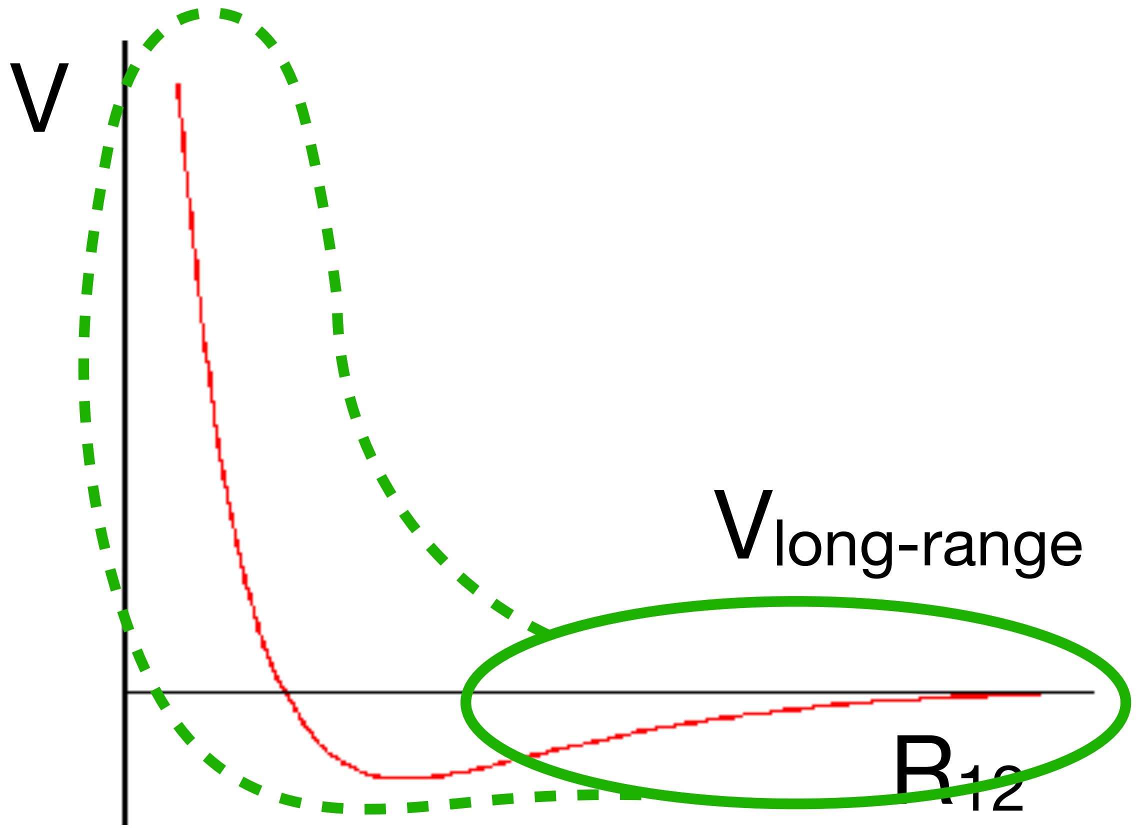
permanent electrostatics

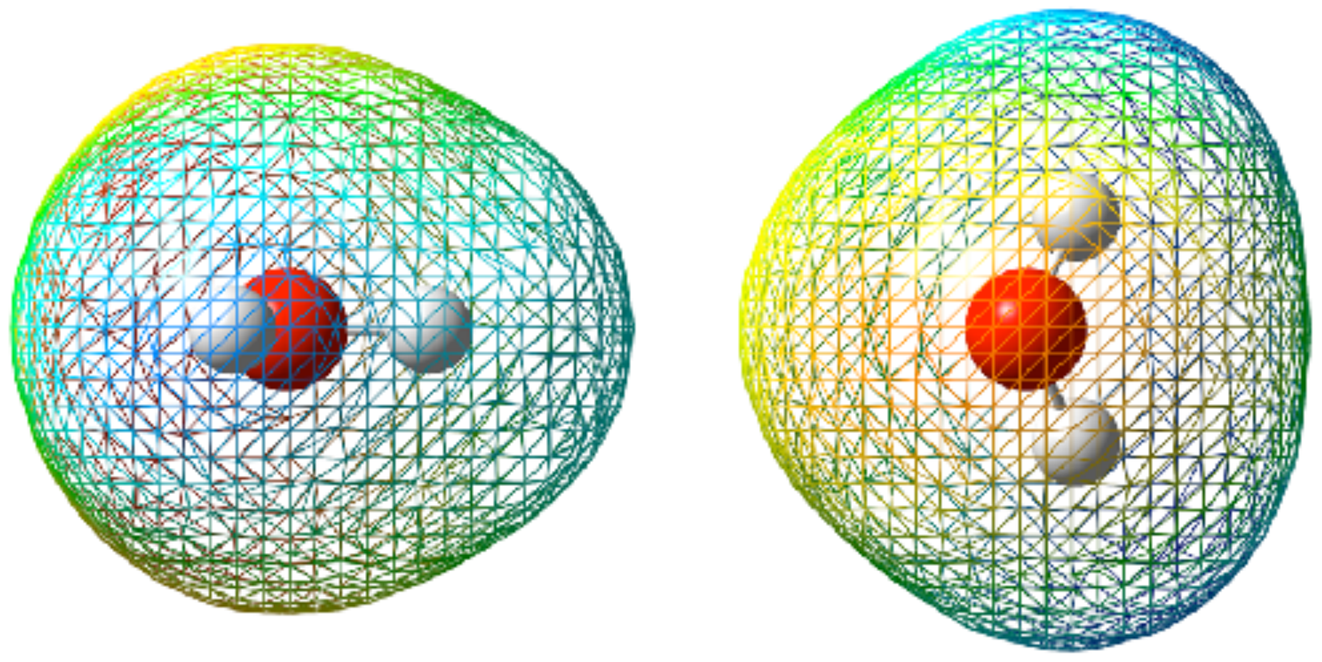
2B dispersion

$$V^{2B} = s_2 V^{2B}_{sr} + V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$

2B induction

from physics of isolated monomers



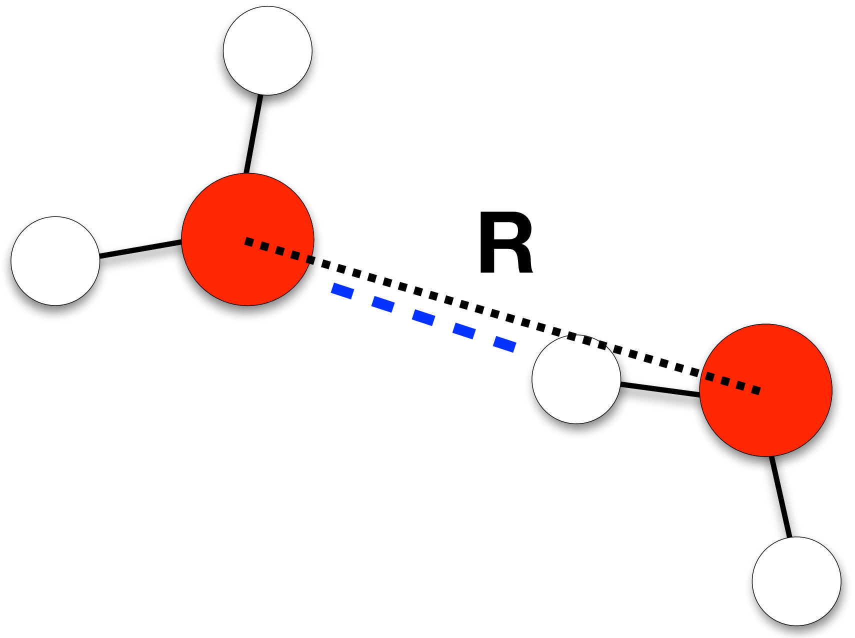


classical representations

Stone, “The Theory of Intermolecular Forces”

Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

Molecular Interactions: A Data-Driven Many-Body Perspective



$$V^{2B} = s_2 V^{2B}_{sr} + V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$

short-range interactions

permanent electrostatics

2B induction

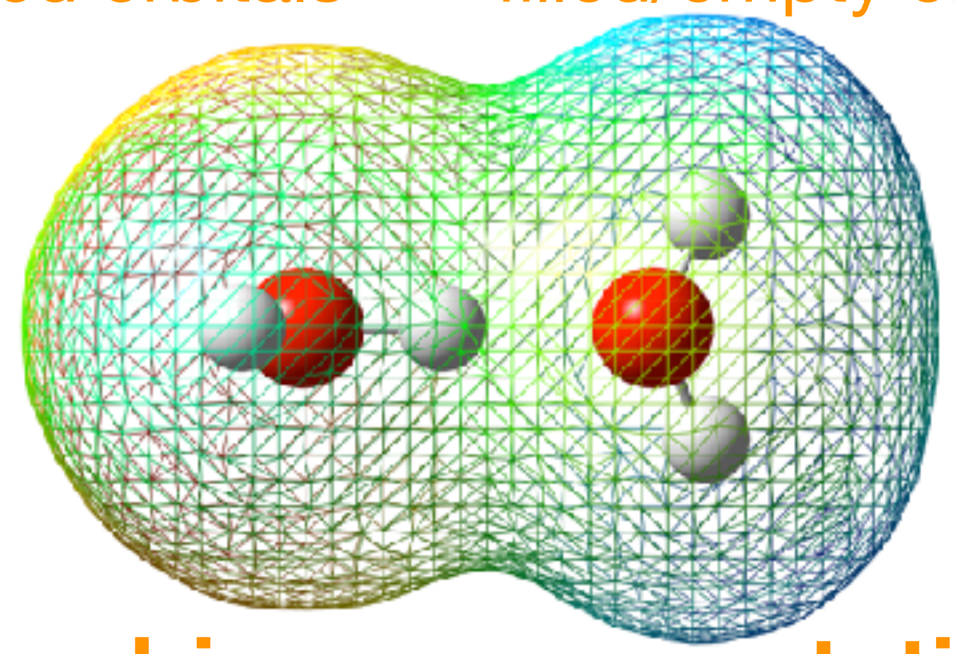
2B dispersion

from physics of isolated monomers

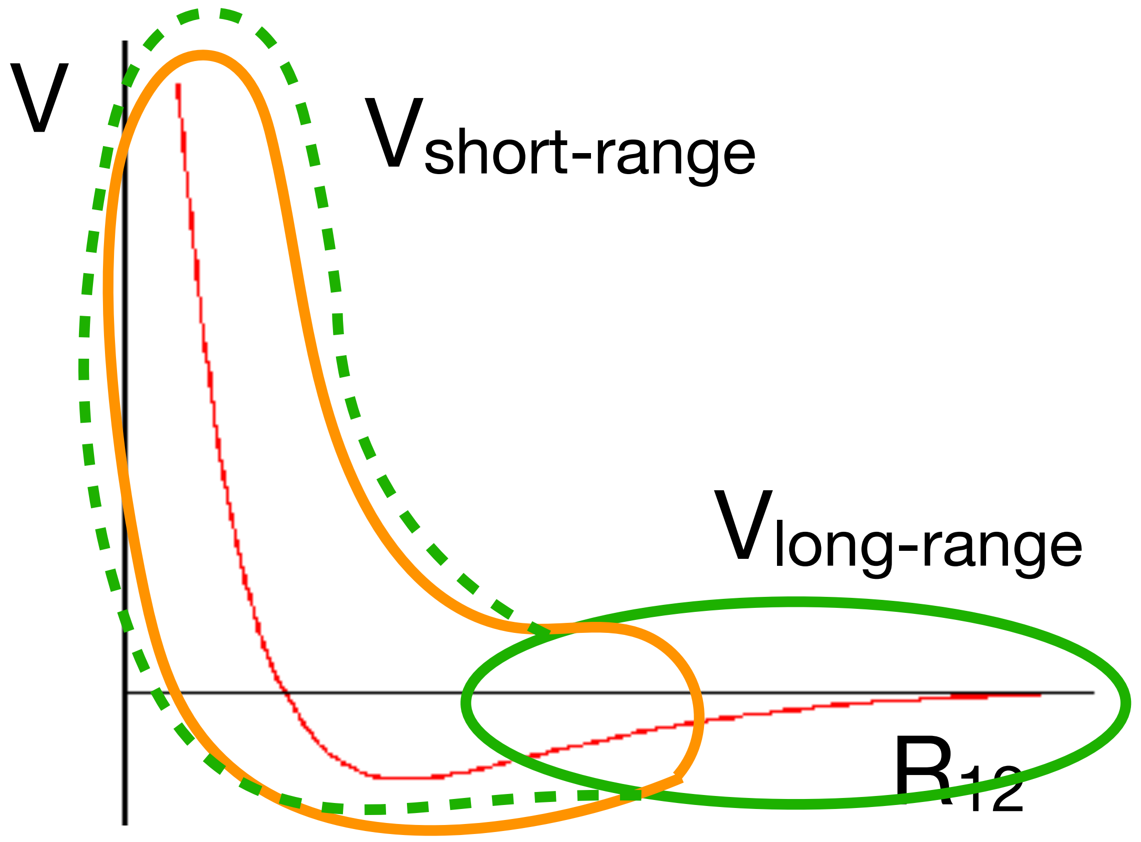
Pauli repulsion & charge transfer

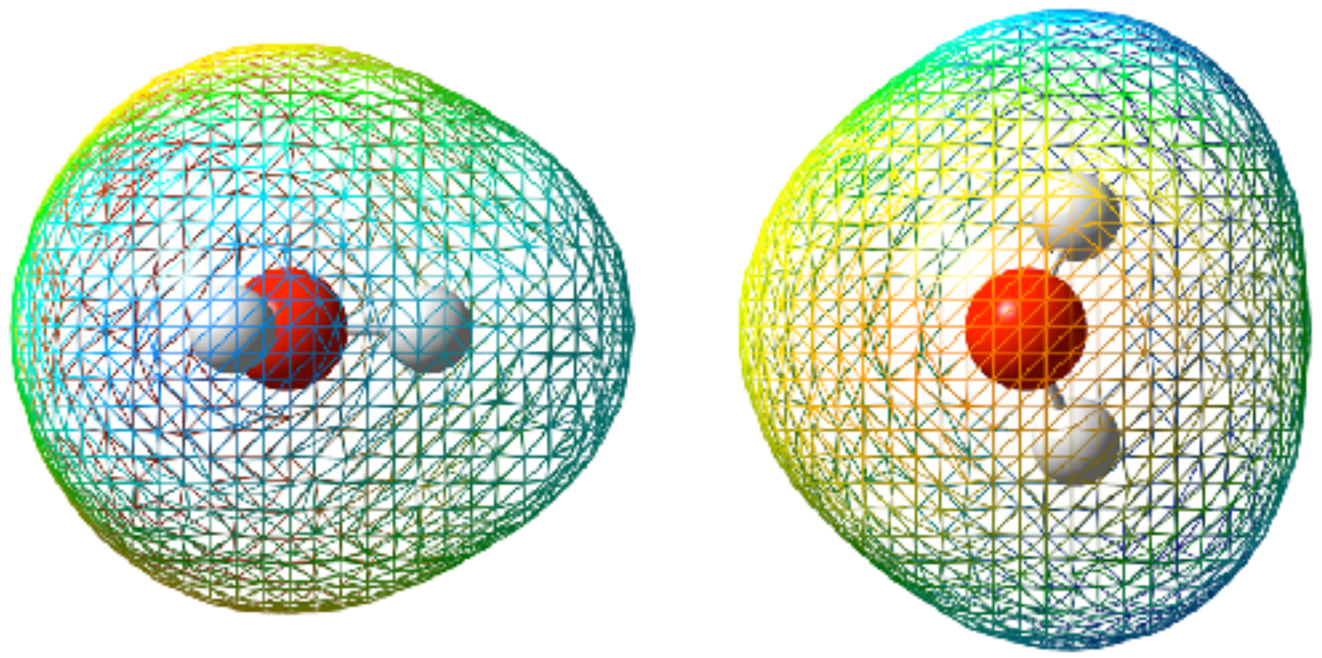
filled/filled orbitals

filled/empty orbitals



data-driven representations derived from CCSD(T)



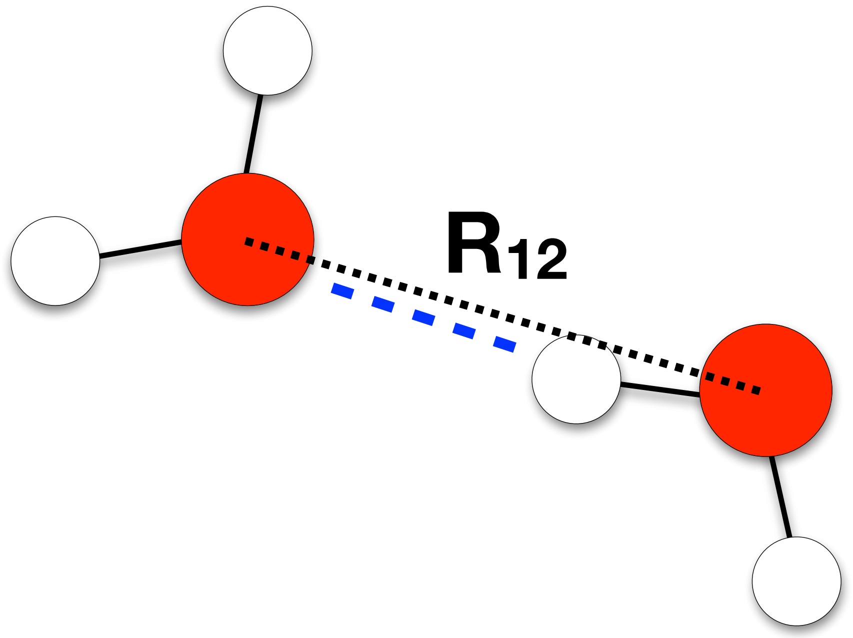


classical representations

Stone, "The Theory of Intermolecular Forces"

Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

Representability of Low-Order Many-Body Interactions



permanent electrostatics

2B dispersion

$$V^{2B} = s_2 V^{2B}_{sr} + V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$

short-range interactions

2B induction

from physics of isolated monomers

Permutationally invariant polynomials

Braams & Bowman, Int. Rev. Phys. Chem. 28, 577 (2009)

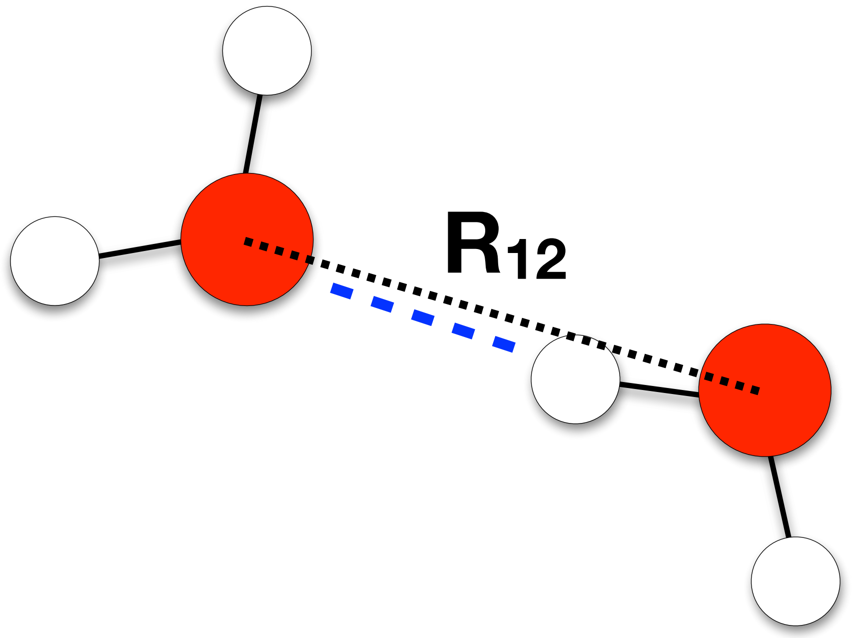
$$V^{2B}_{sr} = V^{2B}_{PIP} = \sum_l A_l \eta_l = A_1\xi_a + A_2\xi_b + A_3\xi_c + \ldots + A_i\xi_a\xi_b + A_j\xi_a\xi_c + A_k\xi_b\xi_c + \ldots + A_q\xi_a\xi_b\xi_c + \ldots + \ldots$$

Morse variables: $\xi_a = e^{-kr_a}$

$a = r_m - r_n$

Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

Representability of Low-Order Many-Body Interactions



$$V^{2B} = s_2 V^{2B}_{sr} + V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$

short-range interactions

permanent electrostatics

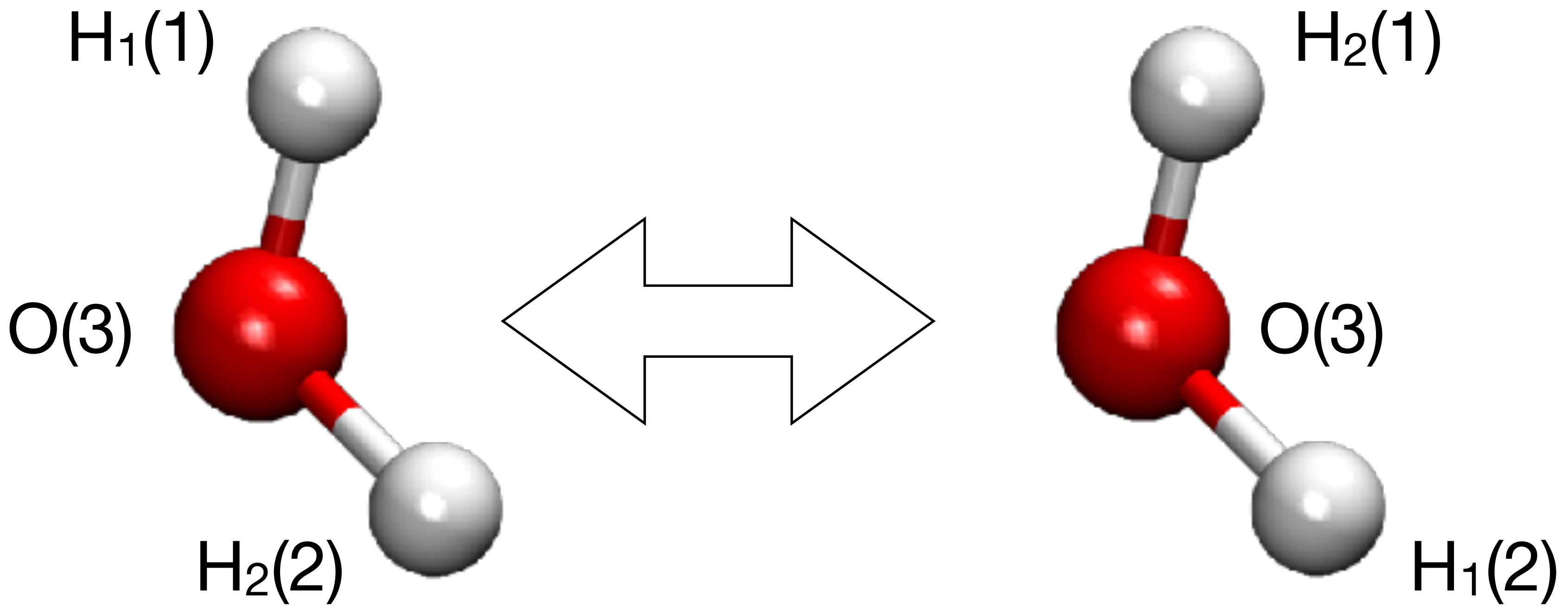
2B induction

2B dispersion

from physics of isolated monomers

Permutationally invariant polynomials

Braams & Bowman, Int. Rev. Phys. Chem. 28, 577 (2009)



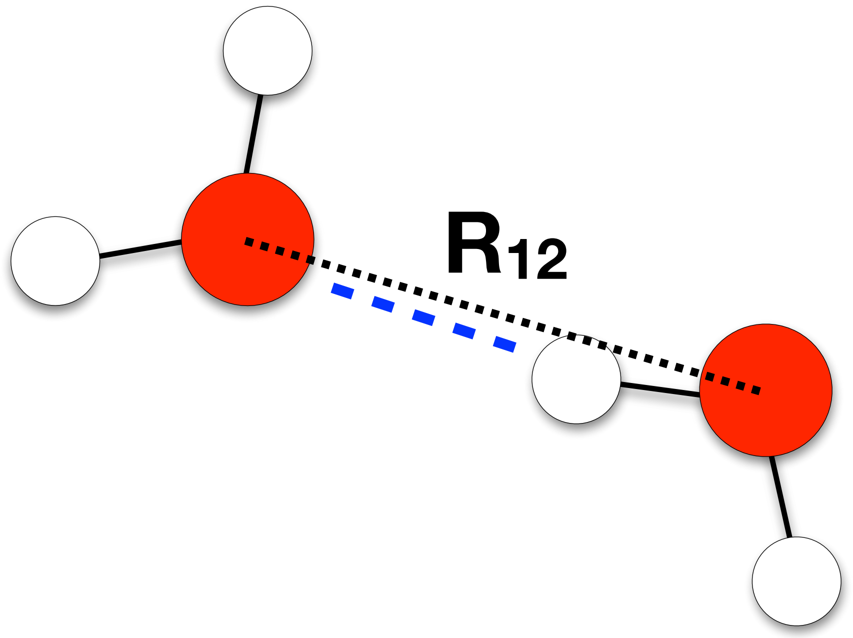
$$\eta_u = A_u \xi_{12}\xi_{13}\xi_{23} \qquad \eta_v = A_v \xi_{12}\xi_{23}\xi_{13}$$

$\Downarrow$ symmetrize

$$\eta_s = A_s \xi_{12}[\xi_{13}\xi_{23} + \xi_{23}\xi_{13}]$$

Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

Representability of Low-Order Many-Body Interactions



$$V^{2B} = s_2 V^{2B}_{sr} + V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$

short-range interactions

permanent electrostatics

2B induction

2B dispersion

from physics of isolated monomers

Permutationally invariant polynomials

Braams & Bowman, Int. Rev. Phys. Chem. 28, 577 (2009)

$$V^{2B}_{sr} = V^{2B}_{PIP} = \sum_l A_l \eta_l = A_1\xi_a + A_2\xi_b + A_3\xi_c + \ldots + A_i\xi_a\xi_b + A_j\xi_a\xi_c + A_k\xi_b\xi_c + \ldots + A_q\xi_a\xi_b\xi_c + \ldots + \ldots$$

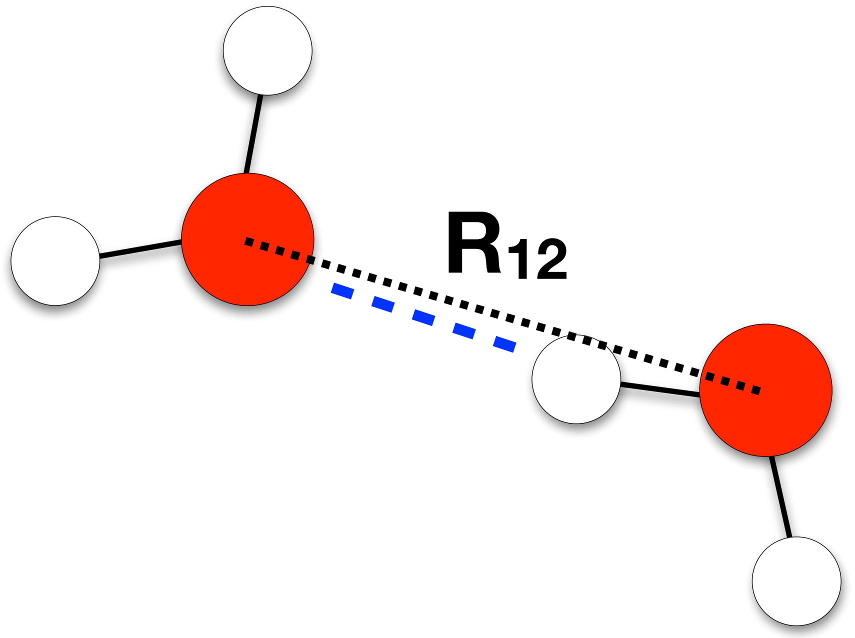
Morse variables: $\xi_a = e^{-kr_a}$

$a = r_m - r_n$

Ridge regression: $\chi^2 = \sum_{n \in S} [V^{2B}_{PIP} - V^{2B}_{sr,CCSD(T)}]^2 + \Gamma^2 \sum_l A_l^2$

Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

Representability of Low-Order Many-Body Interactions



permanent electrostatics

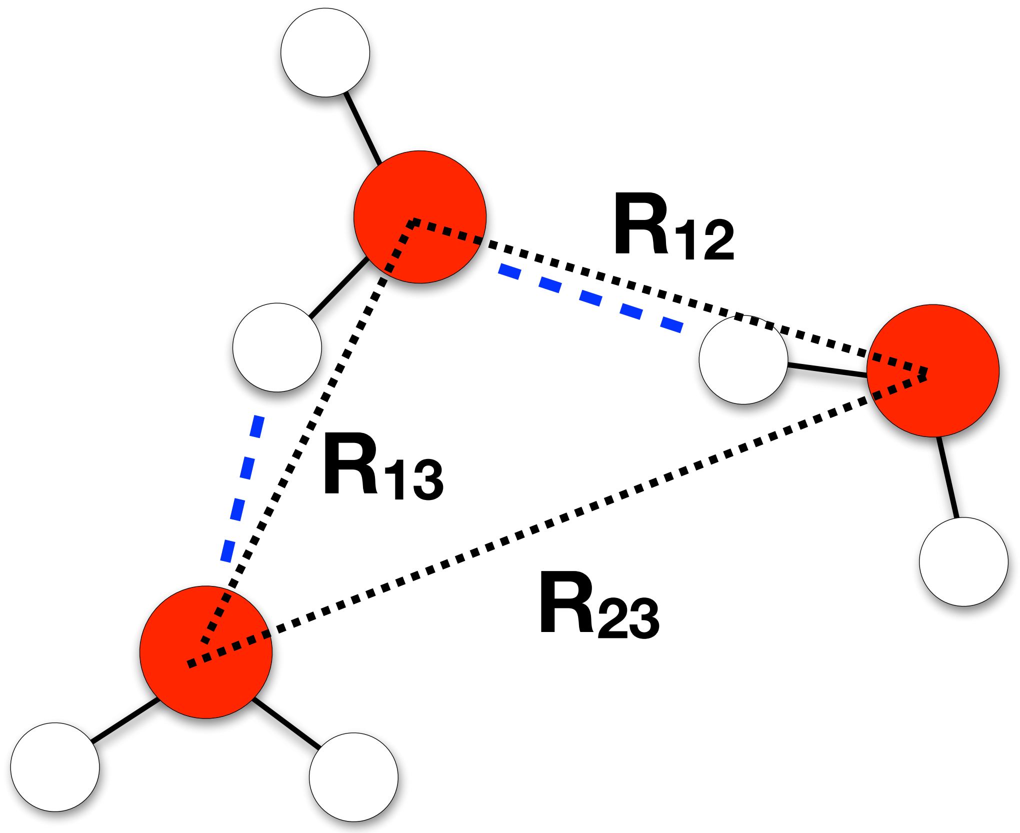
2B dispersion

$$V^{2B} = s_2 V^{2B}_{sr} + V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$

short-range interactions

2B induction

from physics of isolated monomers



3B dispersion

$$V^{3B} = s_3 V^{3B}_{sr} + V^{3B}_{ind} + V^{3B}_{disp}$$

short-range interactions

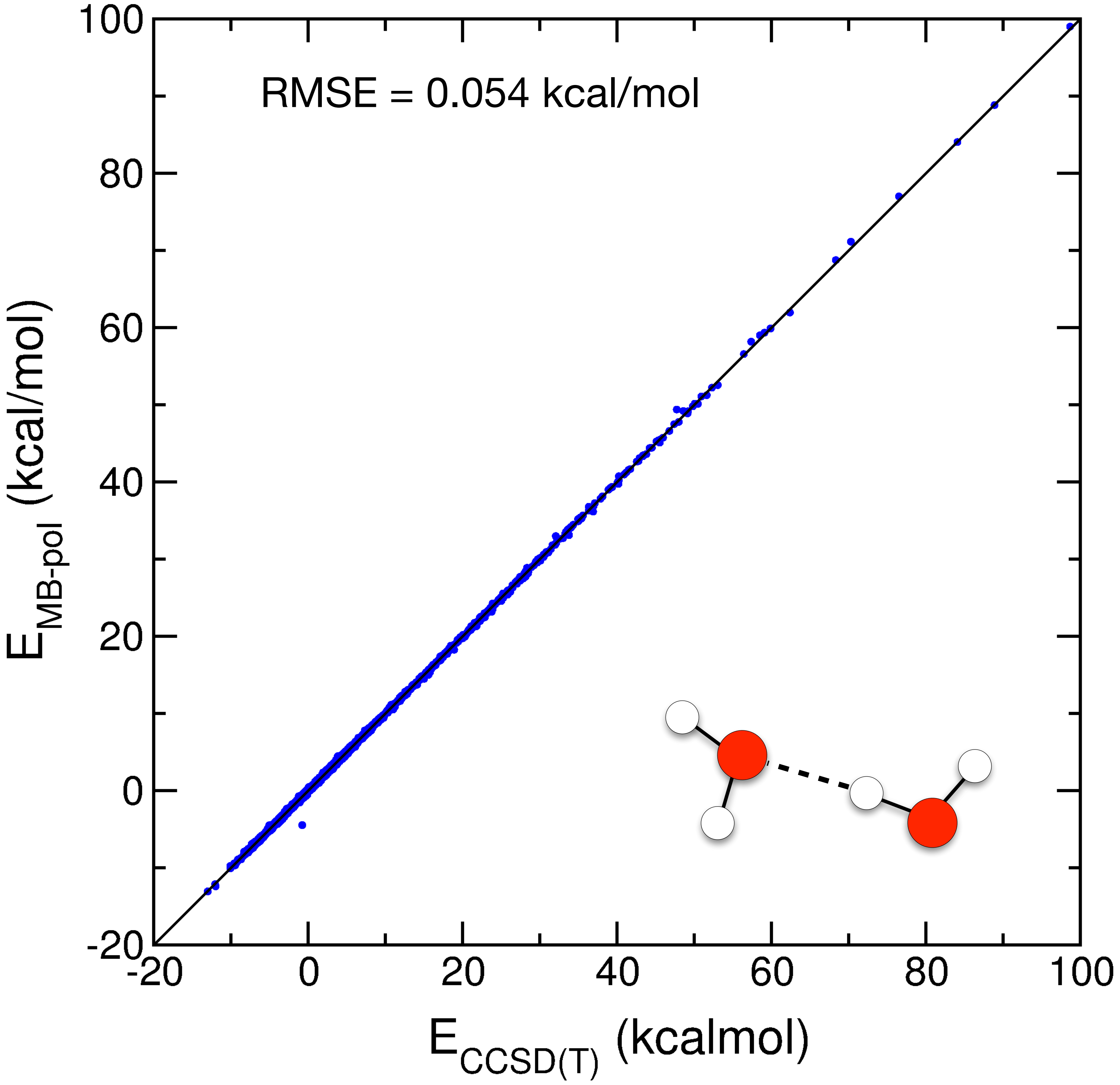
2B induction

Babin, Leforestier & FP, J. Chem. Theory Comput. 9, 5395 (2013); Babin, Medders & FP, J. Chem. Theory Comput. 10, 1599 (2014)

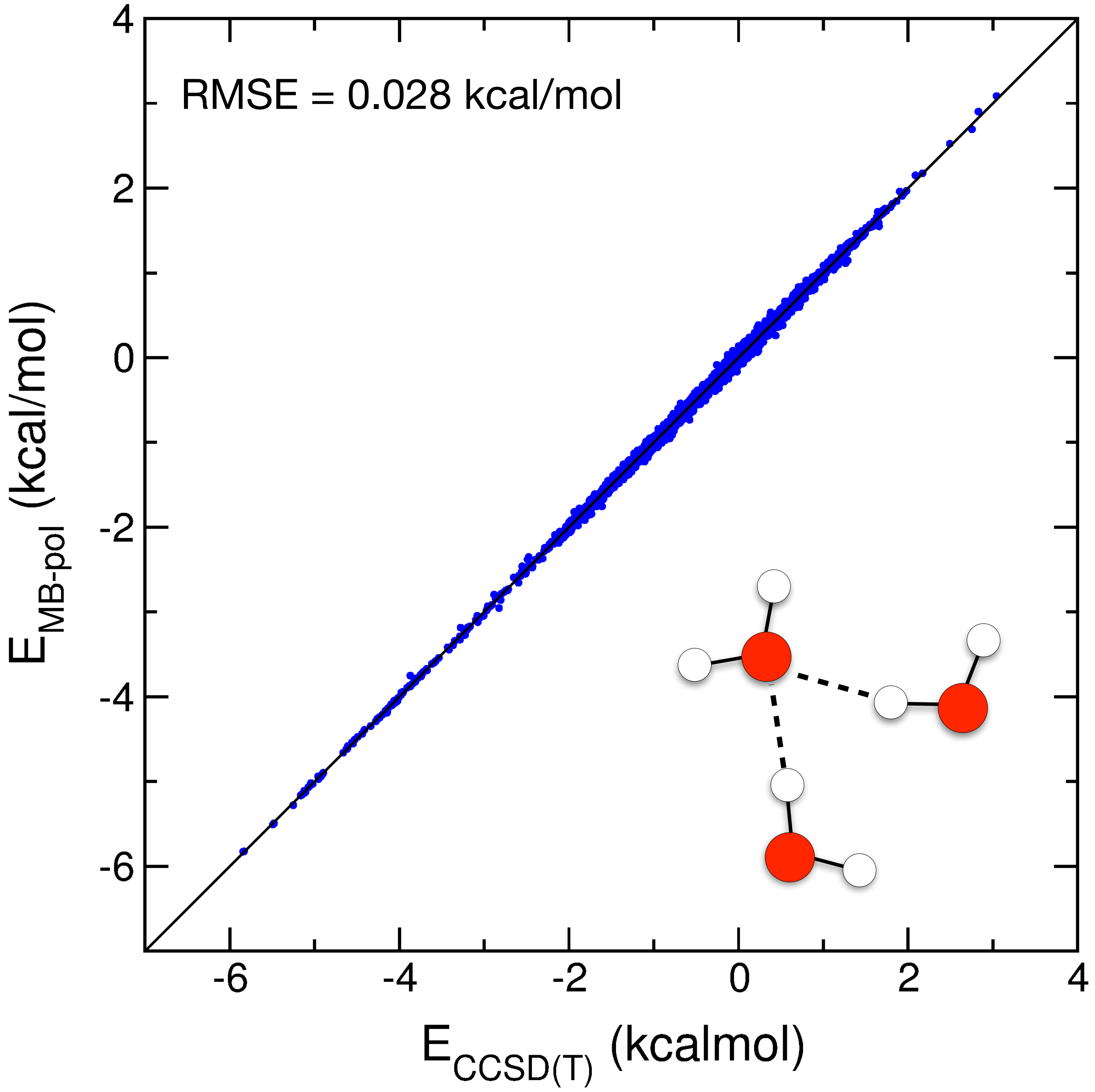
Many-Body Models: Do They Work?

$$E_N = \sum_{i}^{N} V^{1B}(i) + \boxed{\sum_{i<j}^{N} V^{2B}(i,j)} + \boxed{\sum_{i<j<k}^{N} V^{3B}(i,j,k)} + \cdots + V^{NB}(1,\ldots,N)$$

2B energies



3B energies

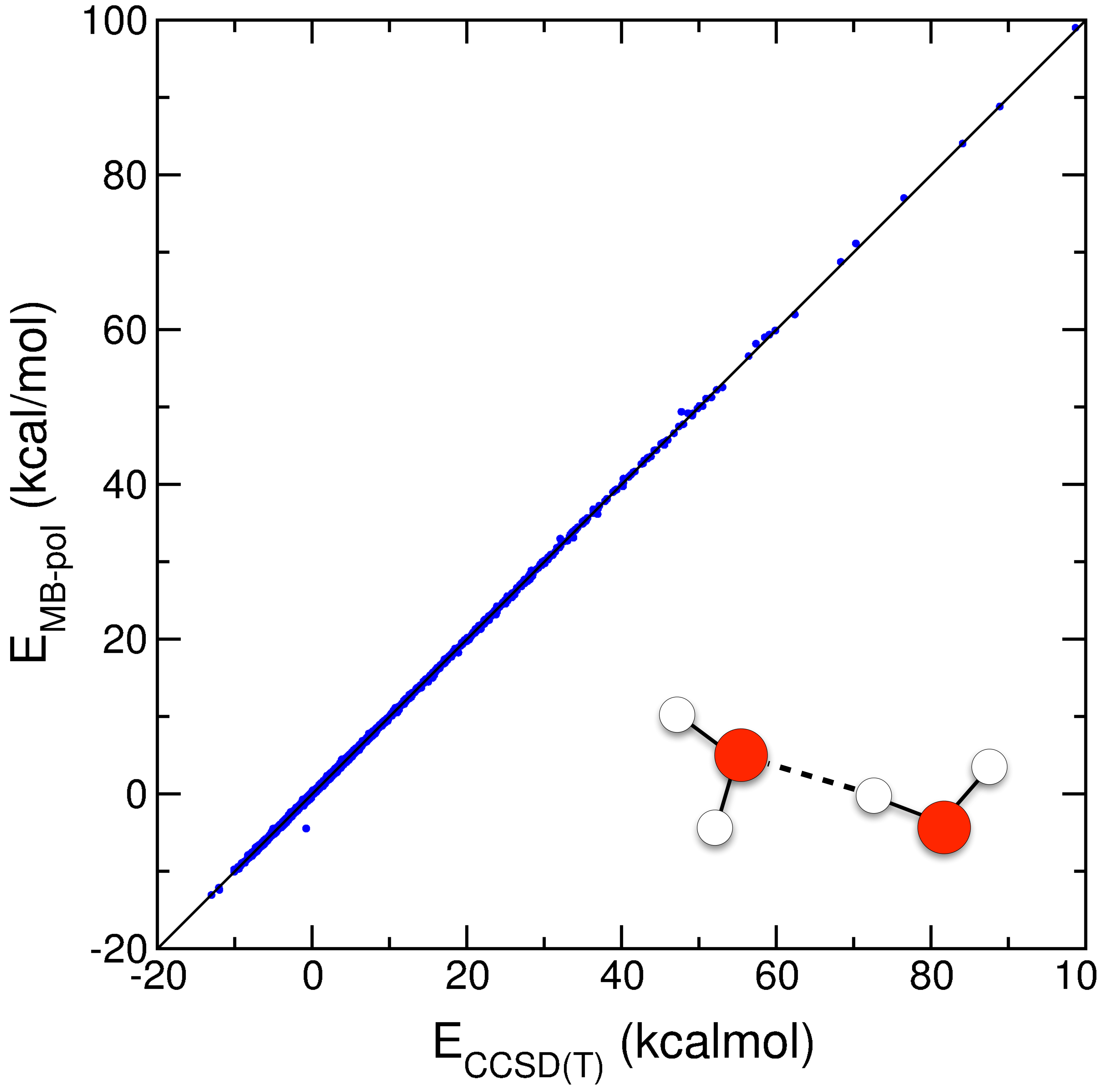


FP, Acc. Chem. Res. 49, 1844 (2016)

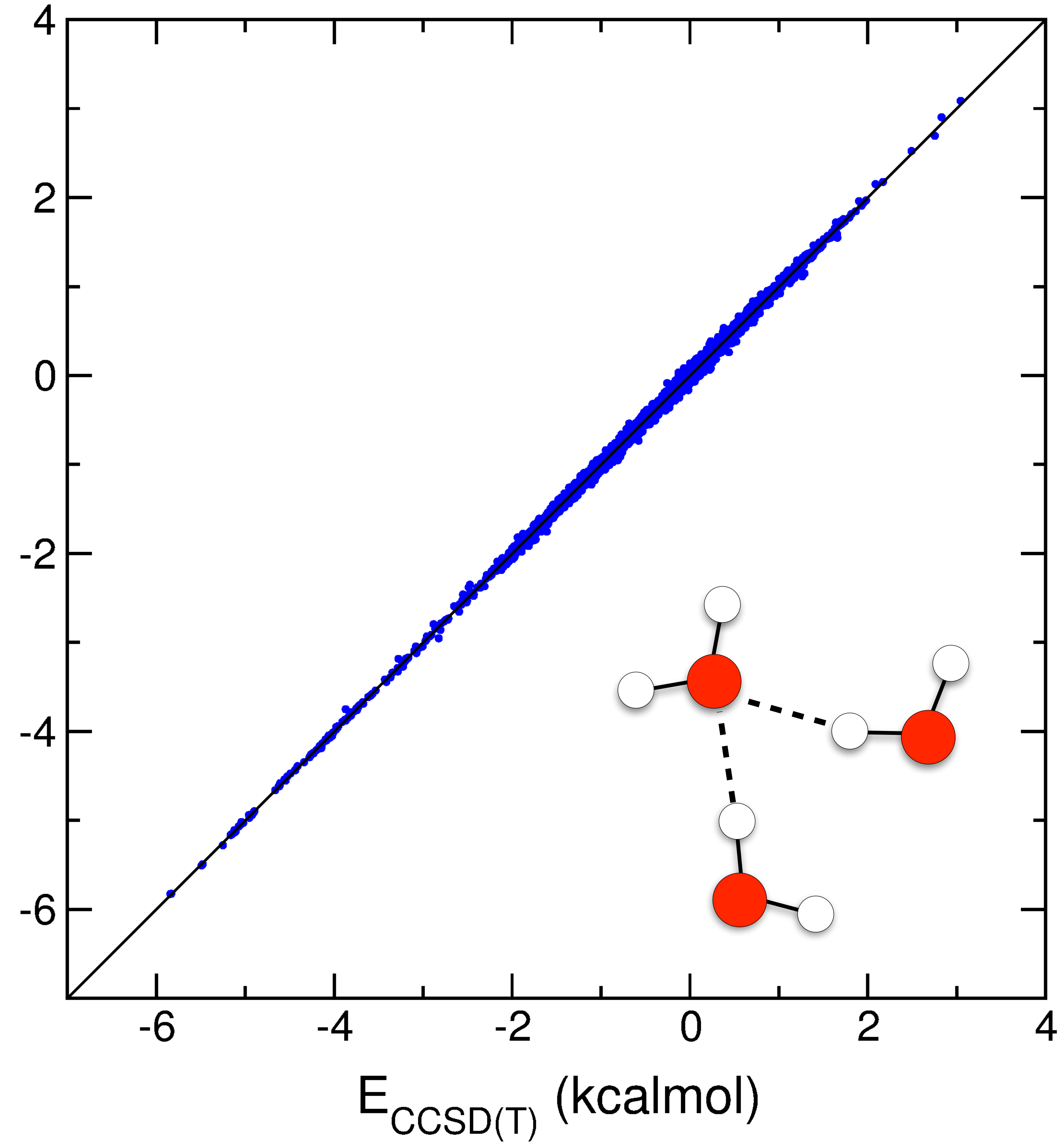
Many-Body Models: Do They Work?

$$E_N = \sum_{i}^{N} V^{1B}(i) + \sum_{i<j}^{N} V^{2B}(i,j) + \sum_{i<j<k}^{N} V^{3B}(i,j,k) + \cdots + V^{NB}(1,\ldots,N)$$

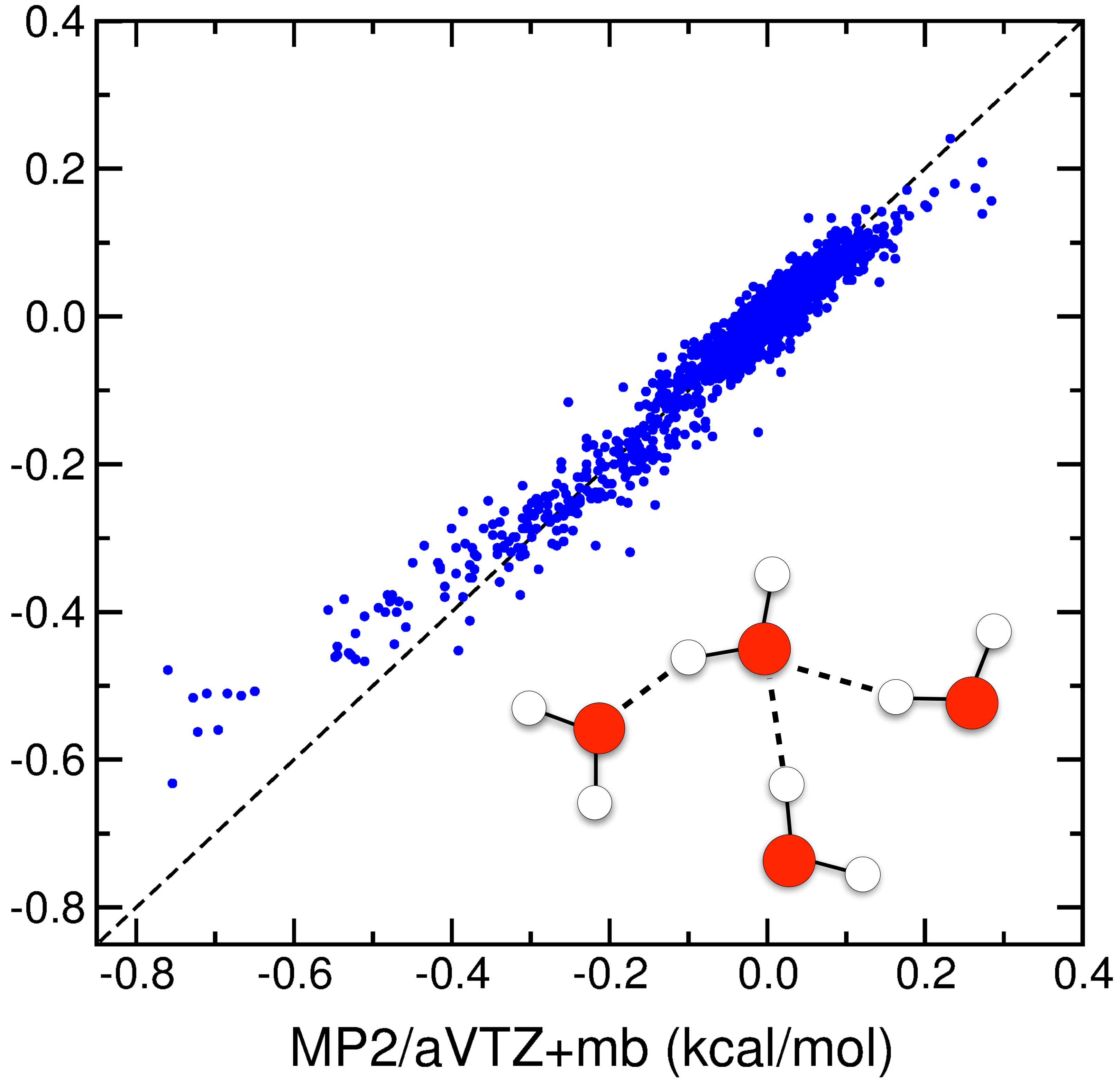
2B energies



3B energies

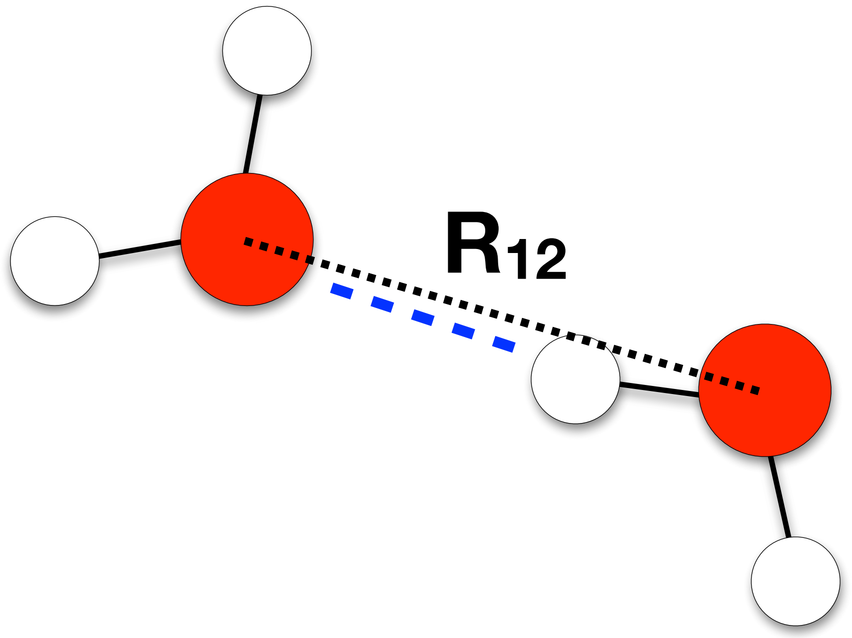


4B energies



FP, Acc. Chem. Res. 49, 1844 (2016)

Representability of Low-Order Many-Body Interactions



$$V^{2B} = s_2 \boxed{V^{2B}_{sr}} + \boxed{\boxed{V^{2B}_{elec}} + \boxed{V^{2B}_{ind}} + \boxed{V^{2B}_{disp}}}$$

short-range interactions

permanent electrostatics

2B induction

2B dispersion

from physics of isolated monomers

Permutationally invariant polynomials

Braams & Bowman, Int. Rev. Phys. Chem. 28, 577 (2009)

$$V^{2B}_{sr} = V^{2B}_{PIP} = \sum_{l} A_l \eta_l = A_1\xi_a + A_2\xi_b + A_3\xi_c + \ldots + A_i\xi_a\xi_b + A_j\xi_a\xi_c + A_k\xi_b\xi_c + \ldots + A_q\xi_a\xi_b\xi_c + \ldots + \ldots$$

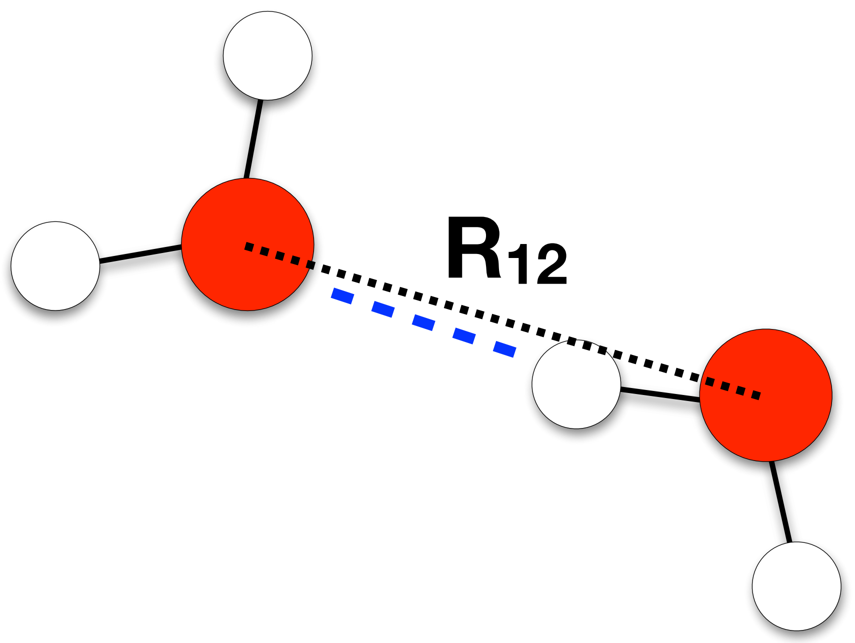
Morse variables: $\xi_a = e^{-kr_a}$

$a = r_m - r_n$

Ridge regression: $\chi^2 = \sum_{n \in S} [V^{2B}_{PIP} - V^{2B}_{sr,CCSD(T)}]^2 + \Gamma^2 \sum_{l} A_l^2$

Nguyen, Székely, Imbalzano, Behler, Csányi, Ceriotti & FP, J. Chem. Phys. 148, 241725 (2018)

Representability of Low-Order Many-Body Interactions



permanent electrostatics

2B dispersion

$$V^{2B} = s_2 V^{2B}_{sr} + V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$

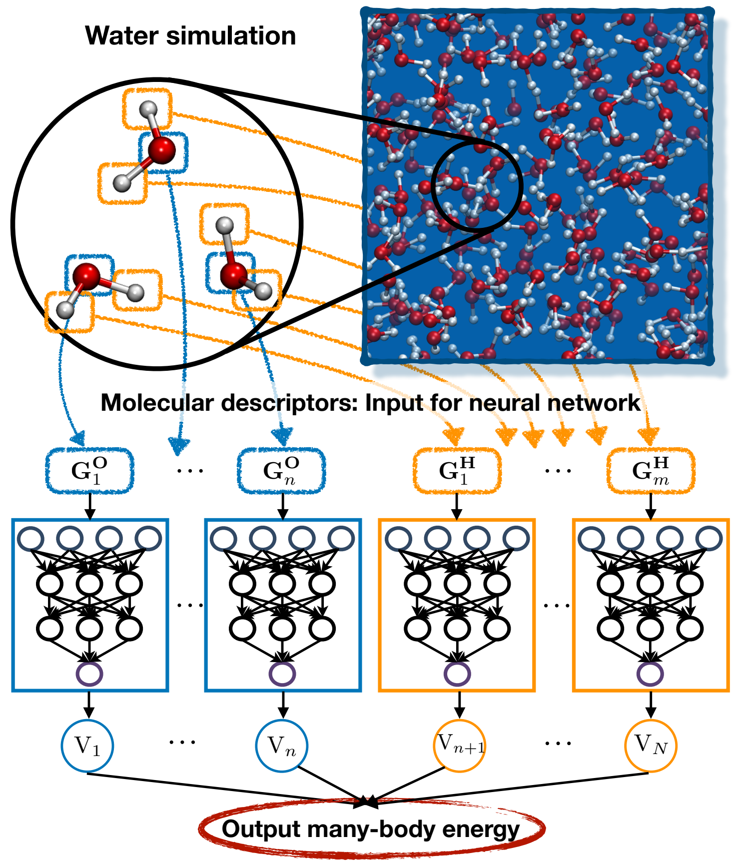
short-range interactions

2B induction

from physics of isolated monomers

Neural Networks

Behler & Parrinello, Phys. Rev. Lett. 98, 146401 (2007)



radial functions

$$G_i^{rad} = \sum_{j \neq i} e^{-(R_{ij} - R_s)^2}$$

angular functions

$$G_i^{ang} = 2^{1-\zeta} \sum_{\substack{j \neq i \\ k \neq i,j}} (1 + \lambda \cos \theta_{ijk})^{\zeta} \times e^{-\eta (R_{ij} + R_{ik} + R_{jk})^2}$$

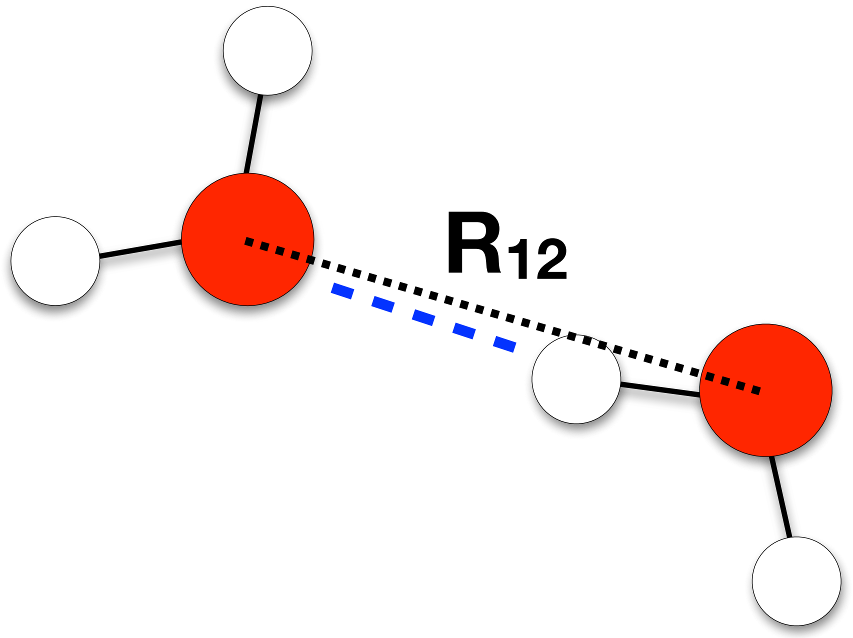
1 input, 3 hidden, and 1 output layers

2 networks: one for H and one for O

activation functions: hyperbolic tangent / linear

Nguyen, Székely, Imbalzano, Behler, Csányi, Ceriotti & FP, J. Chem. Phys. 148, 241725 (2018)

Representability of Low-Order Many-Body Interactions



$$V^{2B} = s_2 V^{2B}_{sr} + V^{2B}_{elec} + V^{2B}_{ind} + V^{2B}_{disp}$$

permanent electrostatics

2B dispersion

short-range interactions

2B induction

from physics of isolated monomers

Gaussian Approximation Potentials

Bartók, Payne, Kondor & Csányi, Phys. Rev. Lett. 104, 136403 (2010)

atomic energies

$$\varepsilon\left(\boldsymbol{R}\right) = \sum_k b_k K(\boldsymbol{R}, \boldsymbol{R}_k),$$

kernel

$$K(\boldsymbol{R}, \boldsymbol{R}') = \left(\sum_{\alpha\beta n_1 n_2 l} p^{\alpha\beta}_{n_1 n_2 l}(\boldsymbol{R}) p^{\alpha\beta}_{n_1 n_2 l}(\boldsymbol{R}') \right)^{\zeta},$$

local atomic neighbor densities

$$\rho^{\alpha}_i(\boldsymbol{r}) = \sum_{nlm} c^{\alpha}_{nlm} g_n(|\boldsymbol{r}|) Y_{lm}(\hat{\boldsymbol{r}}).$$

Nguyen, Székely, Imbalzano, Behler, Csányi, Ceriotti & FP, J. Chem. Phys. 148, 241725 (2018)

Representability of Low-Order Many-Body Interactions

$$E_N = \sum_{i}^{N} V^{1B}(i) + \boxed{\sum_{i<j}^{N} V^{2B}(i,j)} + \boxed{\sum_{i<j<k}^{N} V^{3B}(i,j,k)} + \cdots + V^{NB}(1,\ldots,N)$$

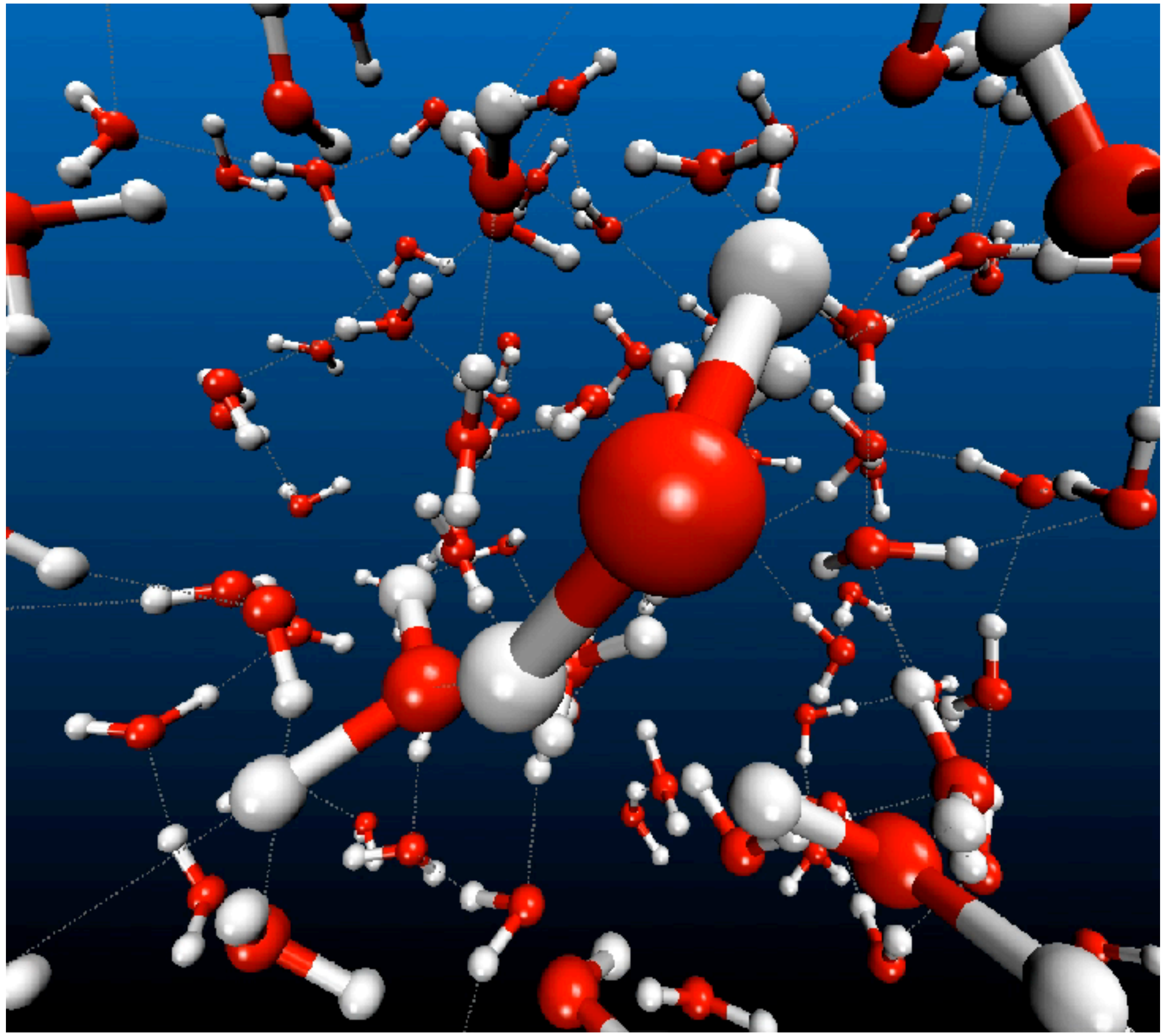
	2B (kcal/mol)			**3B (kcal/mol)**		
	Training	Validation	Test	Training	Validation	Test
PIP	0.0349	0.0449	0.0494	0.0262	0.0463	0.0465
BPNN	0.0493	0.0784	0.0792	0.0318	0.0658	0.0634
GAP	0.0176	0.0441	0.0539	0.0052	0.0514	0.0517

Nguyen, Székely, Imbalzano, Behler, Csányi, Ceriotti & FP, J. Chem. Phys. 148, 241725 (2018)

Molecular Dynamics

MB-pol $\Longleftrightarrow$ Born-Oppenheimer potential energy surface

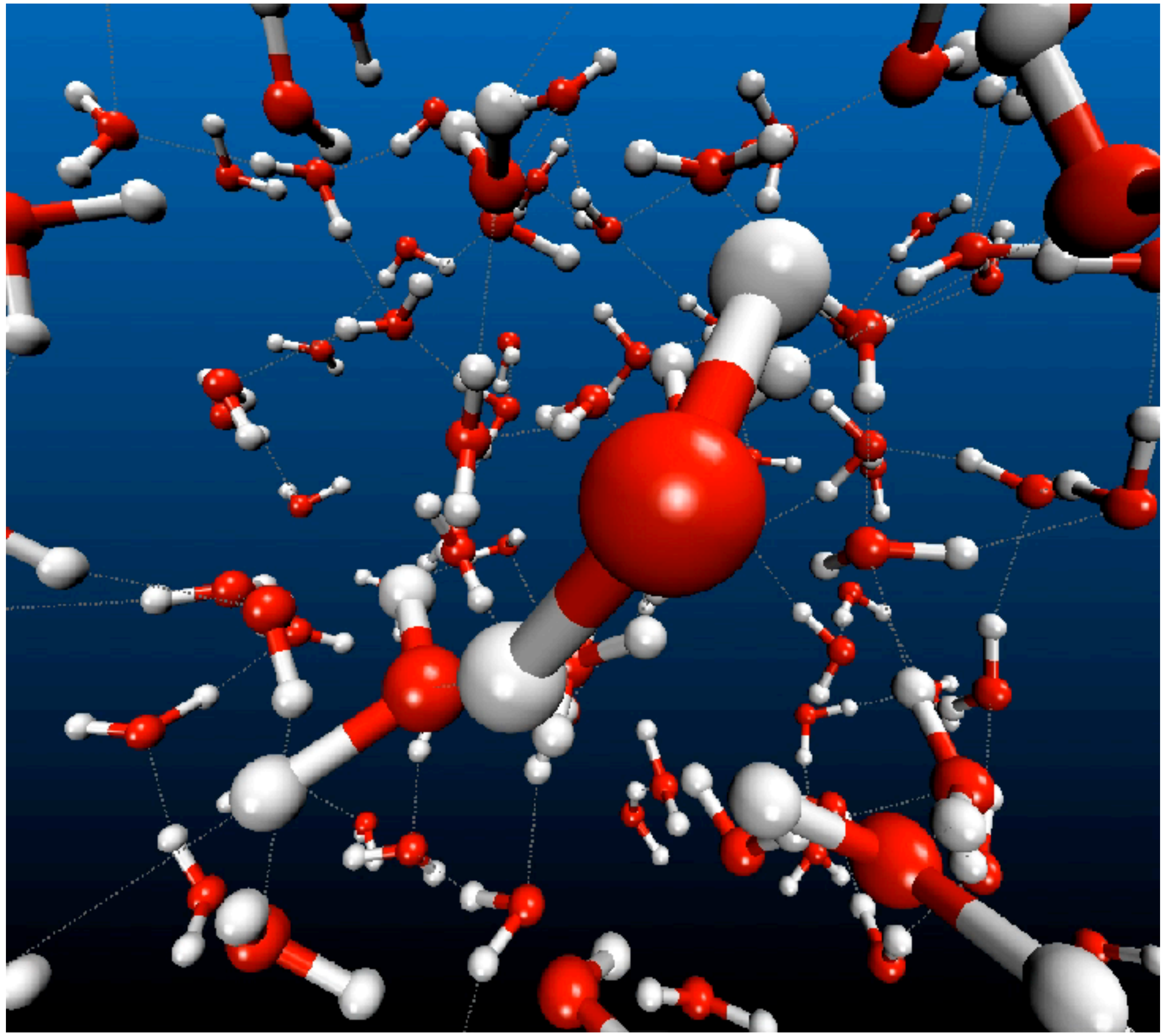
Nuclear quantum effects must be included explicitly in molecular simulations



Molecular Dynamics

MB-pol $\Longleftrightarrow$ Born-Oppenheimer potential energy surface

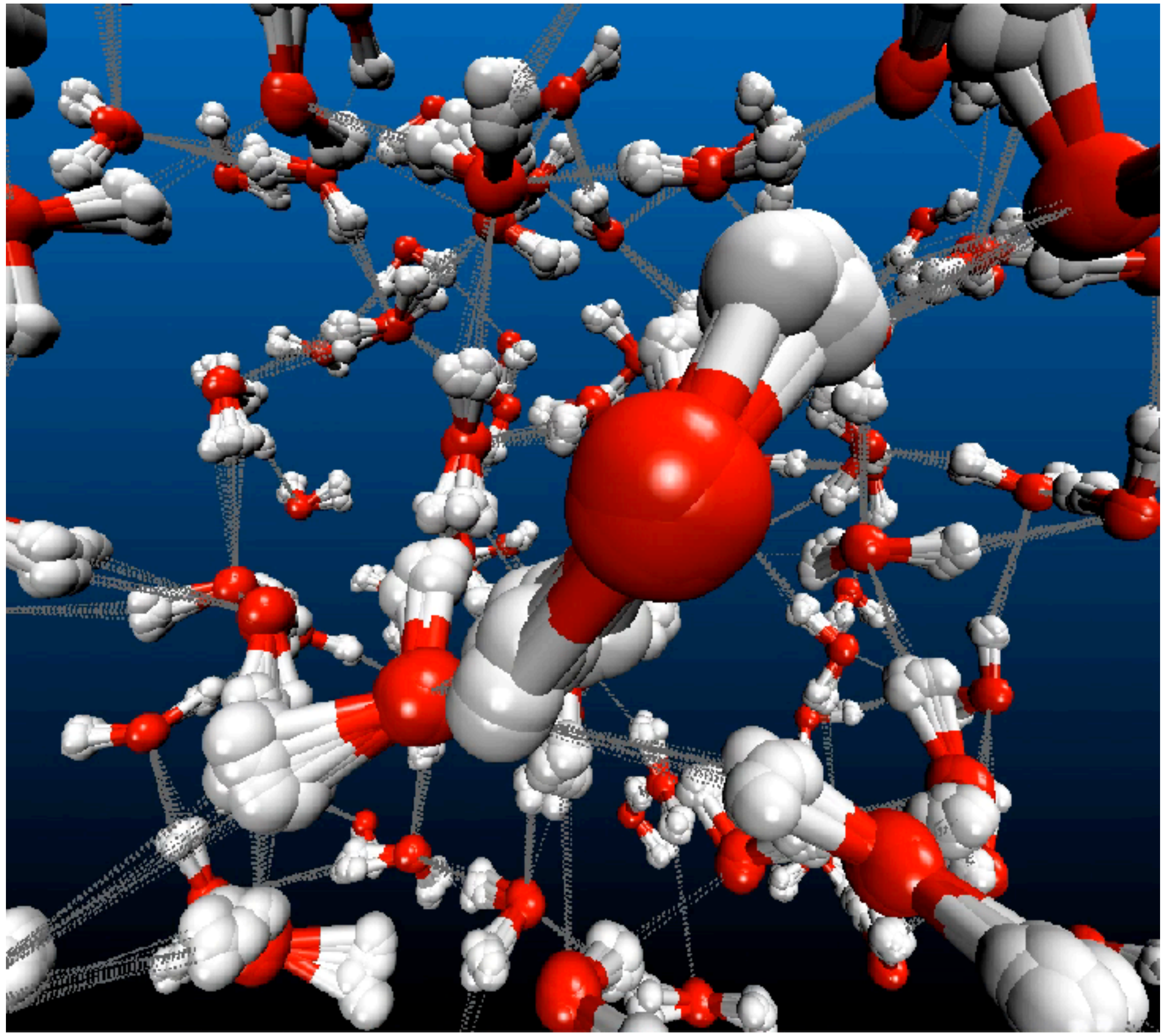
Nuclear quantum effects must be included explicitly in molecular simulations



Molecular Dynamics

MB-pol $\Longleftrightarrow$ Born-Oppenheimer potential energy surface

Nuclear quantum effects must be included explicitly in molecular simulations

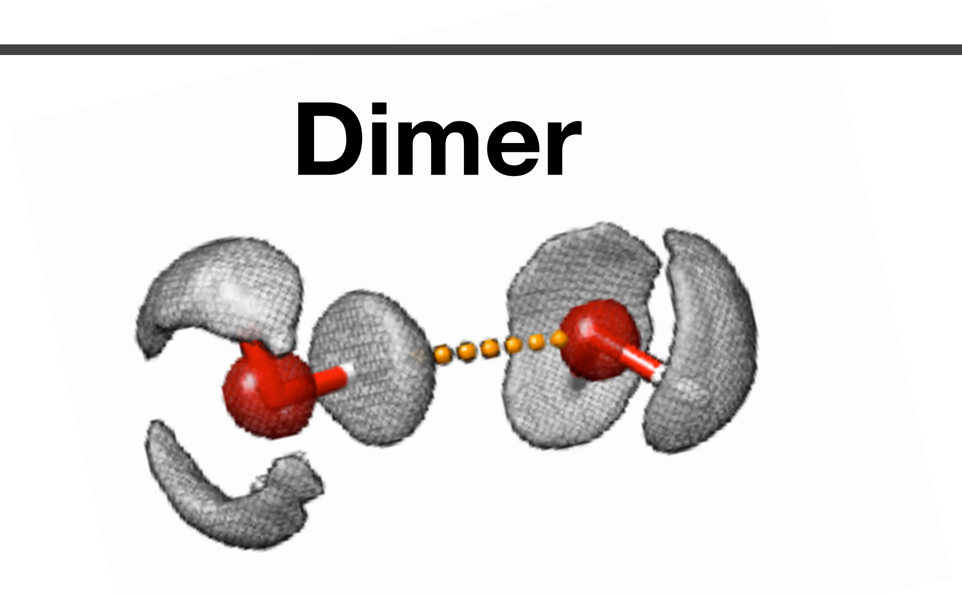


Methods based on:

- Basis set expansions
  with Claude Leforestier (Montpellier), Ryan Steele (Utah)
  (clusters: anharmonic vibrational spectra)
- Path-integral molecular dynamics
  (condensed phase: structure and thermodynamics)
- Centroid molecular dynamics
  (condensed phase: dynamics)
- Ring polymer instanton
  with Jeremy Richardson (ETH)
  (tunneling splittings)

MB-pol: Chemical Accuracy Across All Phases

Dimer



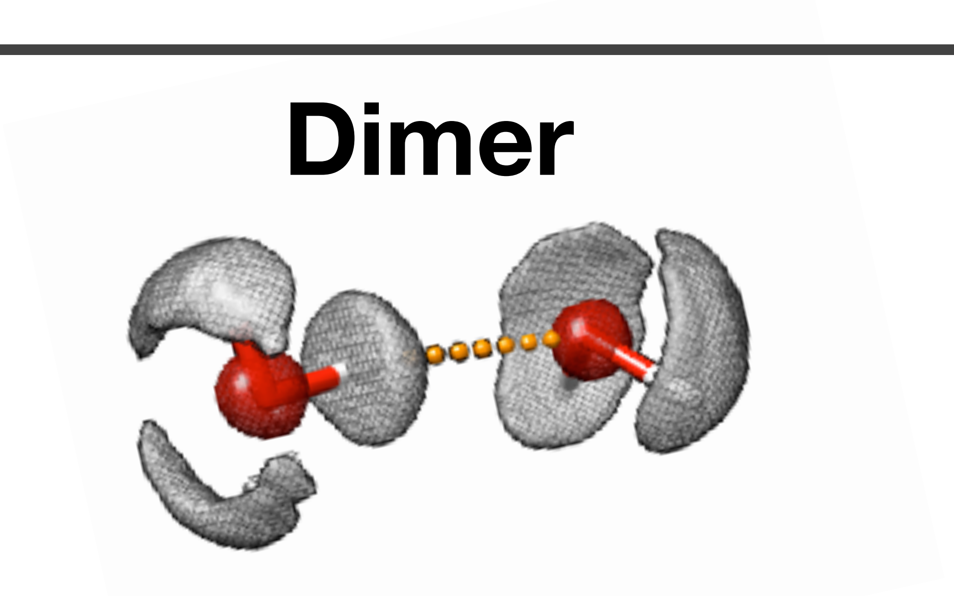
Mode		Experiment (cm^{-1})	MB-pol (cm^{-1})
intermolecular stretch	2	153.62	154.77
	1		149.05
acceptor twist	1		129.49
	2	120.19	119.23
acceptor wag	2	108.89	109.14
	1	107.93	108.76
donor torsion	1		113.18
	2	64.52	61.24
ground state	2	11.18	11.88
	1	0	0

J. Chem. Theory Comput. 9, 4844 (2013)
J. Chem. Theory Comput. 10, 1599 (2014)

FP, Acc. Chem. Res. 49, 1844 (2016)

MB-pol: Chemical Accuracy Across All Phases

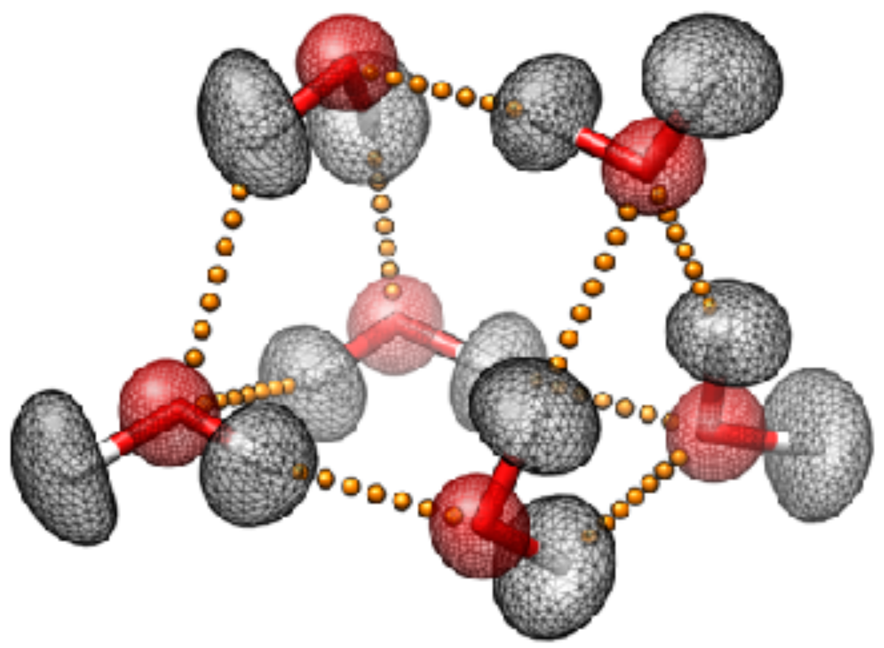
Dimer

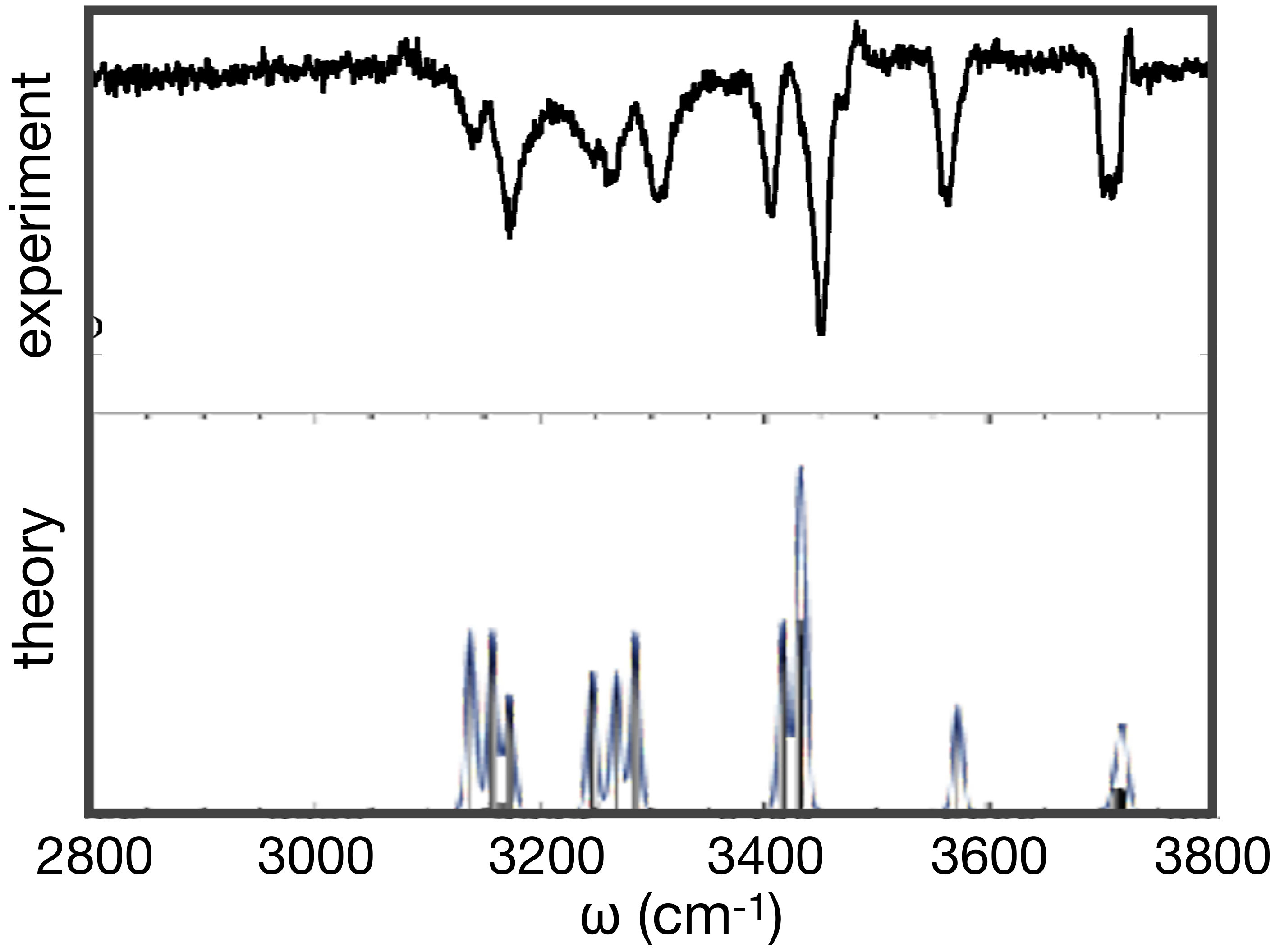


Mode		Experiment (cm^{-1})	MB-pol (cm^{-1})
intermolecular stretch	2	153.62	154.77
	1		149.05
acceptor twist	1		129.49
	2	120.19	119.23
acceptor wag	2	108.89	109.14
	1	107.93	108.76
donor torsion	1		113.18
	2	64.52	61.24
ground state	2	11.18	11.88
	1	0	0

J. Chem. Theory Comput. 9, 4844 (2013)
J. Chem. Theory Comput. 10, 1599 (2014)

Clusters



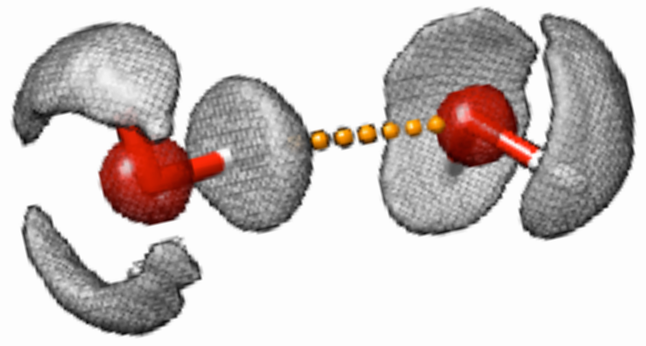


J. Am. Chem. Soc. 134, 11116 (2012)
Chem. Phys. Lett. 580, 1 (2013)
J. Am. Chem. Soc. 139, 7082 (2017)

FP, Acc. Chem. Res. 49, 1844 (2016)

MB-pol: Chemical Accuracy Across All Phases

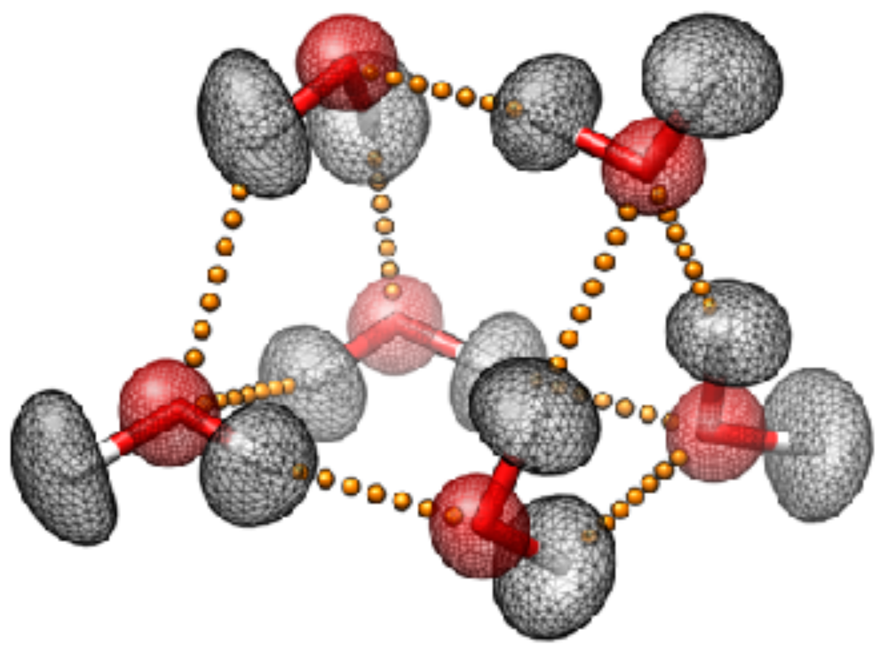
Dimer

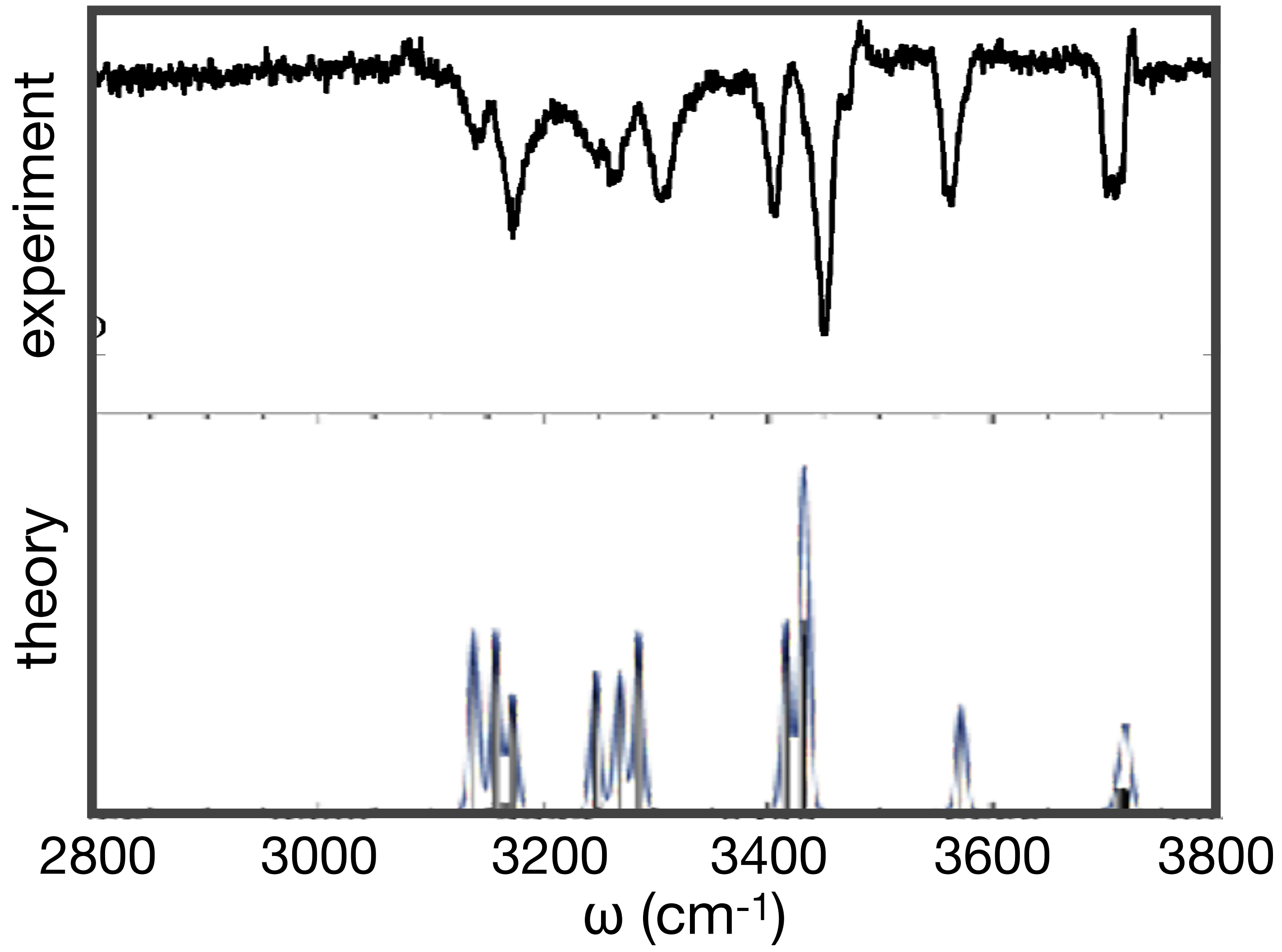


Mode		Experiment (cm^{-1})	MB-pol (cm^{-1})
intermolecular stretch	2	153.62	154.77
	1		149.05
acceptor twist	1		129.49
	2	120.19	119.23
acceptor wag	2	108.89	109.14
	1	107.93	108.76
donor torsion	1		113.18
	2	64.52	61.24
ground state	2	11.18	11.88
	1	0	0

J. Chem. Theory Comput. 9, 4844 (2013)
J. Chem. Theory Comput. 10, 1599 (2014)

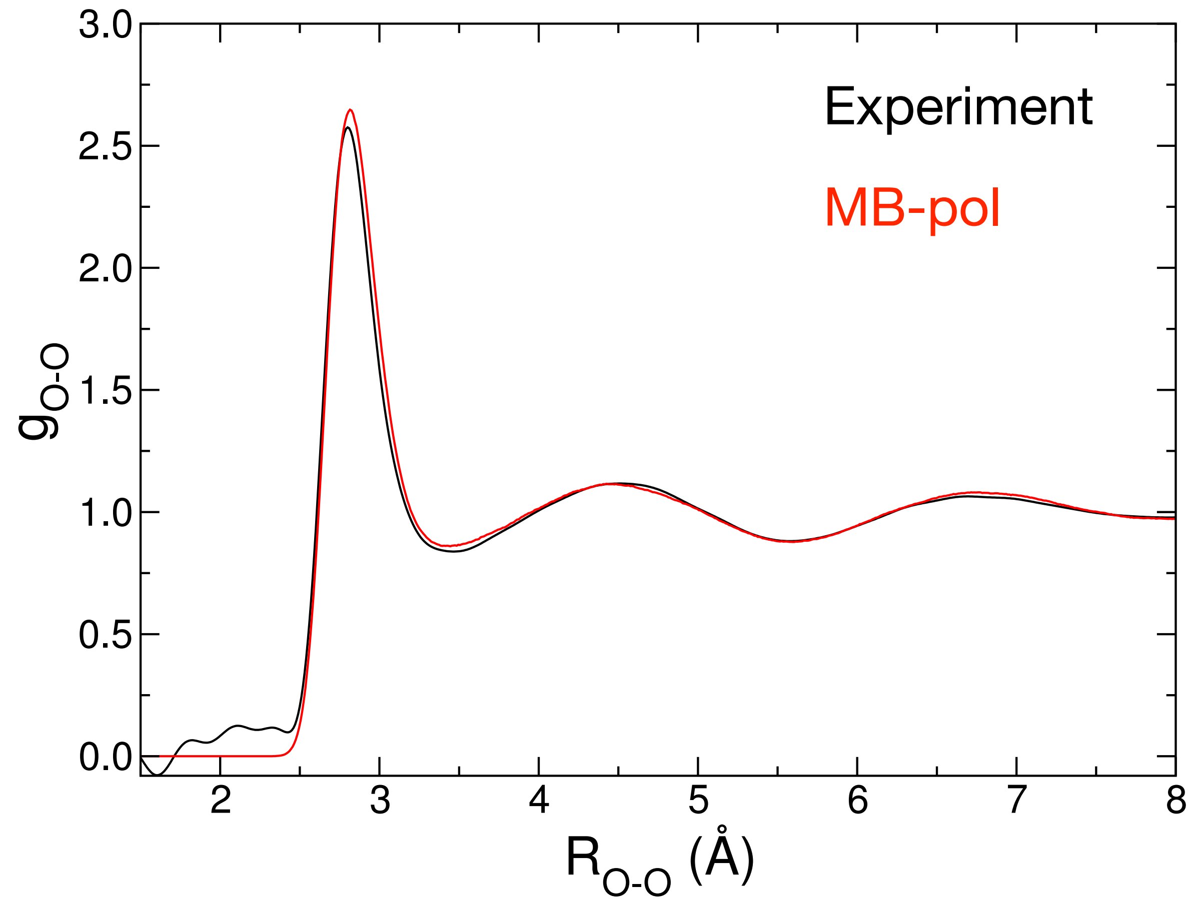
Clusters

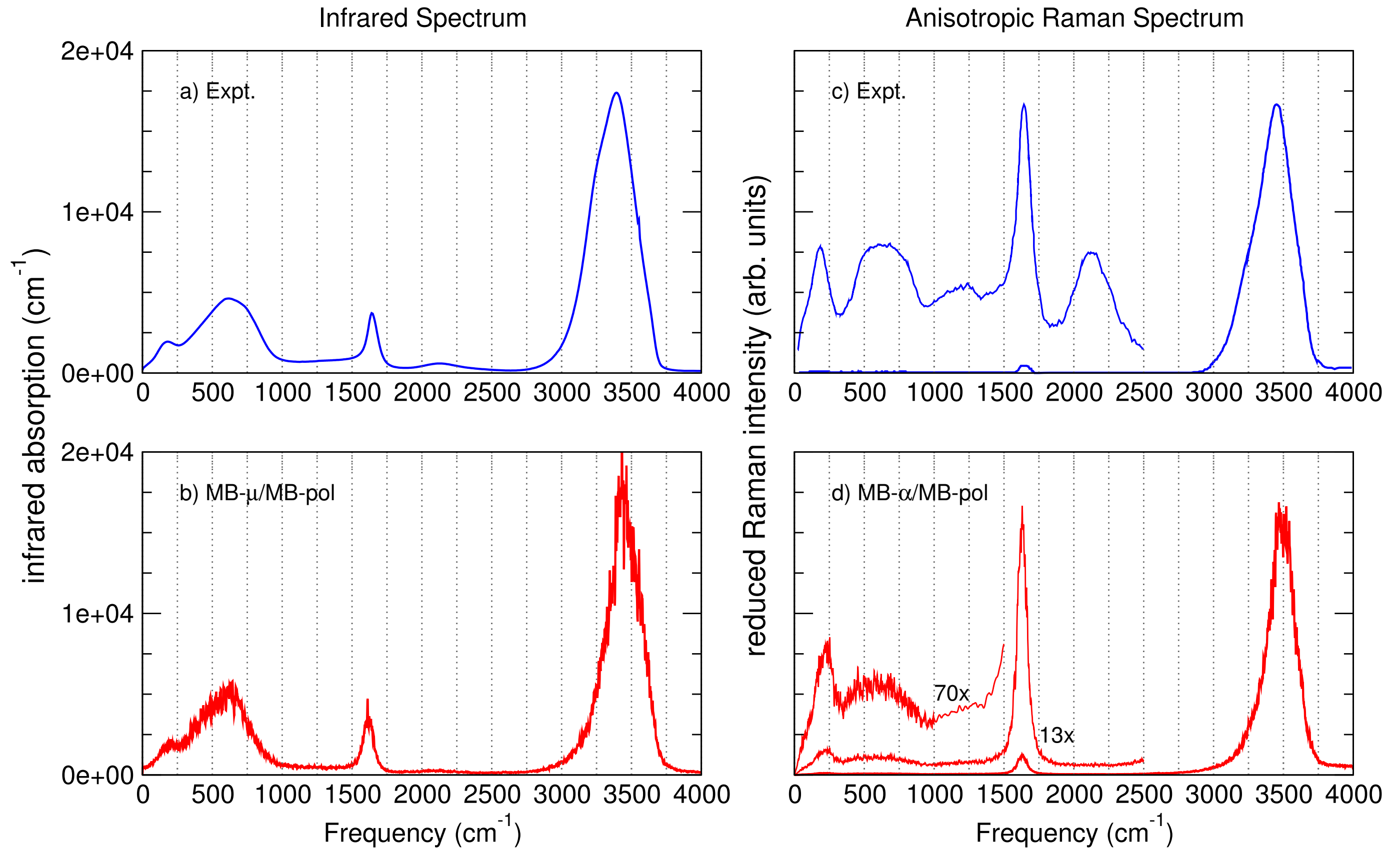




J. Am. Chem. Soc. 134, 11116 (2012)
Chem. Phys. Lett. 580, 1 (2013)
J. Am. Chem. Soc. 139, 7082 (2017)

Bulk



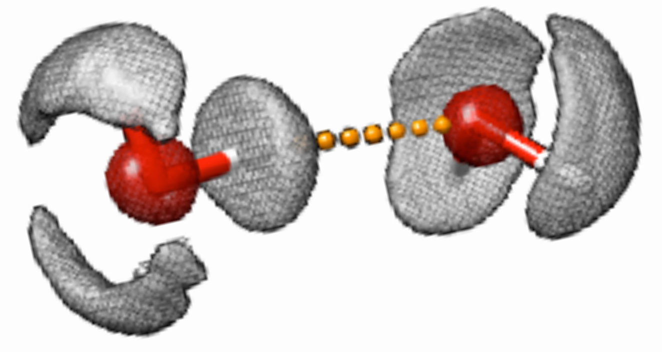


J. Chem. Theory Comput. 10, 2096 (2014)
J. Chem. Theory Comput. 11, 1145 (2014)
J. Phys. Chem. B 120, 8539 (2016)
J. Chem. Phys. 147, 244504 (2017)

FP, Acc. Chem. Res. 49, 1844 (2016)

MB-pol: Chemical Accuracy Across All Phases

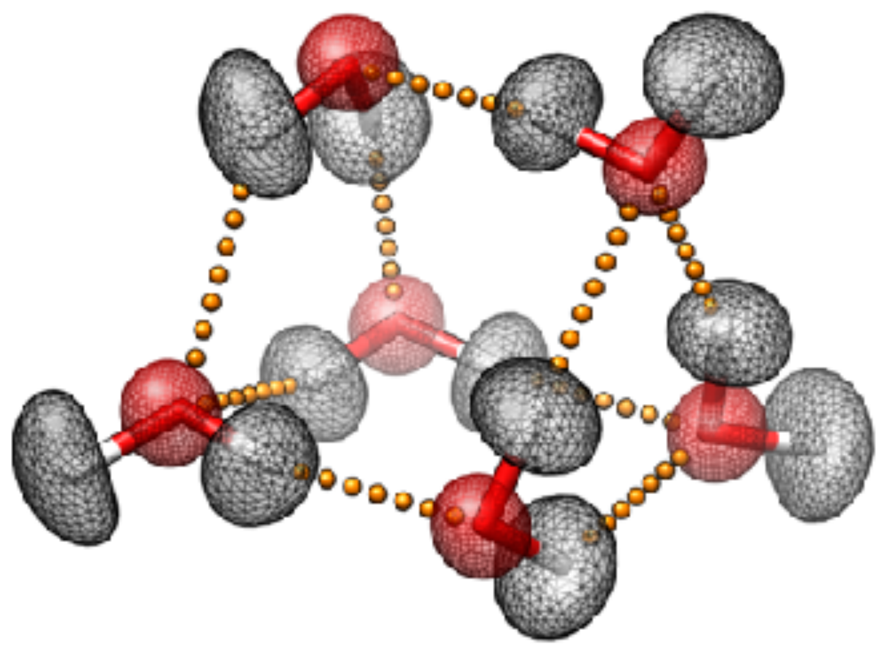
Dimer

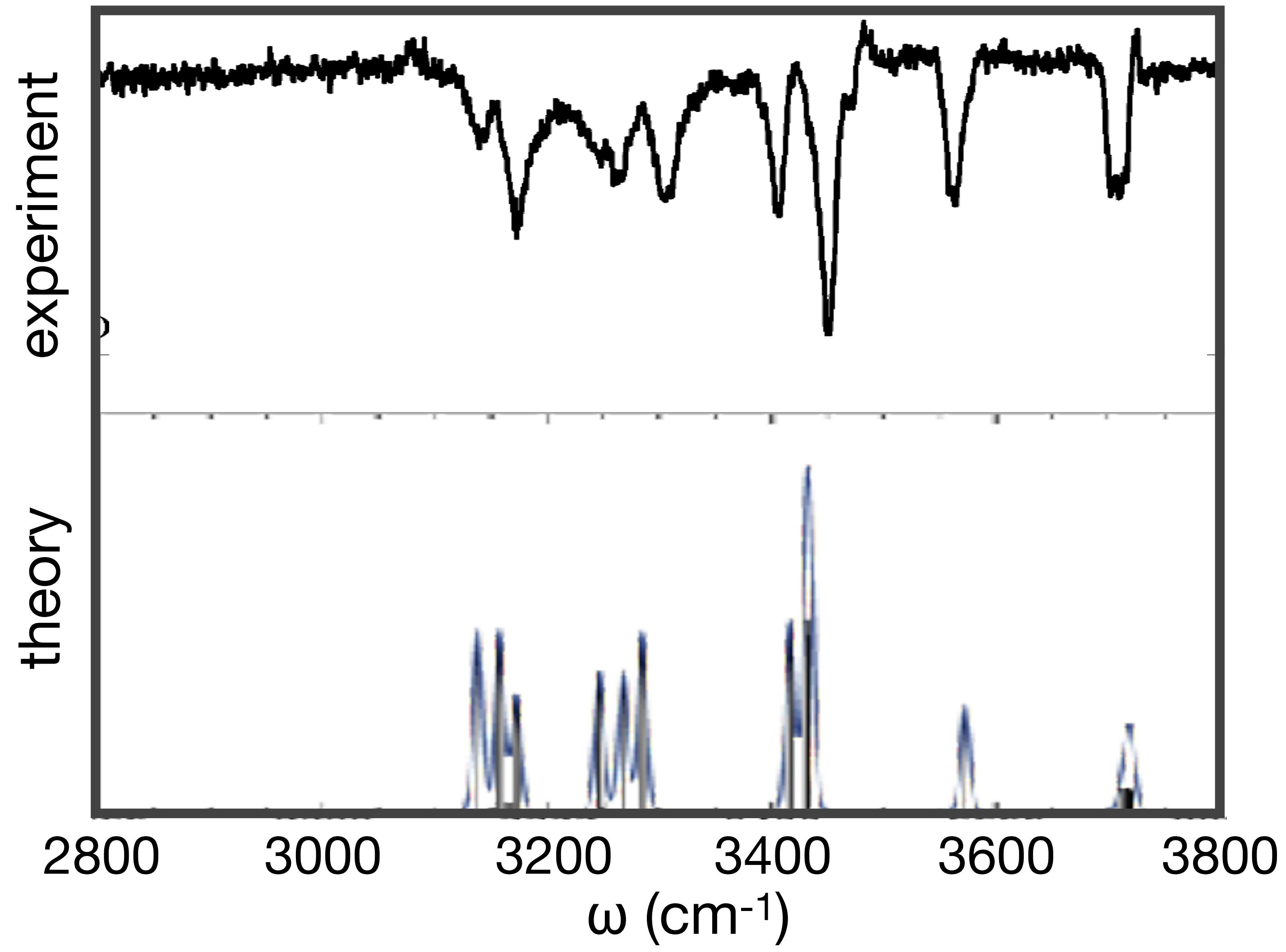


Mode		Experiment (cm^{-1})	MB-pol (cm^{-1})
intermolecular stretch	2	153.62	154.77
	1		149.05
acceptor twist	1		129.49
	2	120.19	119.23
acceptor wag	2	108.89	109.14
	1	107.93	108.76
donor torsion	1		113.18
	2	64.52	61.24
ground state	2	11.18	11.88
	1	0	0

J. Chem. Theory Comput. 9, 4844 (2013)
J. Chem. Theory Comput. 10, 1599 (2014)

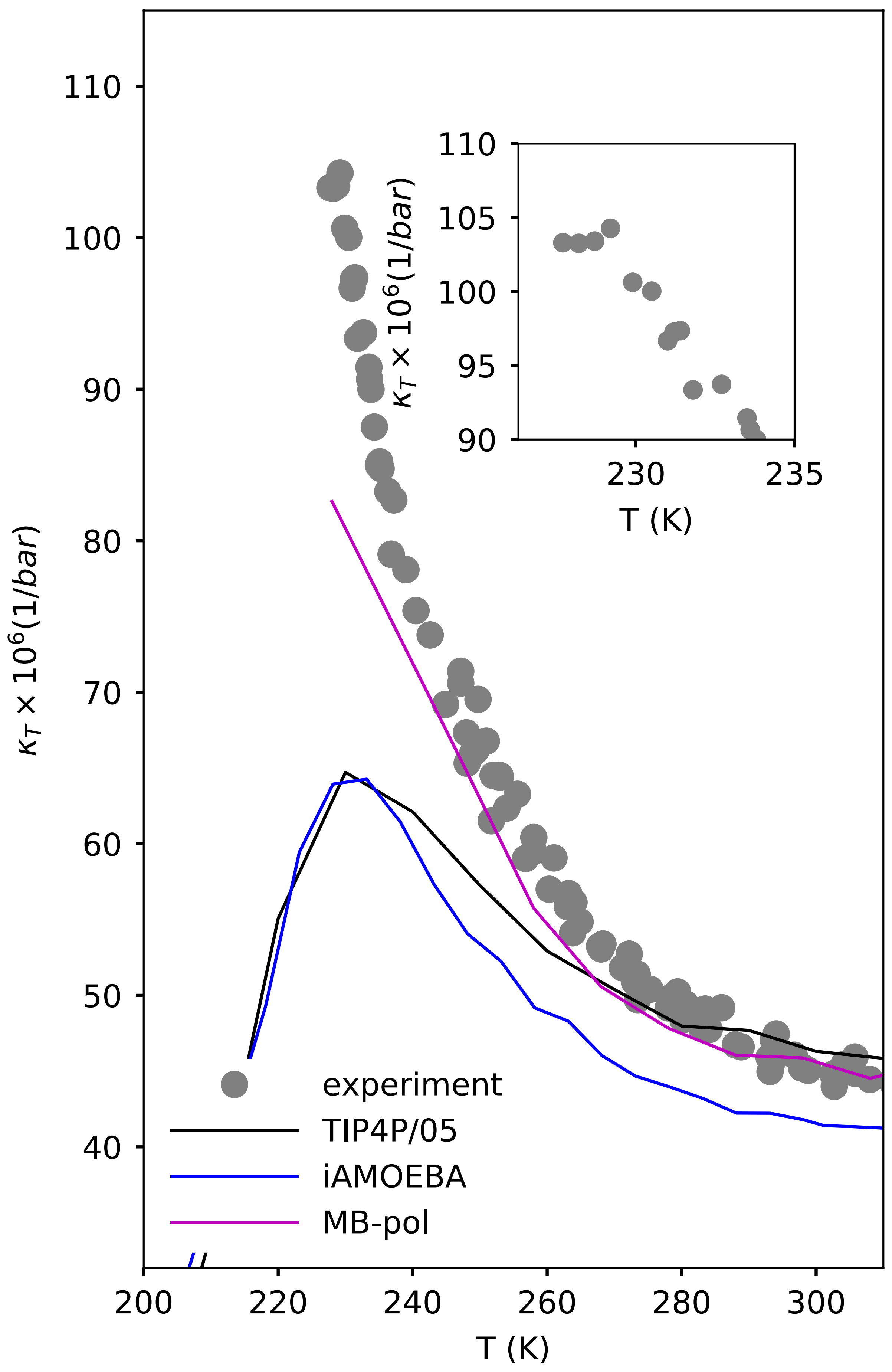
Clusters





J. Am. Chem. Soc. 134, 11116 (2012)
Chem. Phys. Lett. 580, 1 (2013)
J. Am. Chem. Soc. 139, 7082 (2017)

Bulk

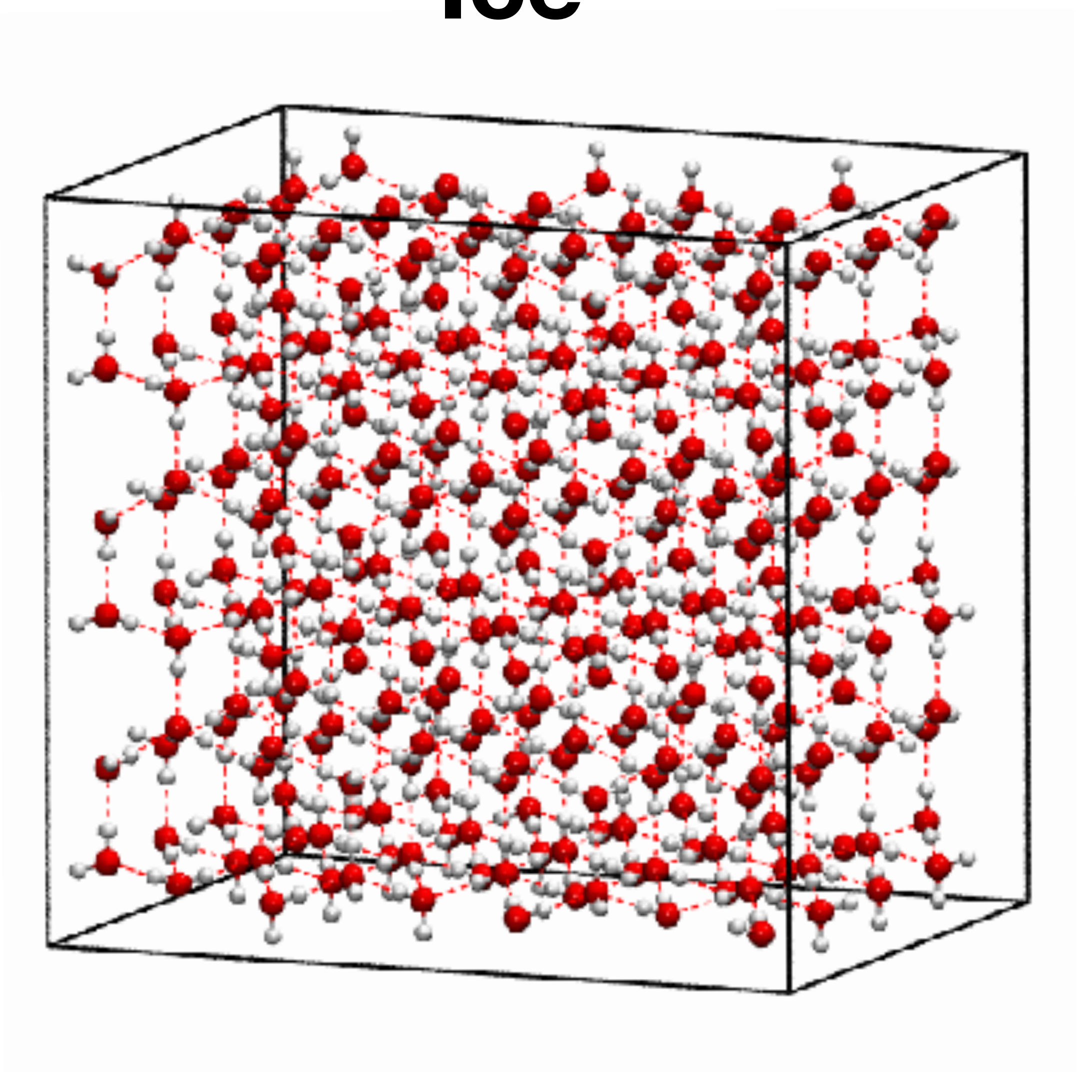


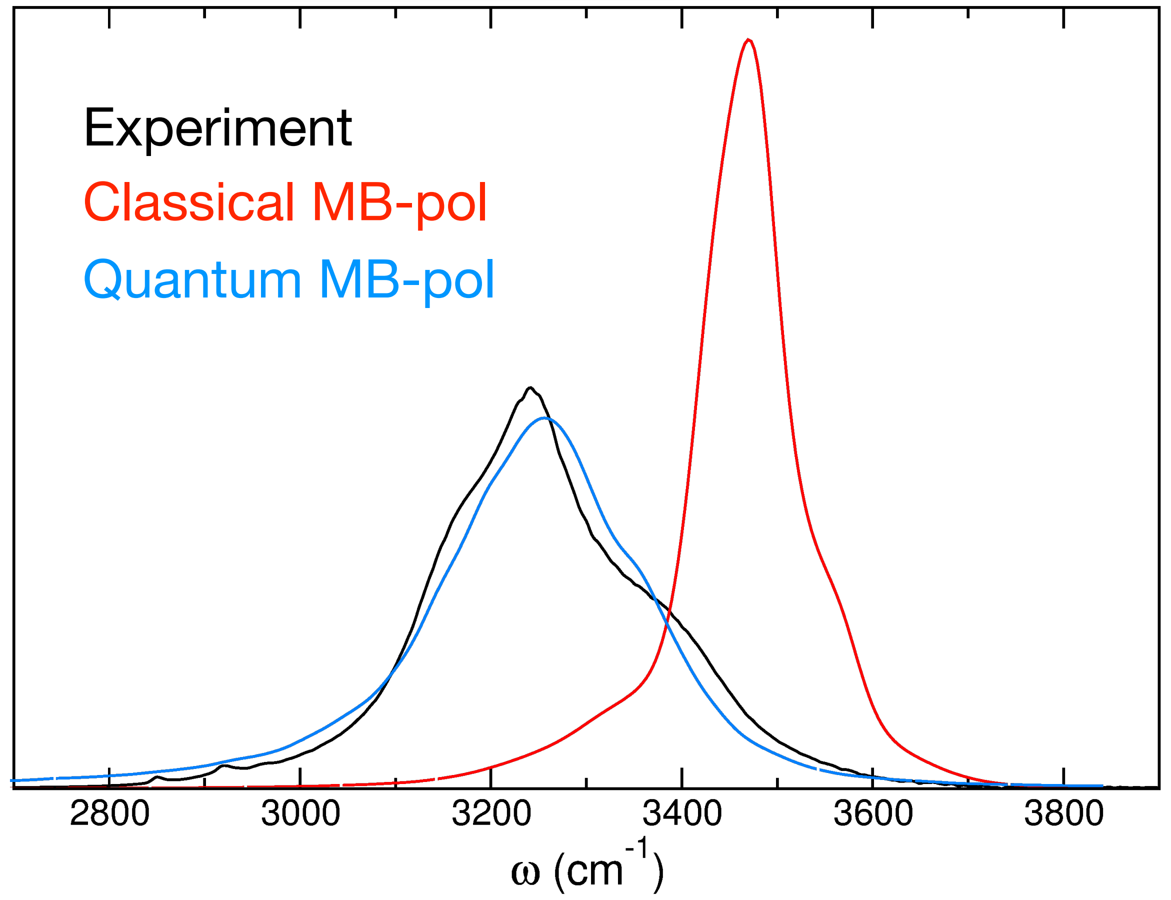
Unpublished
Courtesy Anders Nilsson (Stockholm University)

FP, Acc. Chem. Res. 49, 1844 (2016)

MB-pol: Chemical Accuracy Across All Phases

Ice



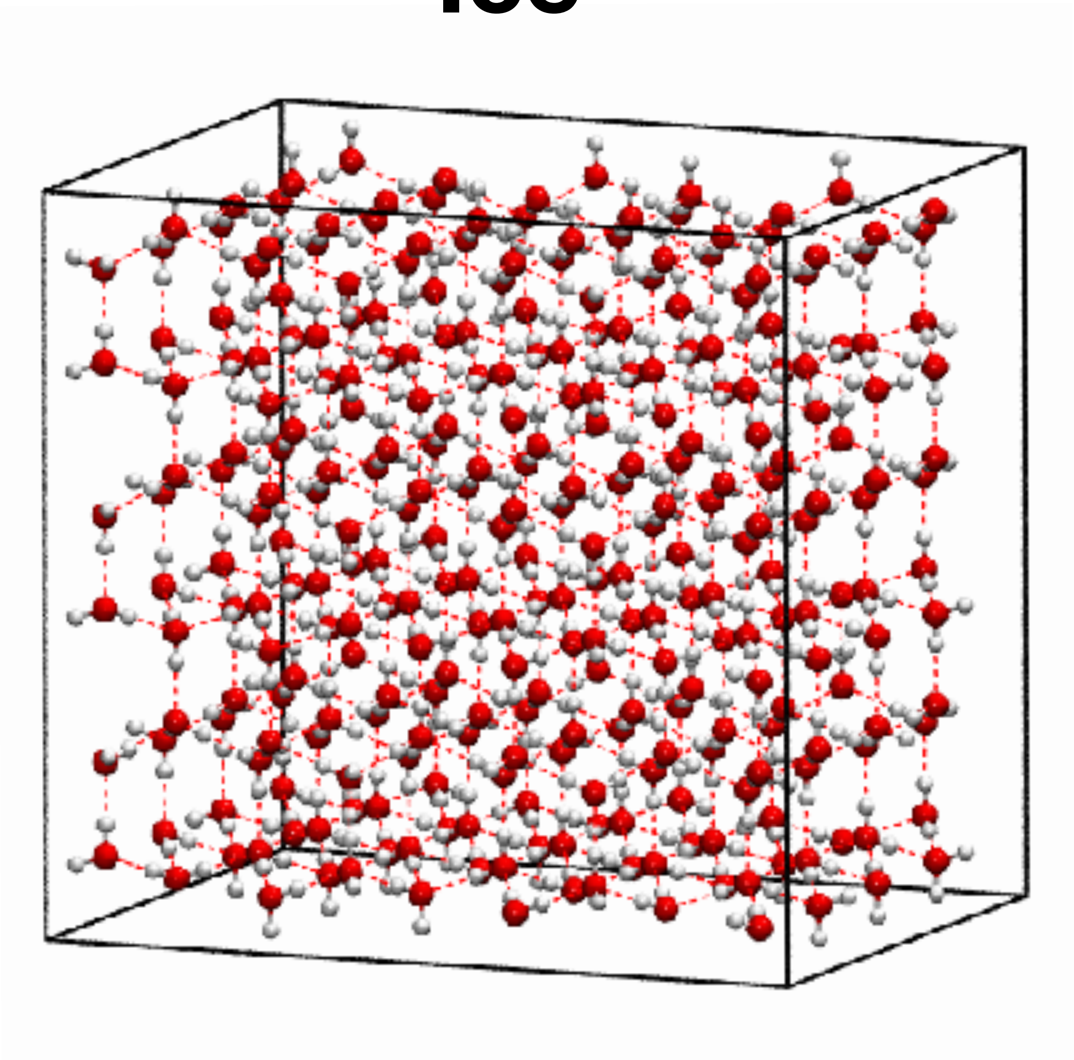


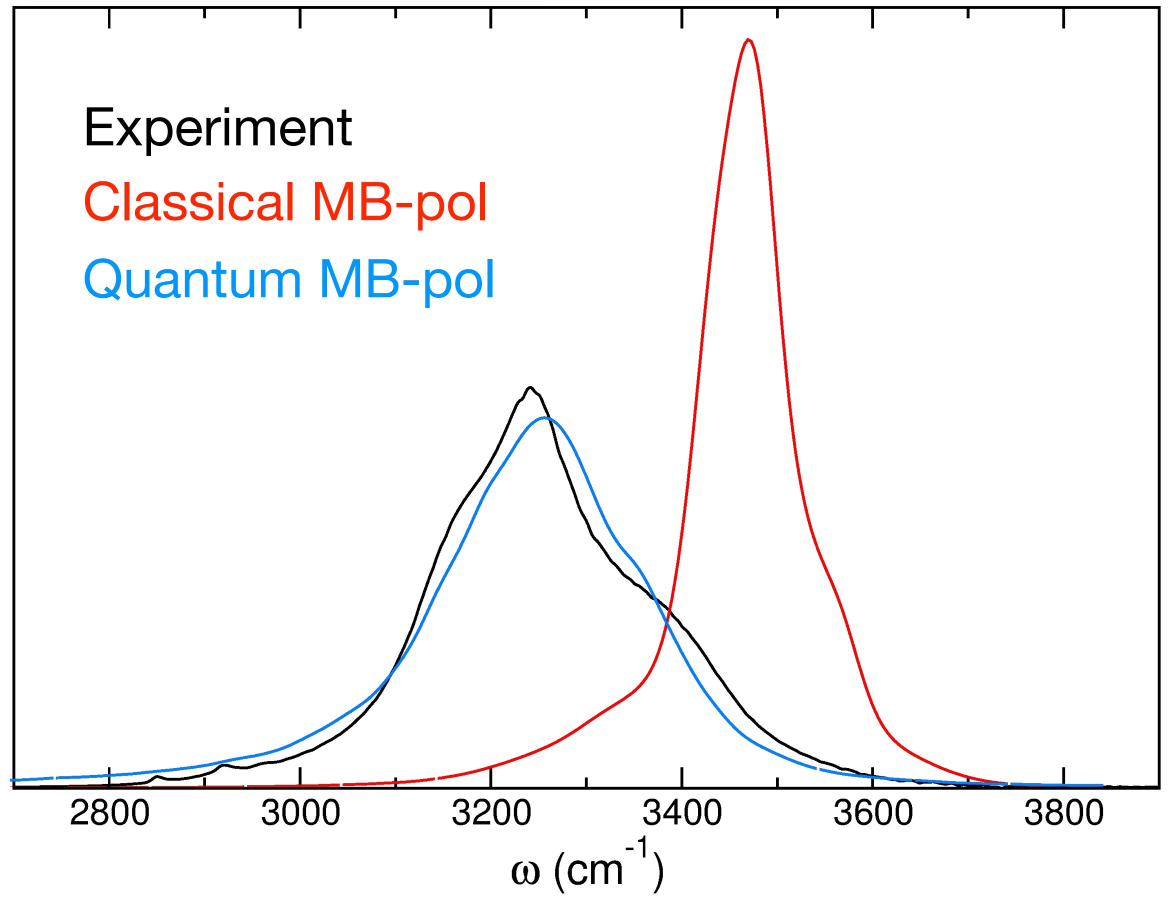
J. Chem. Theory Comput. 13, 1778 (2015)
J. Phys. Chem. Lett. 8, 2579 (2017)

FP, Acc. Chem. Res. 49, 1844 (2016)

MB-pol: Chemical Accuracy Across All Phases

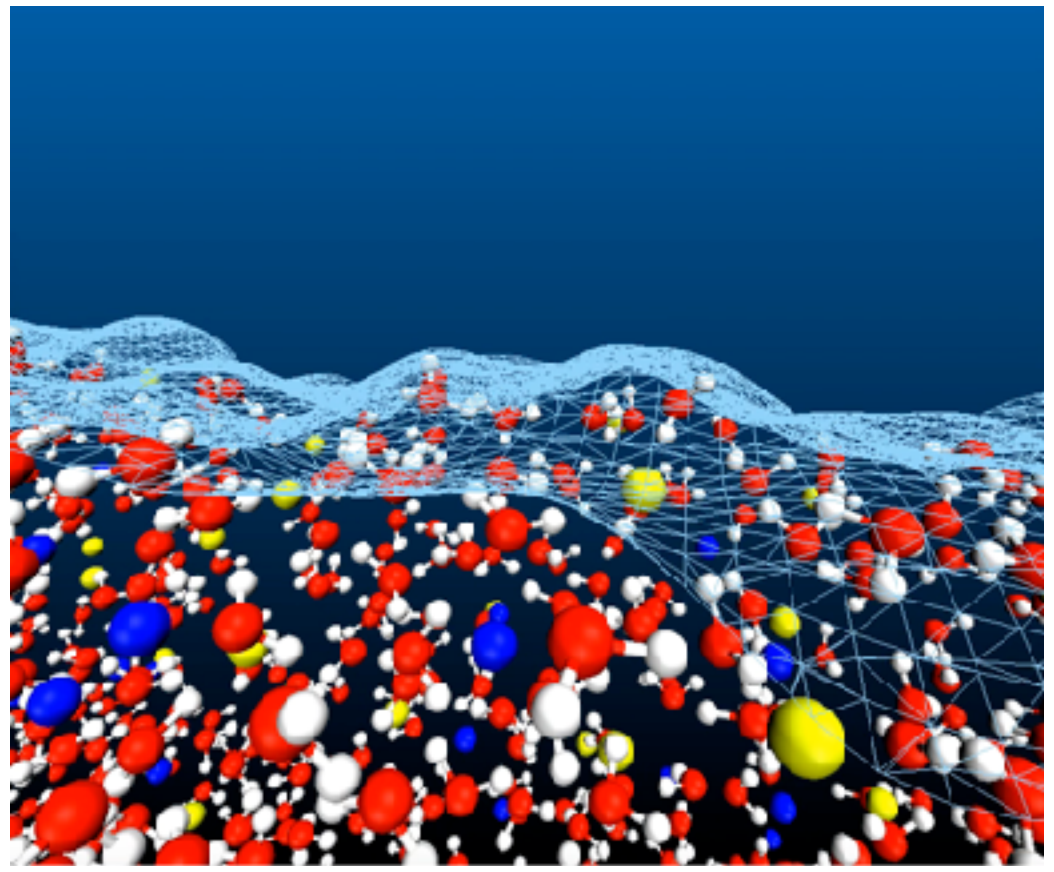
Ice

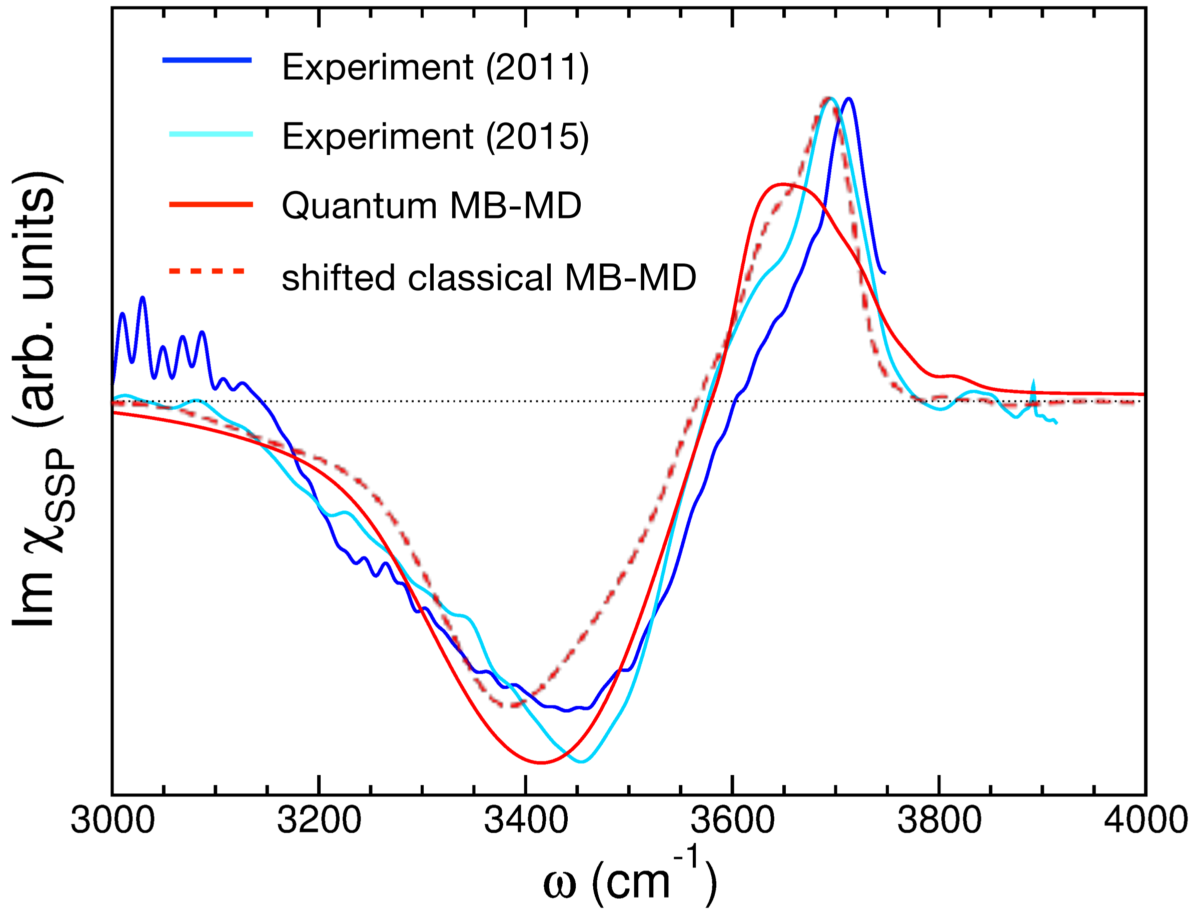




J. Chem. Theory Comput. 13, 1778 (2015)
J. Phys. Chem. Lett. 8, 2579 (2017)

Air/water interface



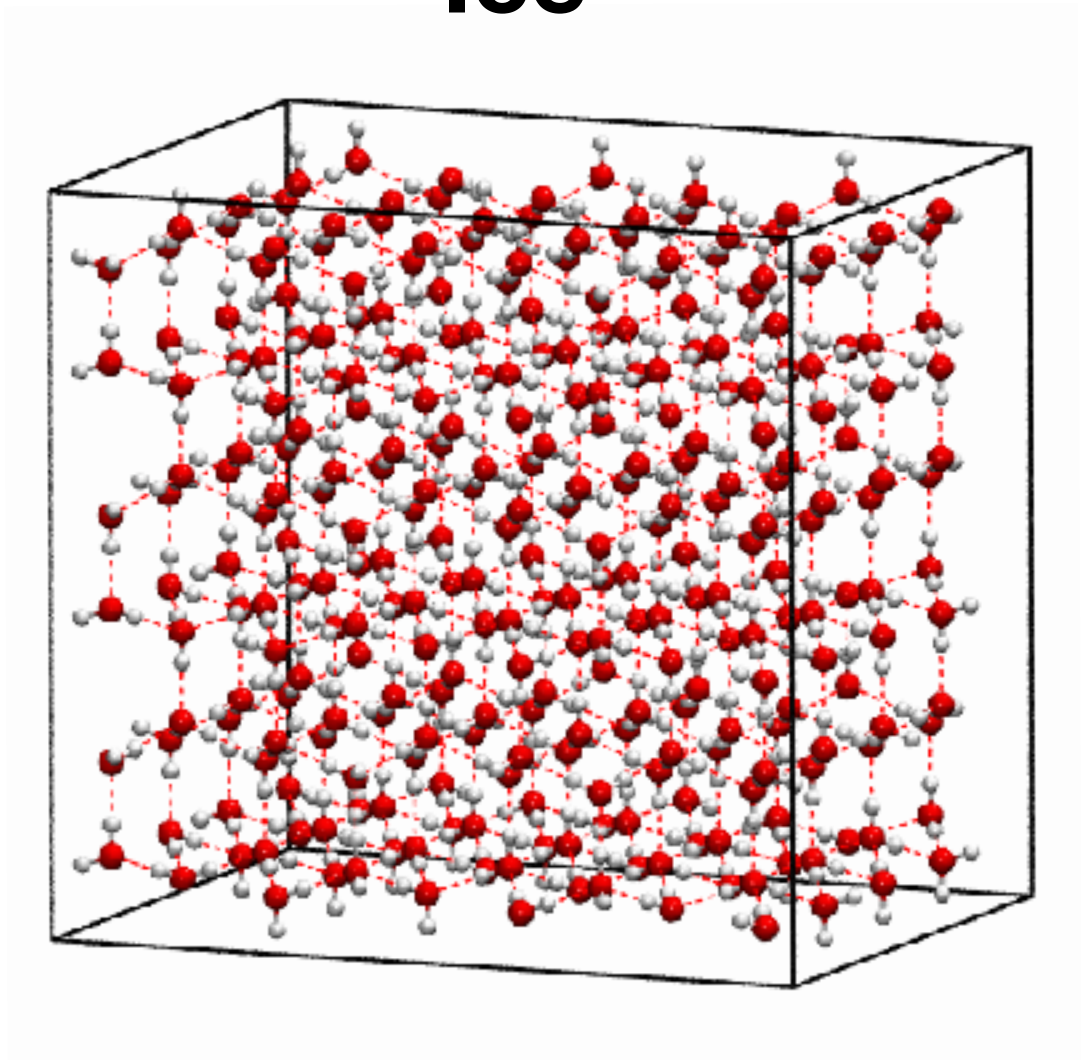


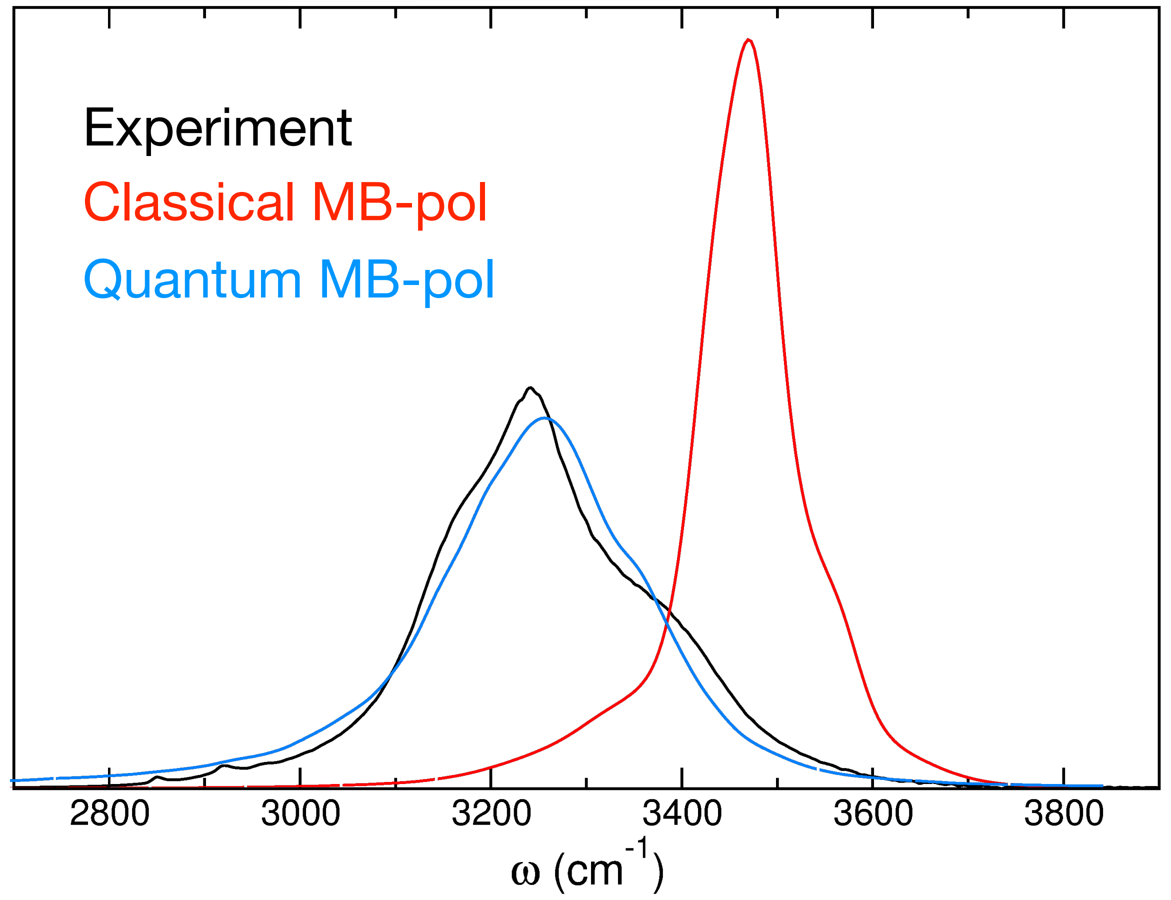
J. Am. Chem. Soc. 138, 3912 (2016)
J. Phys. Chem. B 122, 4356 (2018)

FP, Acc. Chem. Res. 49, 1844 (2016)

MB-pol: Chemical Accuracy Across All Phases

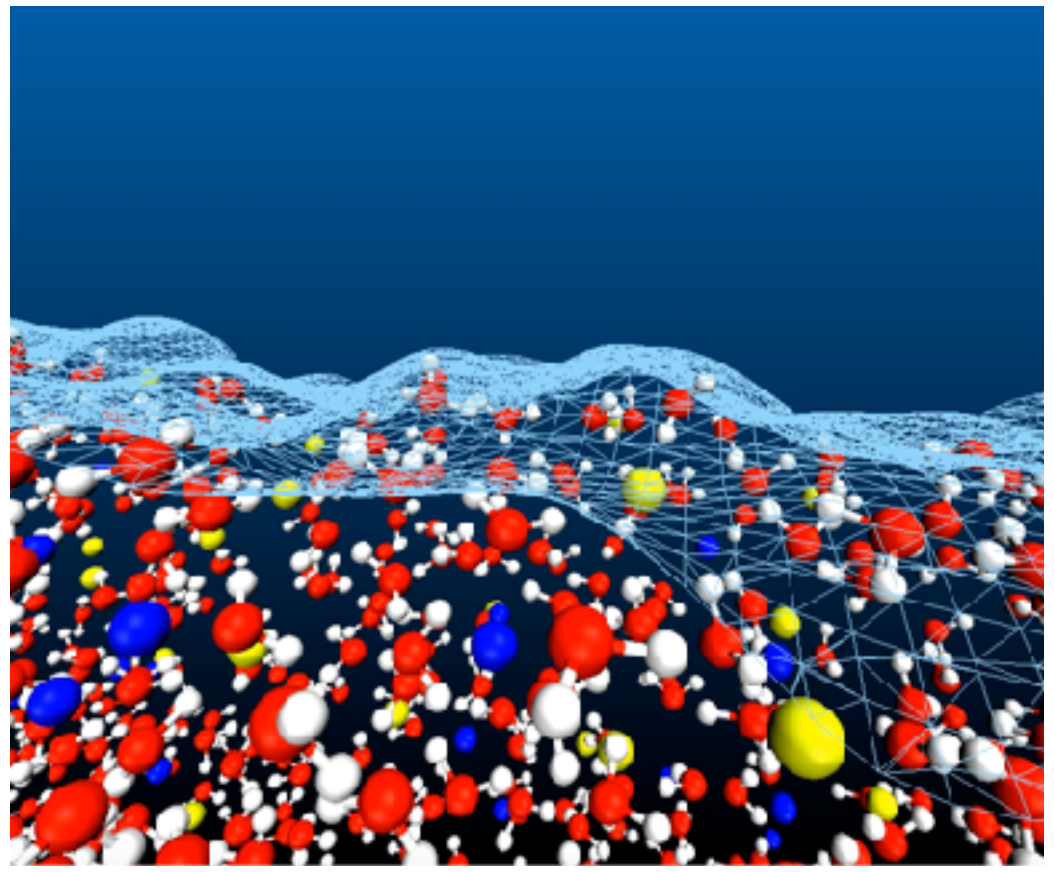
Ice

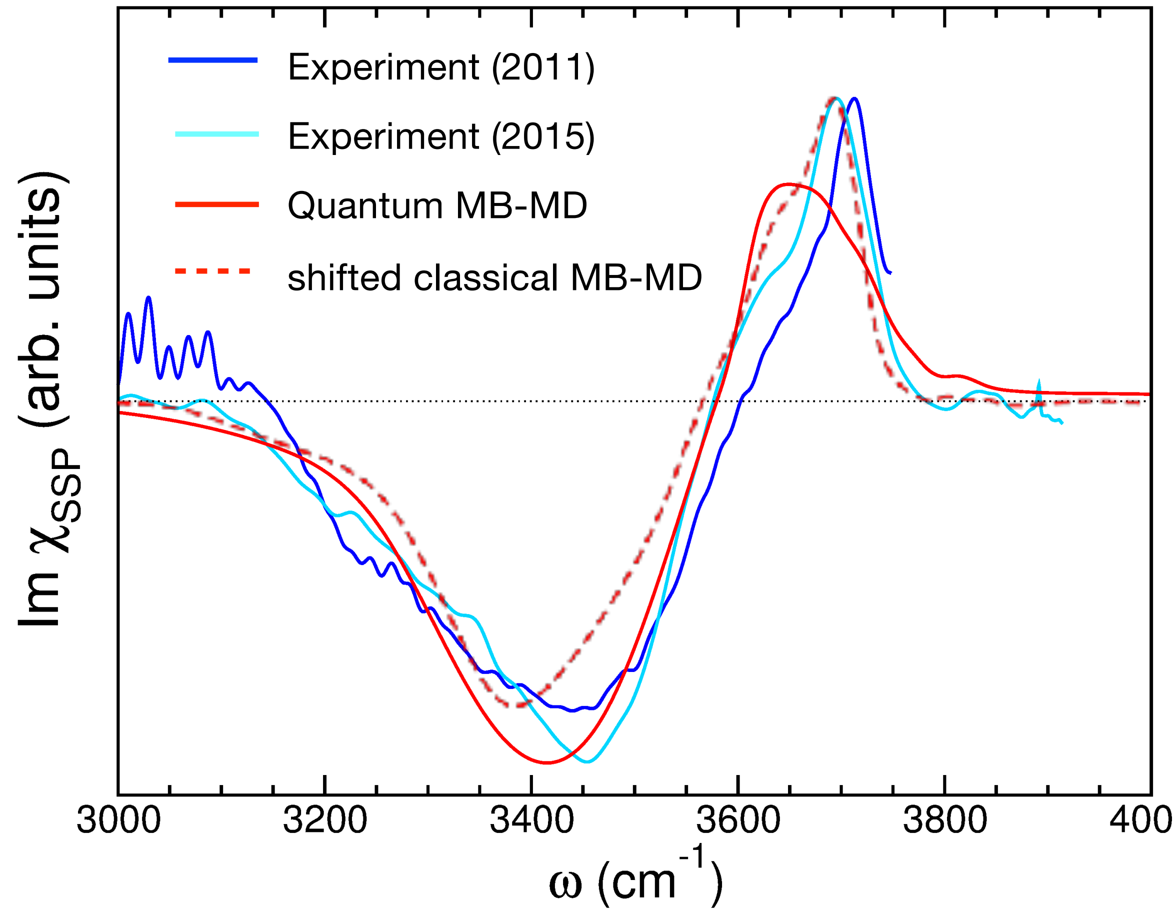




J. Chem. Theory Comput. 13, 1778 (2015)
J. Phys. Chem. Lett. 8, 2579 (2017)

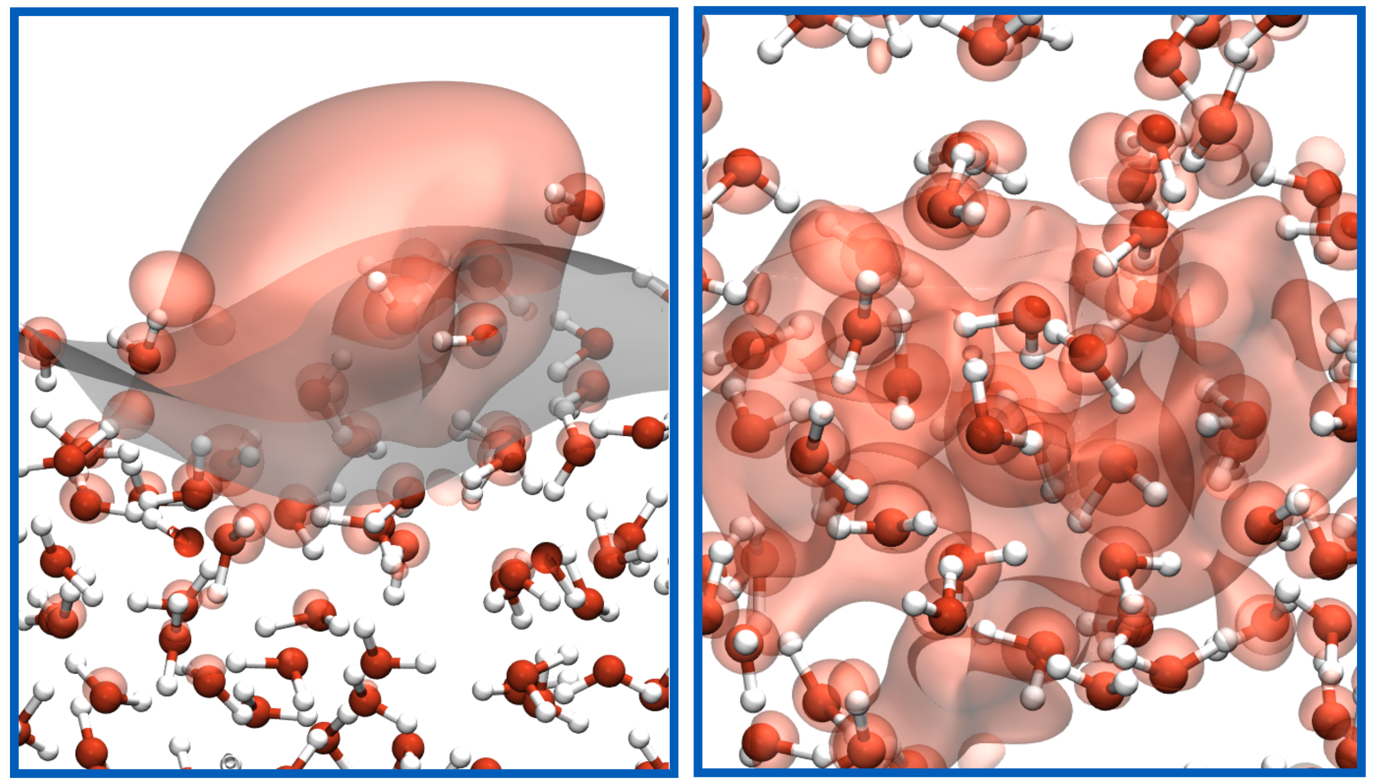
Air/water interface

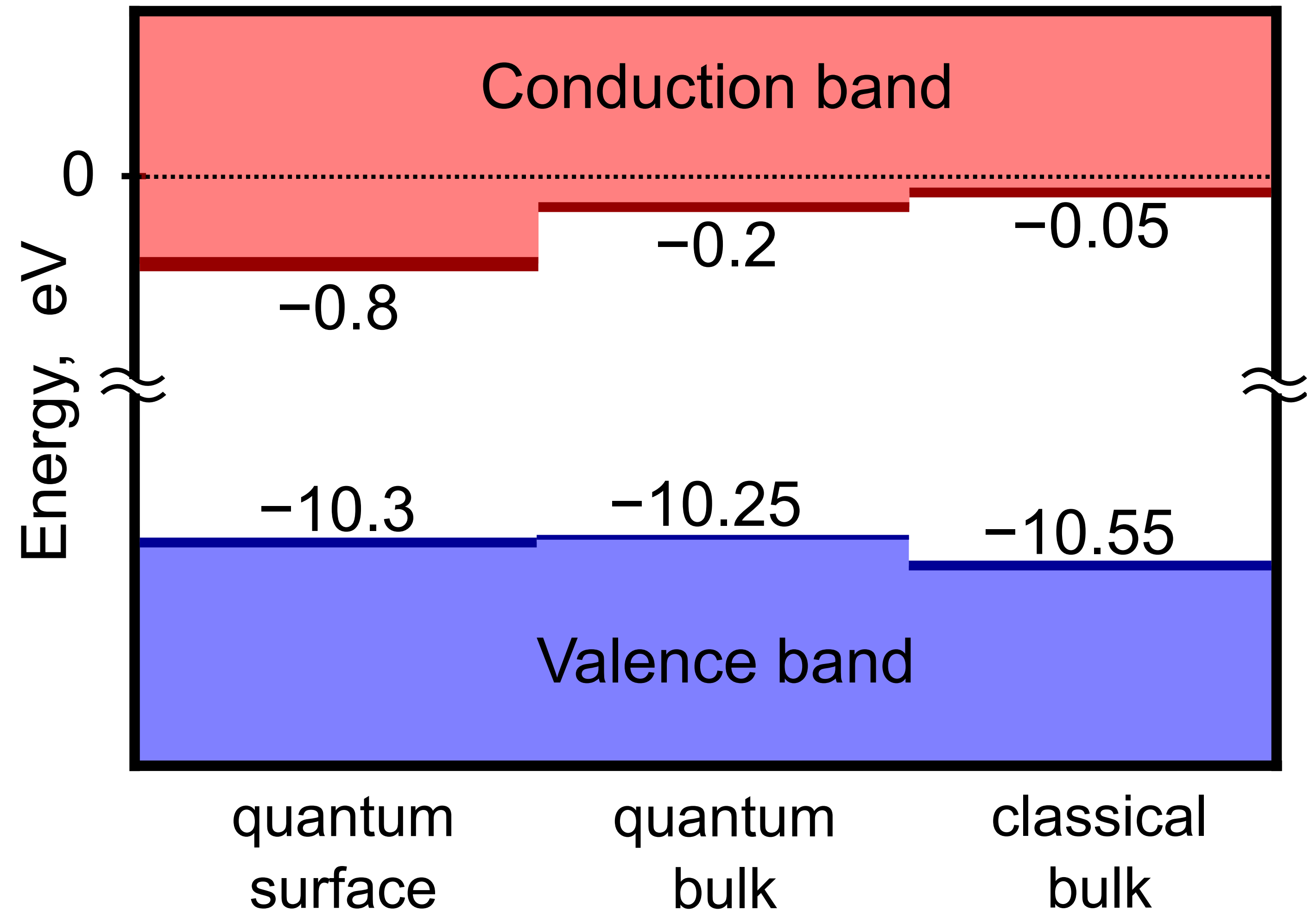




J. Am. Chem. Soc. 138, 3912 (2016)
J. Phys. Chem. B 122, 4356 (2018)

Electron affinity



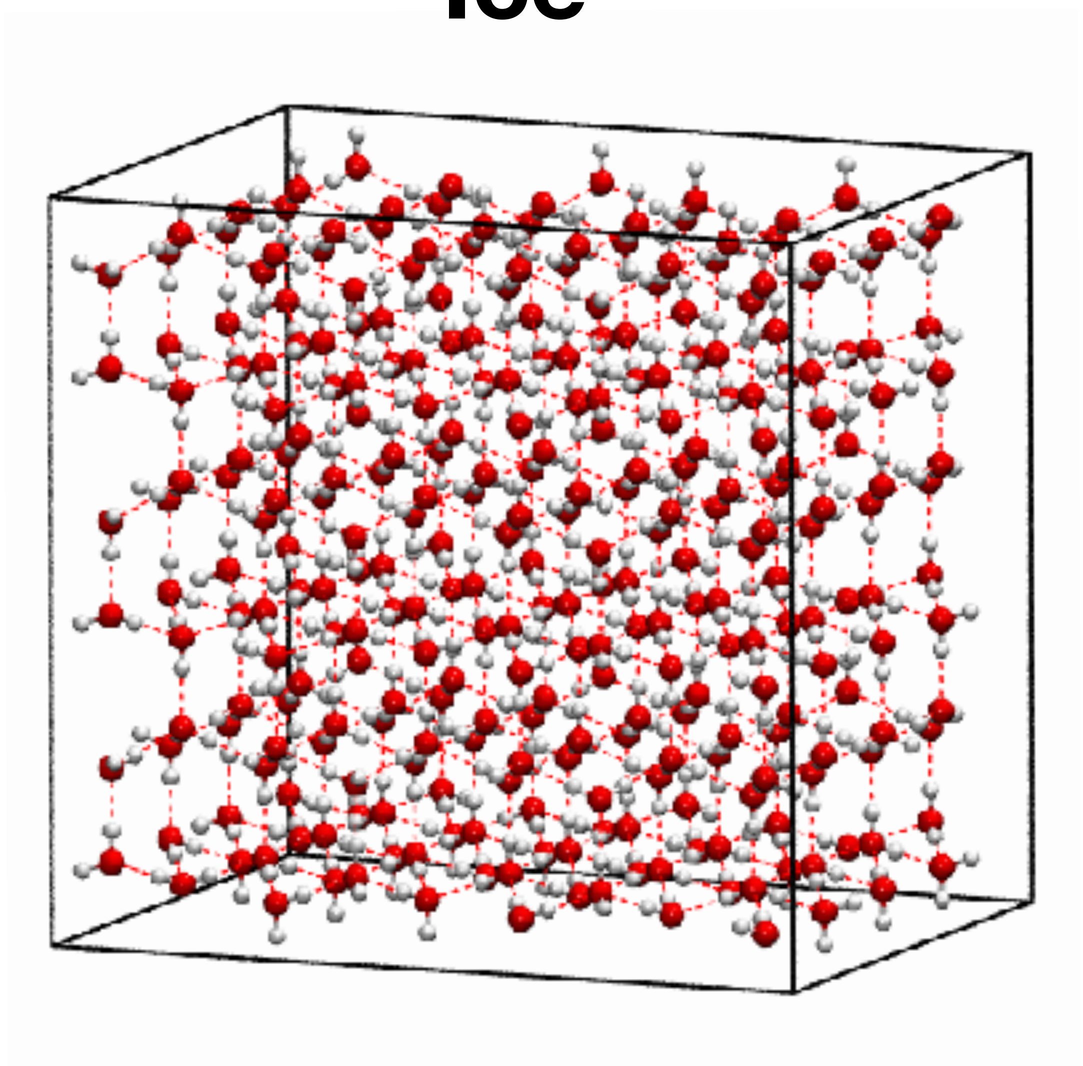


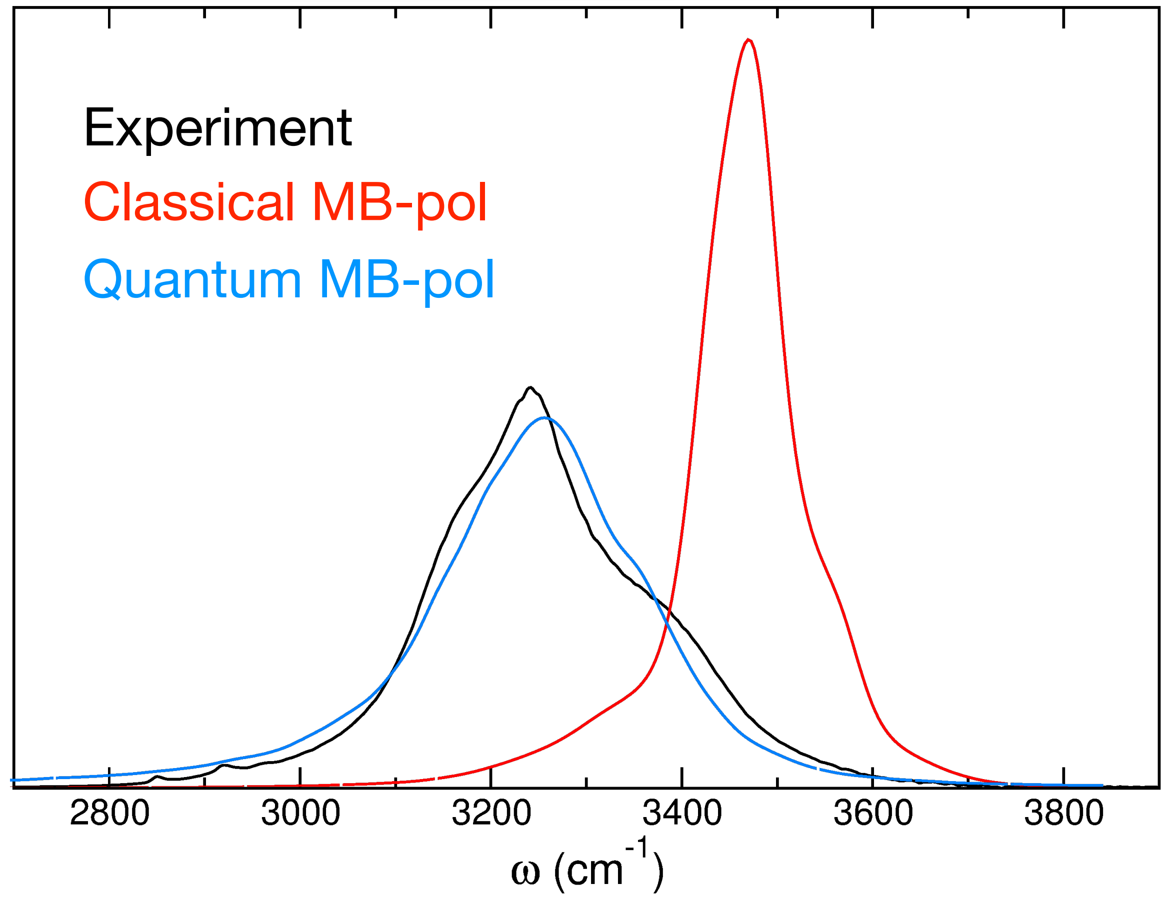
Collaboration with Giulia Galli (Chicago)
Nat. Commun. 9, 247 (2018)

FP, Acc. Chem. Res. 49, 1844 (2016)

MB-pol: Chemical Accuracy Across All Phases

Ice

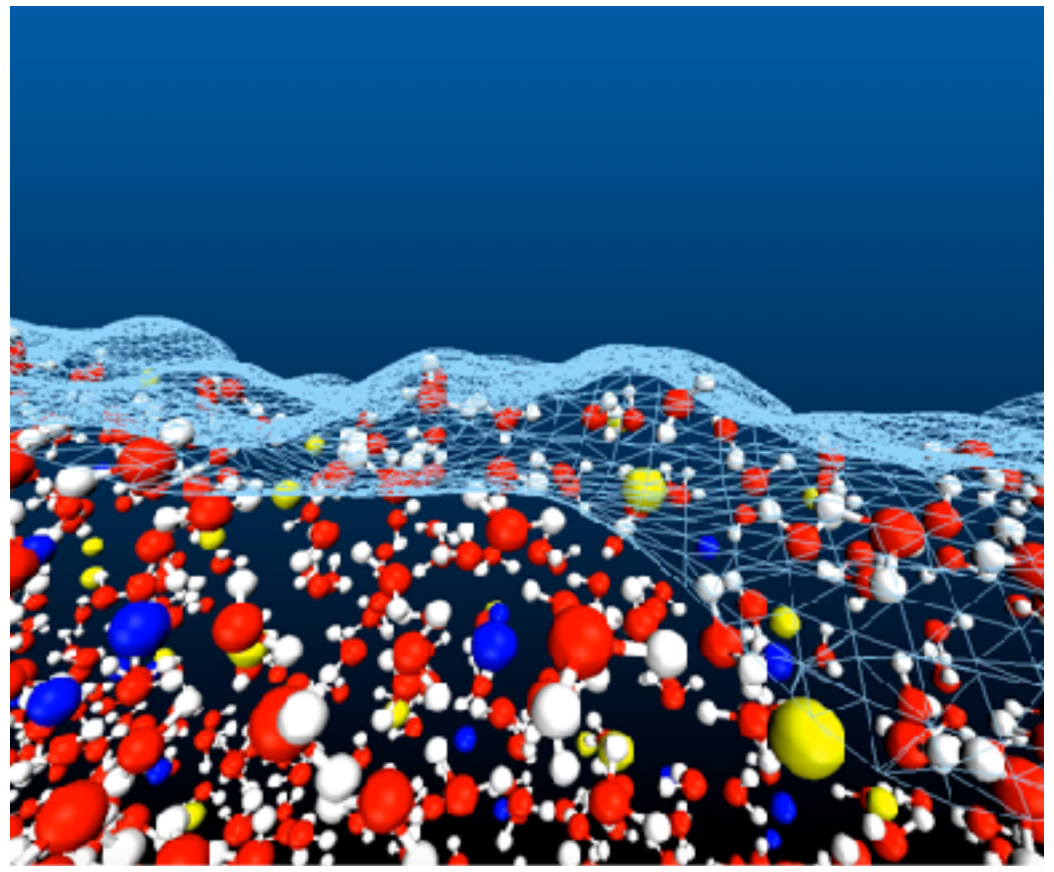


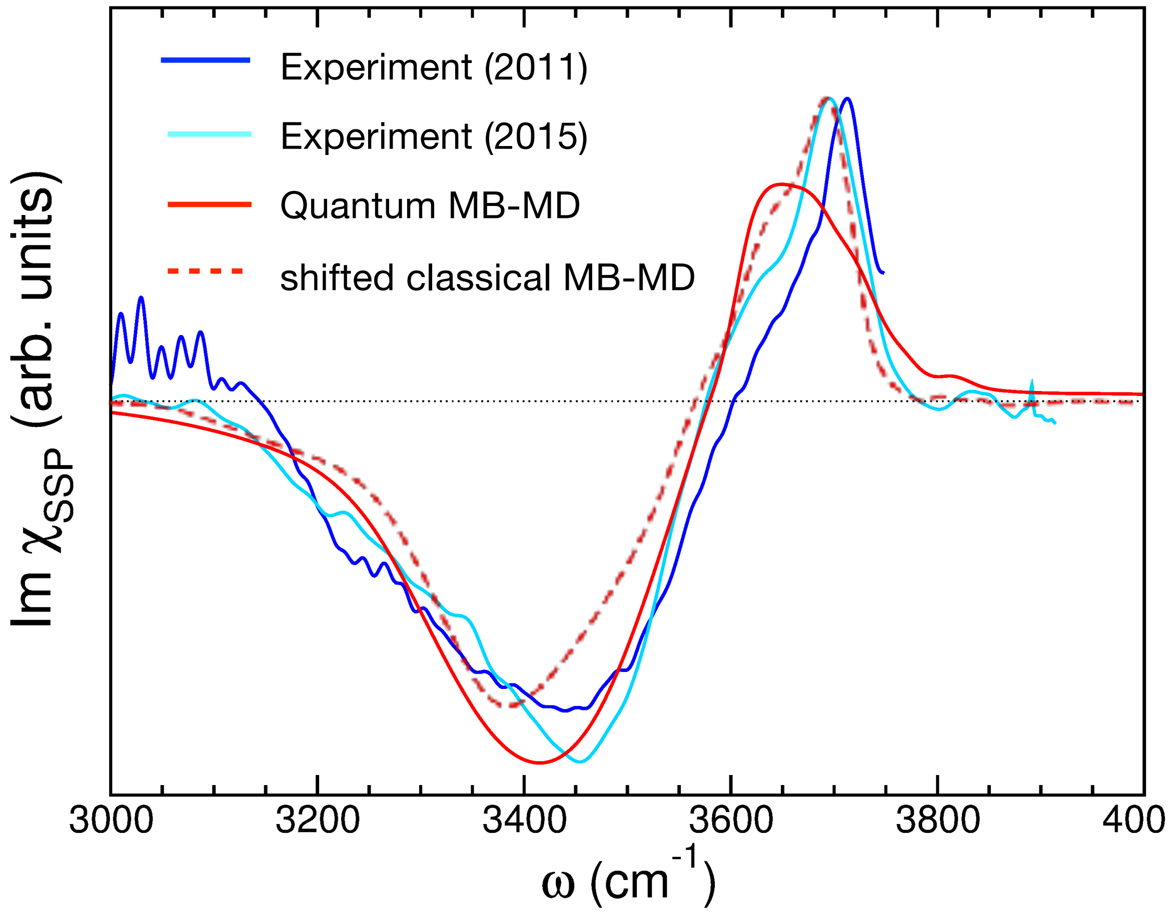


J. Chem. Theory Comput. *13*, *1778 (2015)*

J. Phys. Chem. Lett. *8*, *2579 (2017)*

Air/water interface

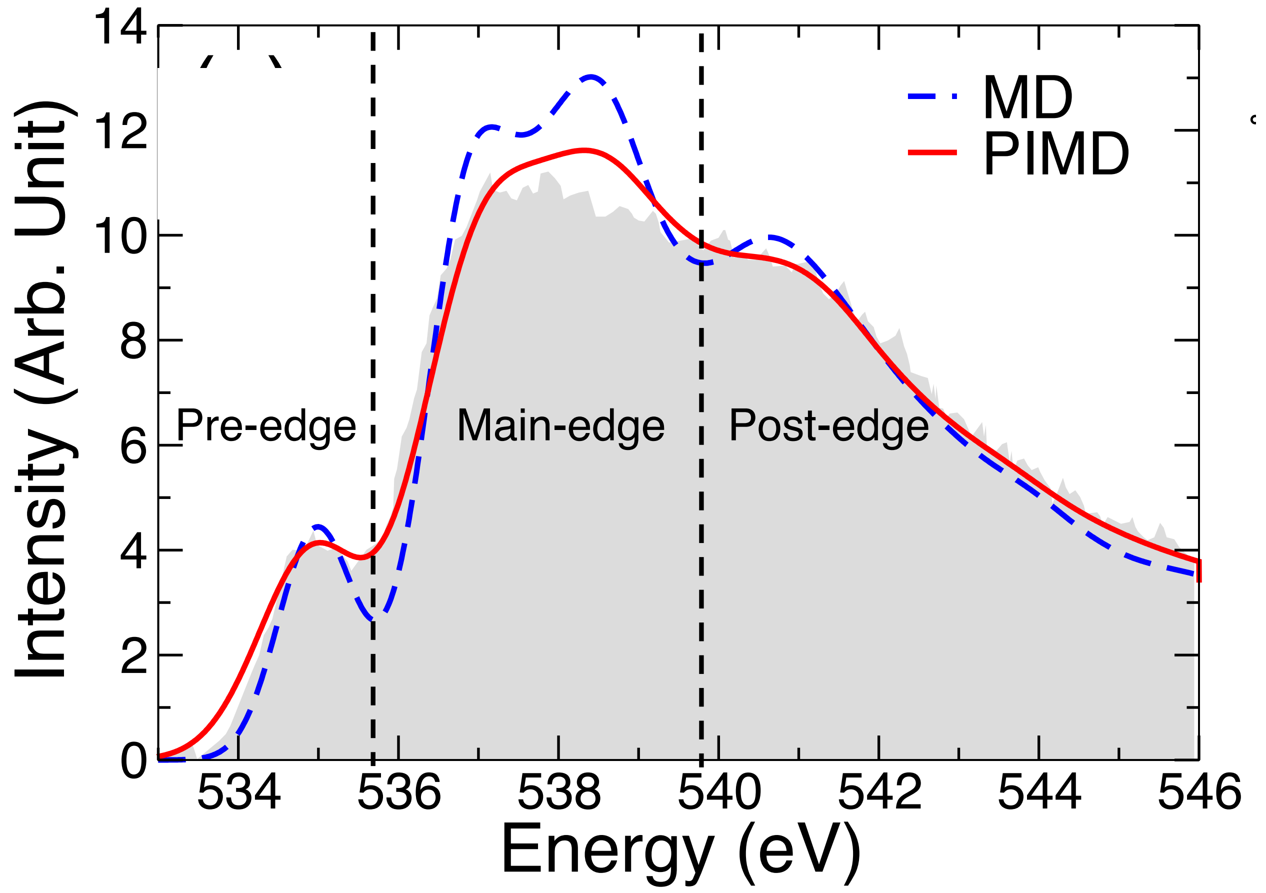


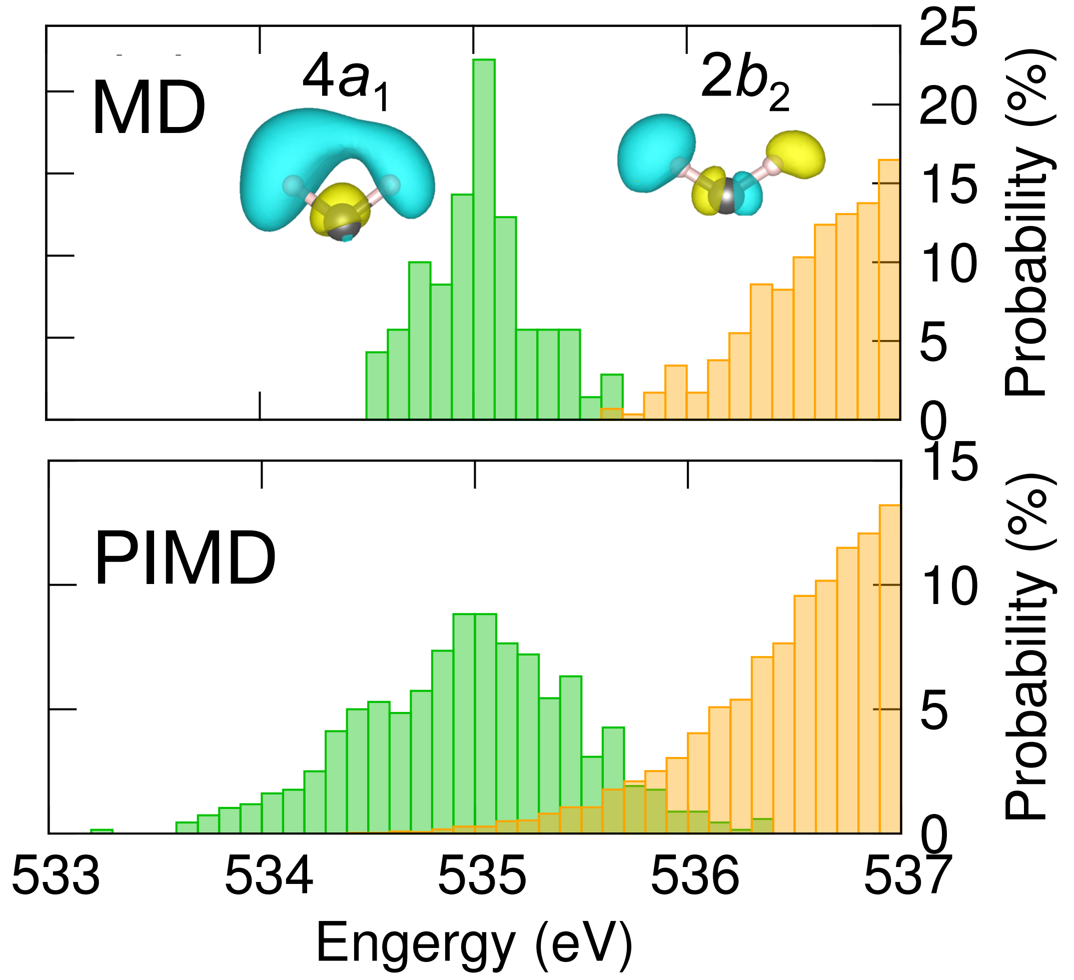


J. Am. Chem. Soc. *138*, *3912 (2016)*

J. Phys. Chem. B *122*, *4356 (2018)*

X-ray absorption





Collaboration with Xifan Wu (Temple)

Phys. Rev. Lett. *121*, *137401 (2018)*

FP, Acc. Chem. Res. *49*, *1844 (2016)*

MB-pol in Perspective

Reference data from monomer, dimer, hexamer, and ice

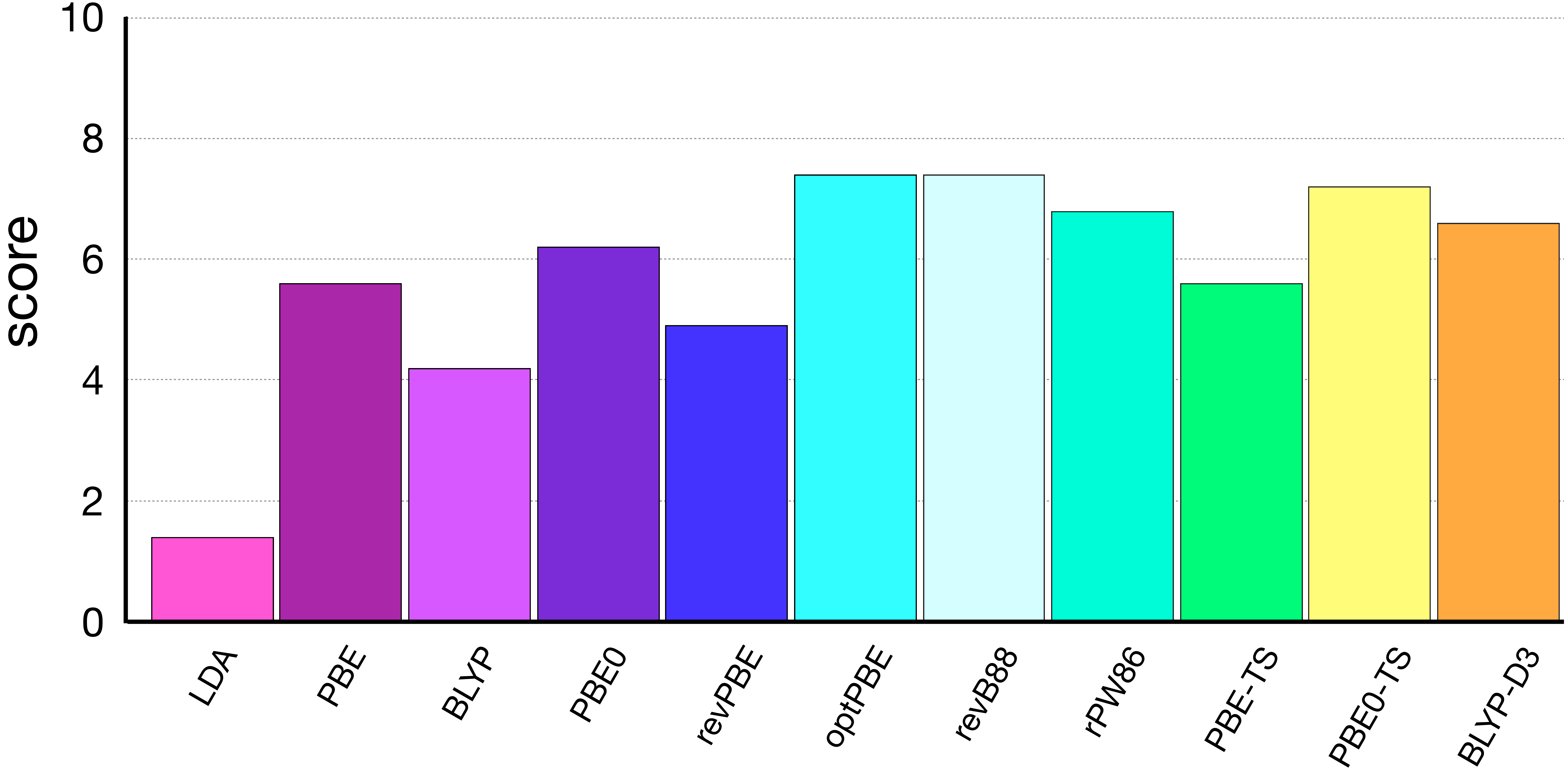
Gillan, Alfé, Michaelides, "Perspective: How Good Is DFT for Water?", J. Chem. Phys. 144, 130901 (2016)

Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

Reference data from monomer, dimer, hexamer, and ice

Gillan, Alfé, Michaelides, "Perspective: How Good Is DFT for Water?", J. Chem. Phys. 144, 130901 (2016)

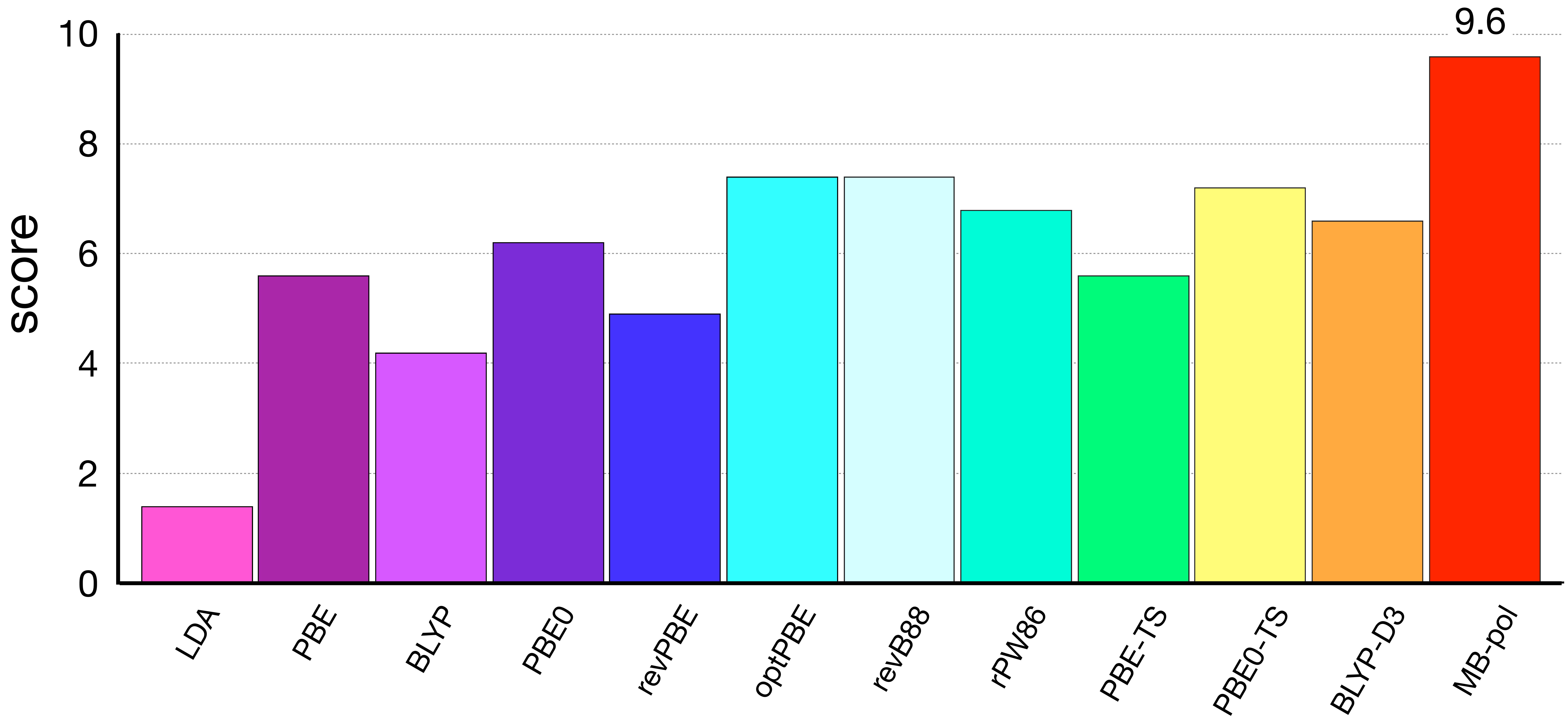


Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

Reference data from monomer, dimer, hexamer, and ice

Gillan, Alfé, Michaelides, "Perspective: How Good Is DFT for Water?", J. Chem. Phys. 144, 130901 (2016)

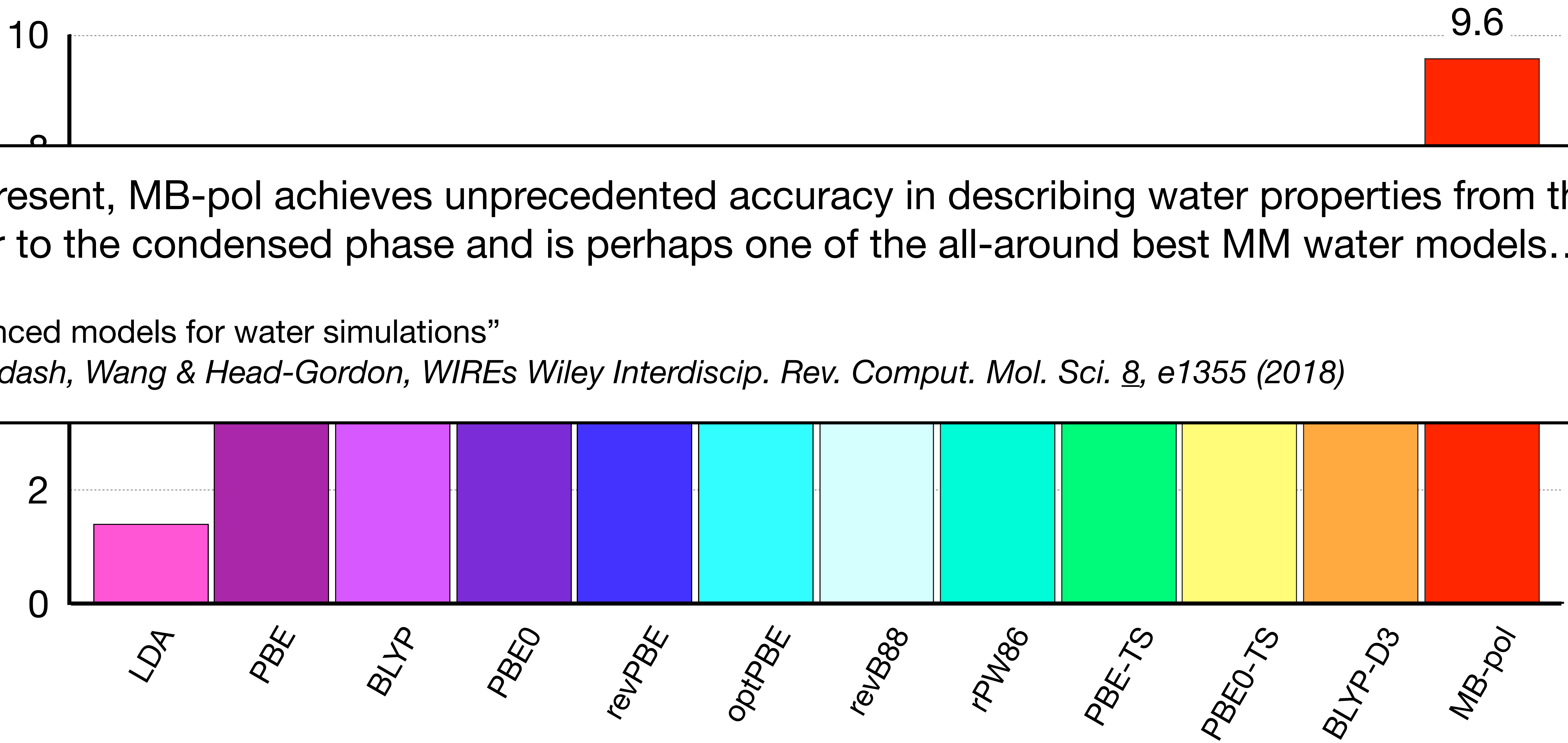


Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

Reference data from monomer, dimer, hexamer, and ice

Gillan, Alfé, Michaelides, “Perspective: How Good Is DFT for Water?”, J. Chem. Phys. 144, 130901 (2016)

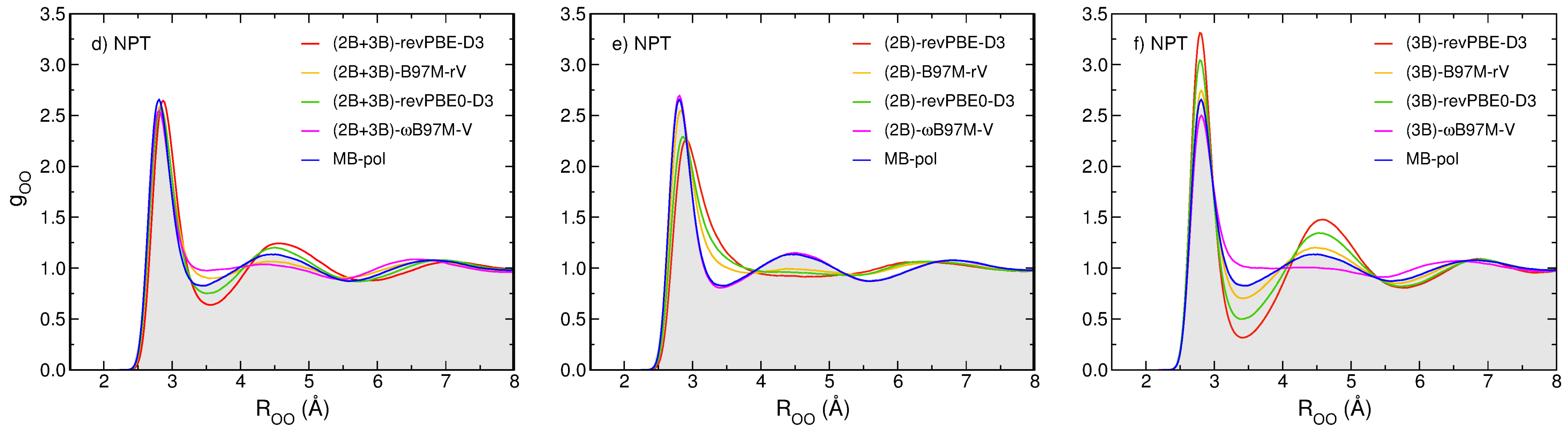


> “At present, MB-pol achieves unprecedented accuracy in describing water properties from the dimer to the condensed phase and is perhaps one of the all-around best MM water models…”
>
> “Advanced models for water simulations”
> *Demerdash, Wang & Head-Gordon, WIREs Wiley Interdiscip. Rev. Comput. Mol. Sci. 8, e1355 (2018)*

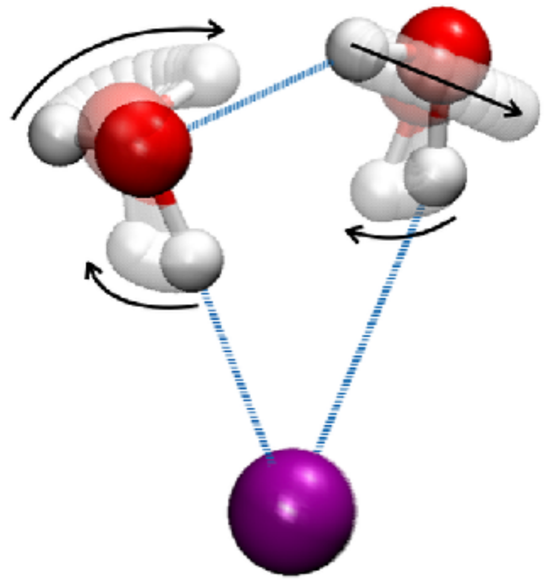
Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

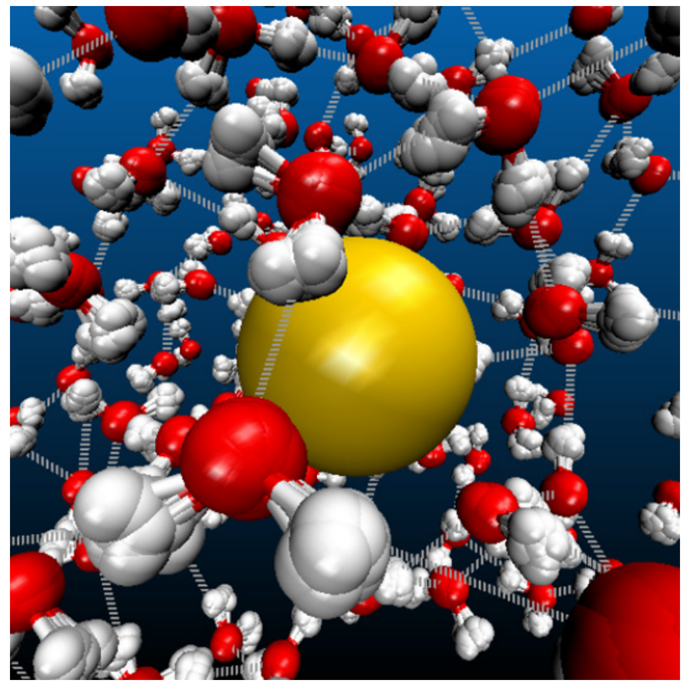
Learning Physics from Many-Body Models

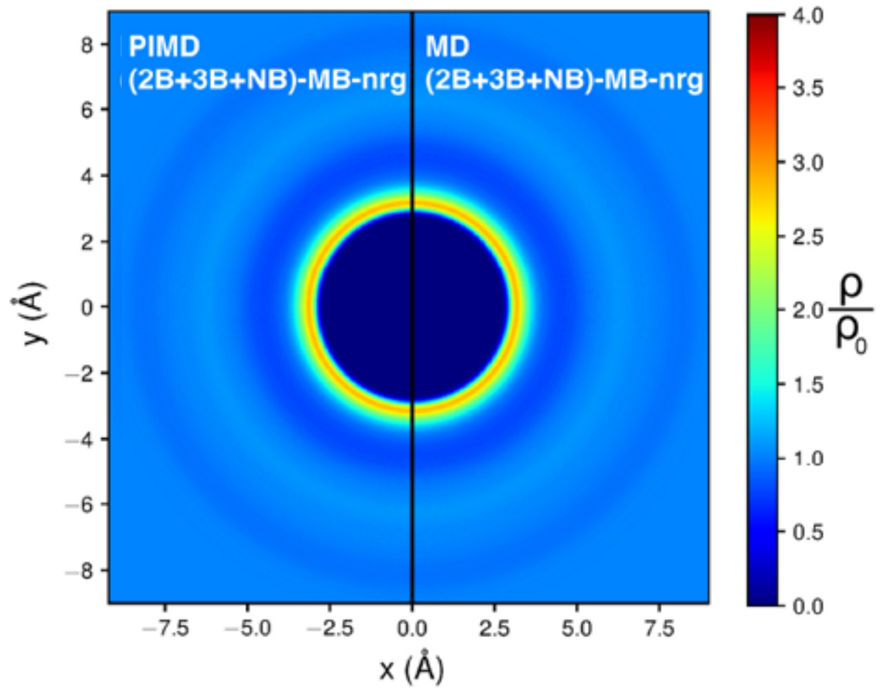
What makes water water?



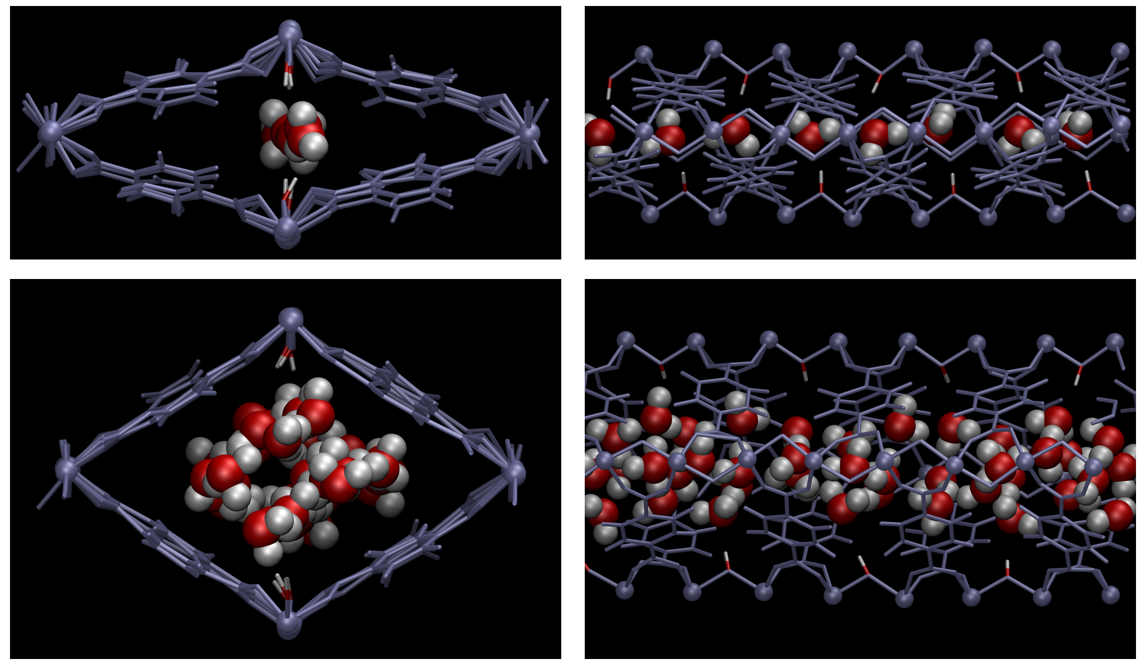
How good are simulations of hydration?

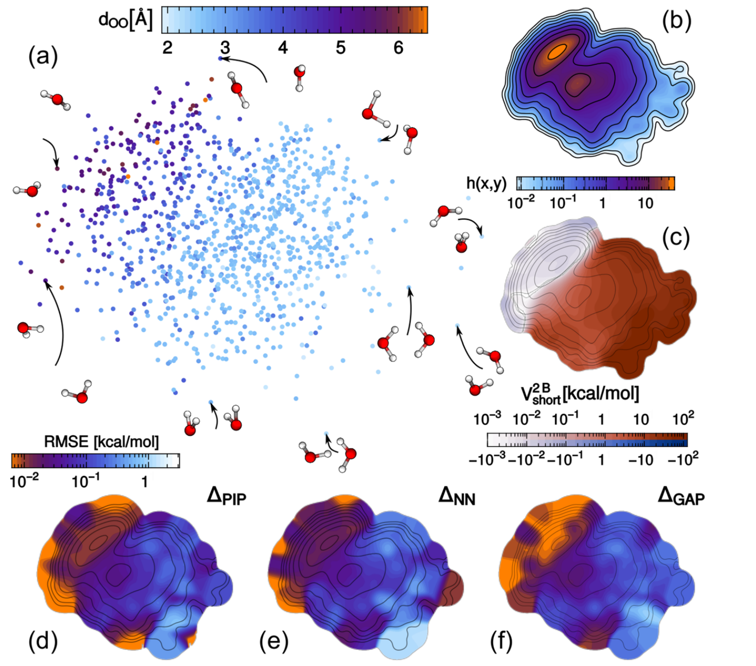






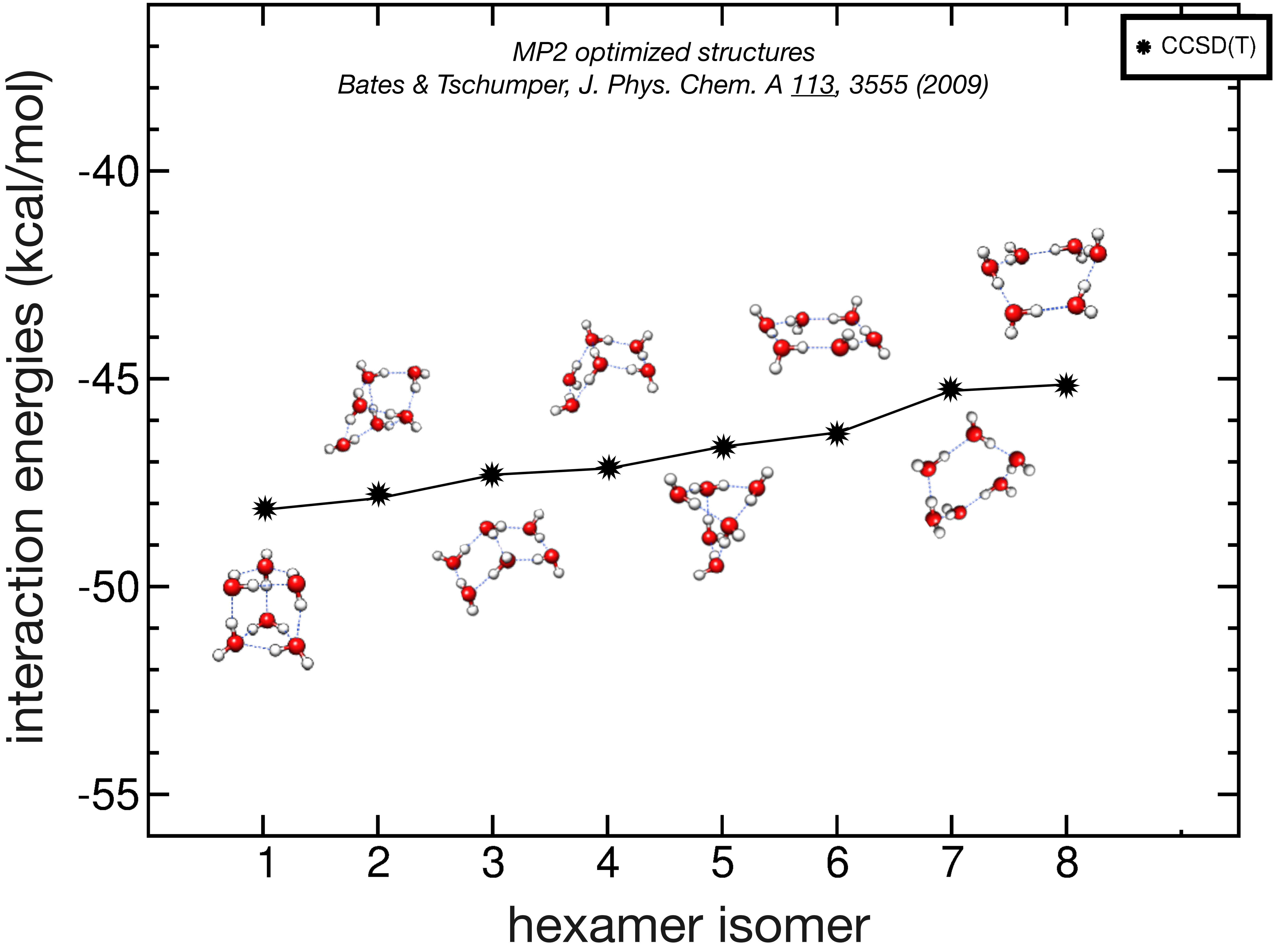
Data-driven many-body models: Quo vadis?





Why MB-pol Works

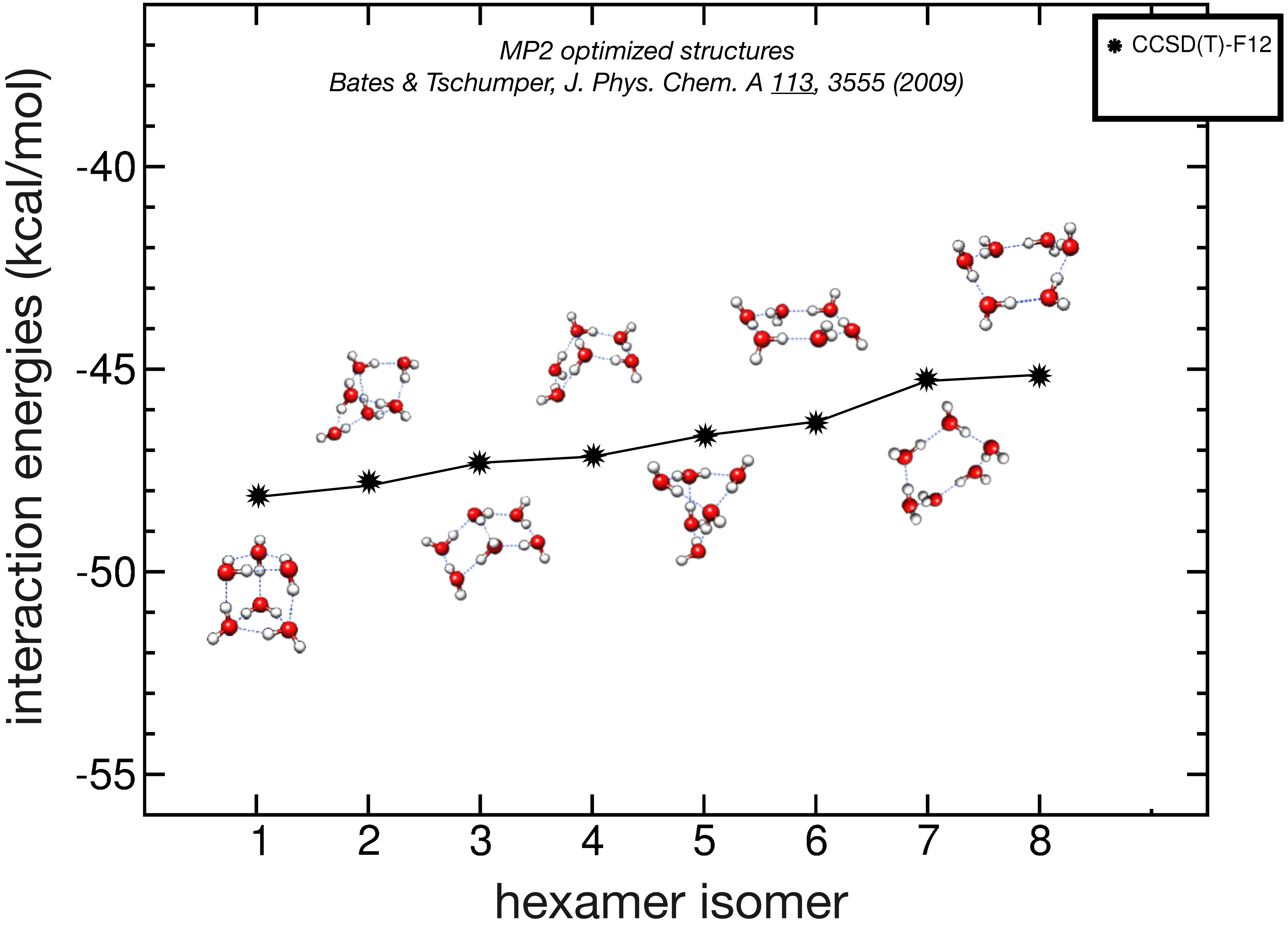
MB-pol vs. CCSD(T): Interaction energies for water clusters



FP, Acc. Chem. Res. 49, 1844 (2016)

Why MB-pol Works

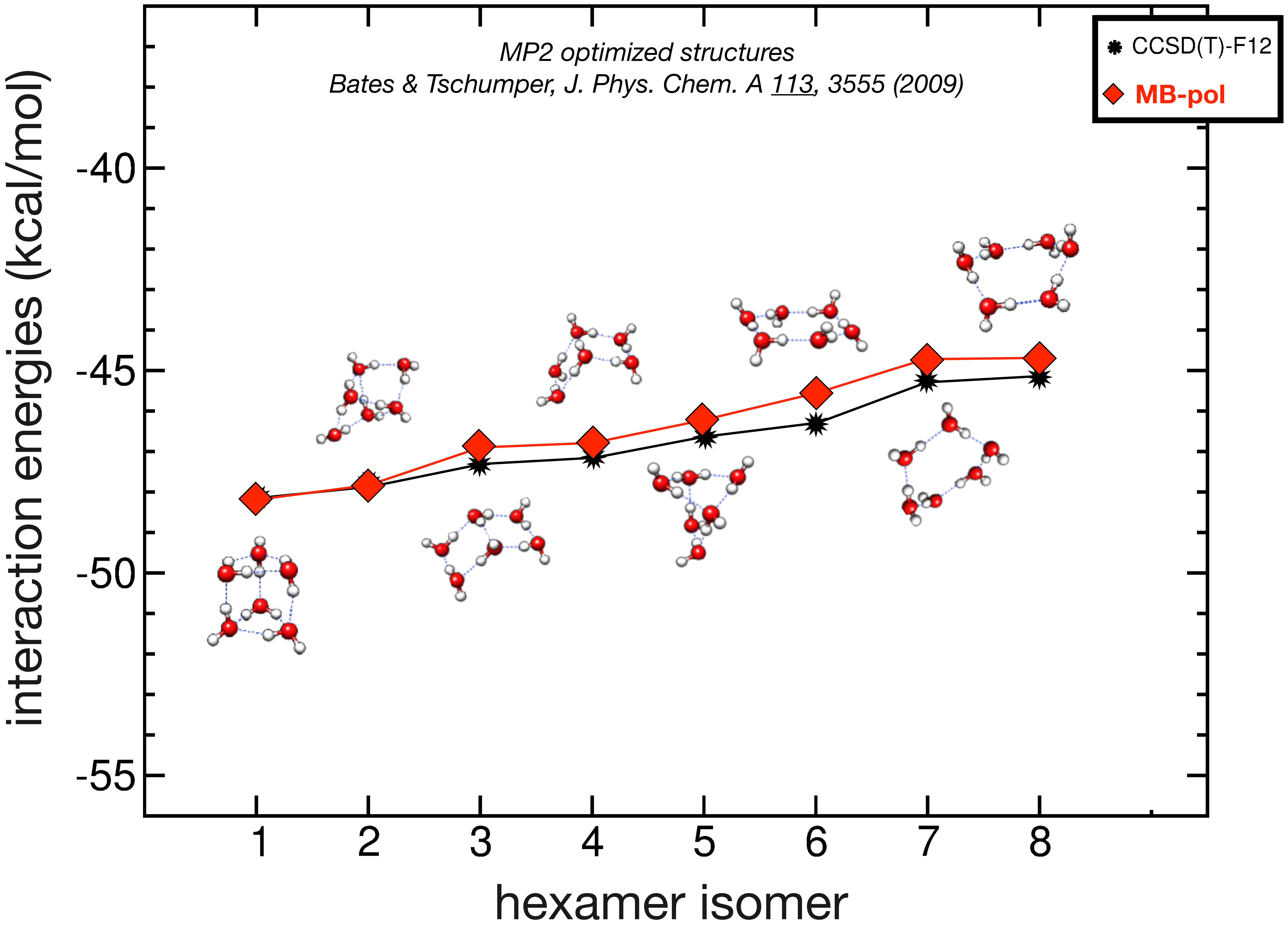
MB-pol vs. CCSD(T): Interaction energies for water clusters



FP, Acc. Chem. Res. 49, 1844 (2016)

Why MB-pol Works

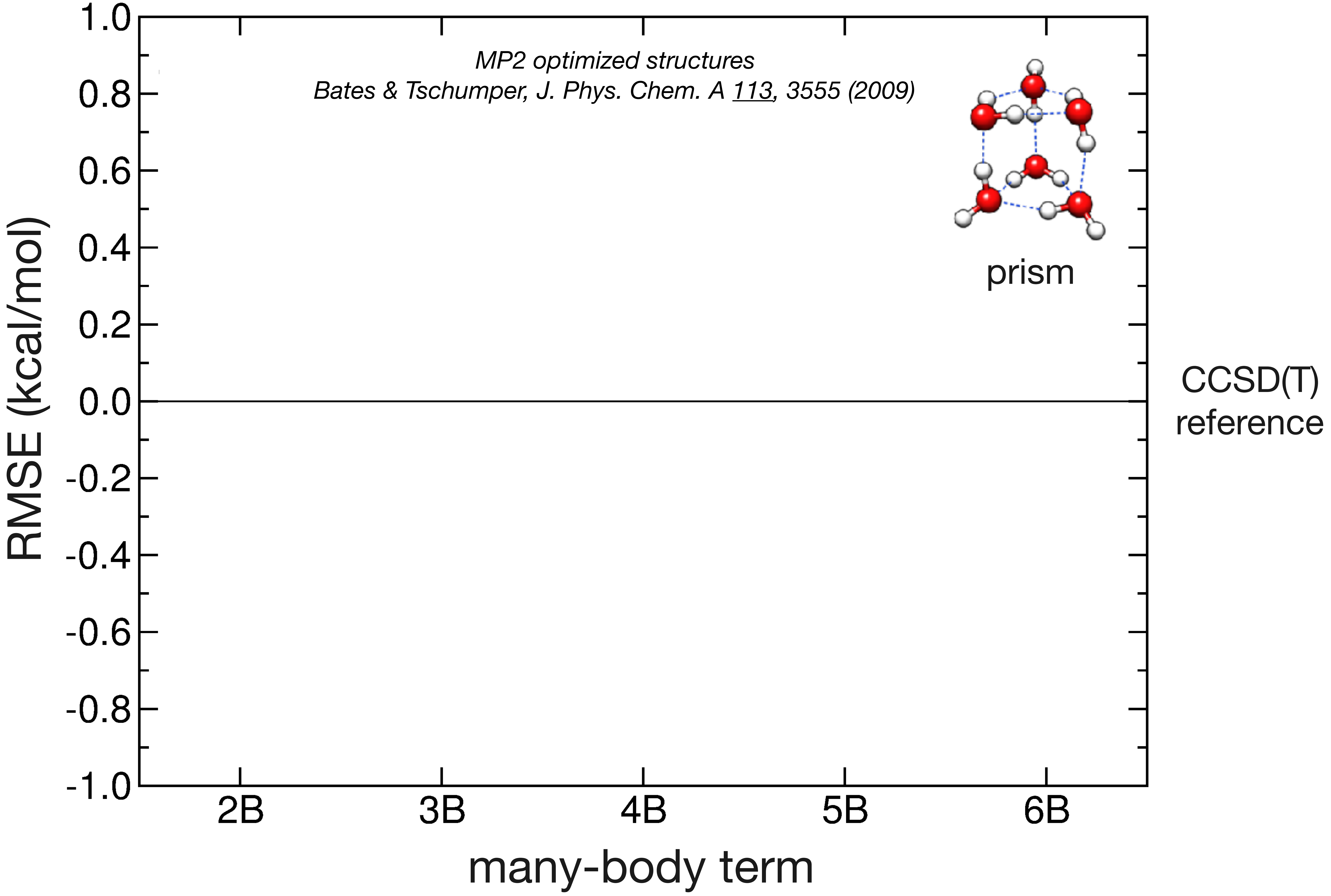
MB-pol vs. CCSD(T): Interaction energies for water clusters



FP, Acc. Chem. Res. 49, 1844 (2016)

Why MB-pol Works

MB-pol vs. CCSD(T): Many-body energy decomposition for water clusters

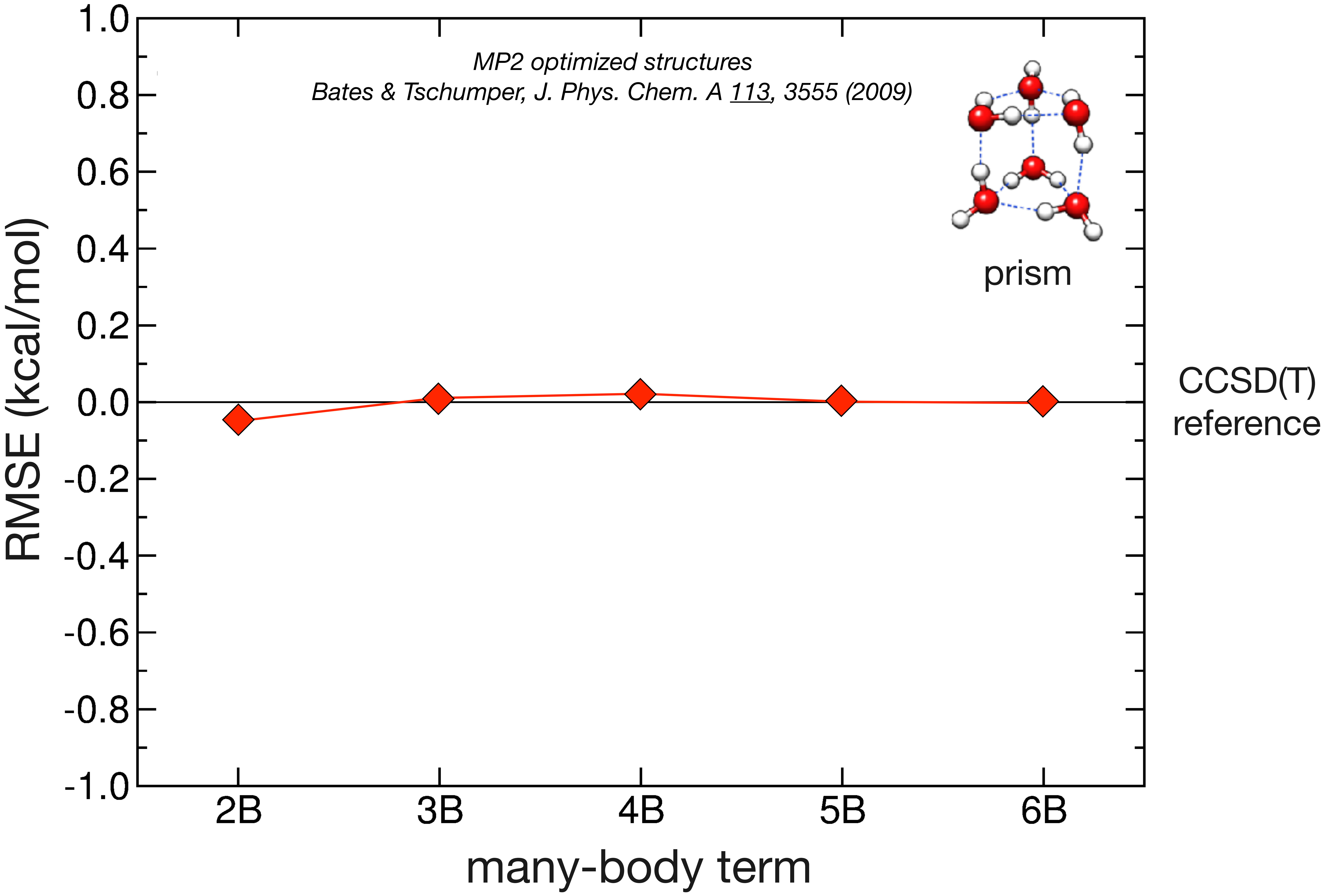


MP2 optimized structures
Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

FP, Acc. Chem. Res. 49, 1844 (2016)

Why MB-pol Works

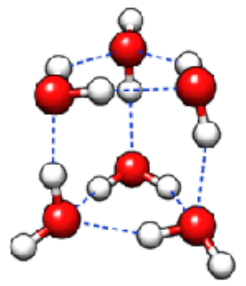
MB-pol vs. CCSD(T): Many-body energy decomposition for water clusters



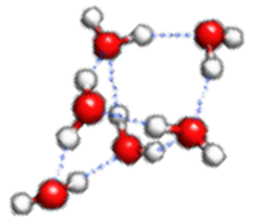
FP, Acc. Chem. Res. 49, 1844 (2016)

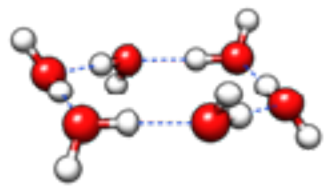
MB-pol in Perspective

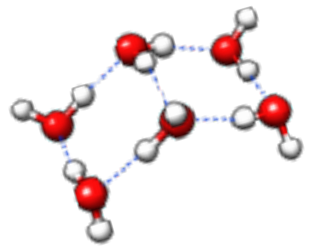
MB-pol vs. DFT: Interaction energies for water clusters

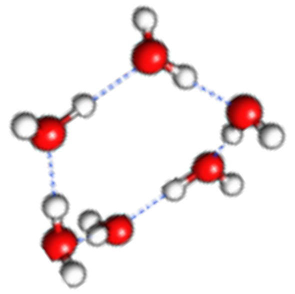

isomer 1

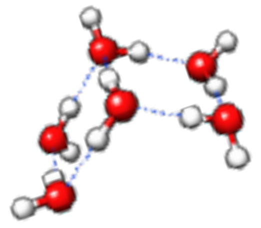

isomer 5

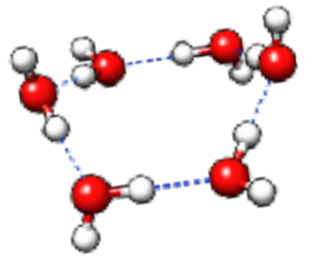

isomer 2

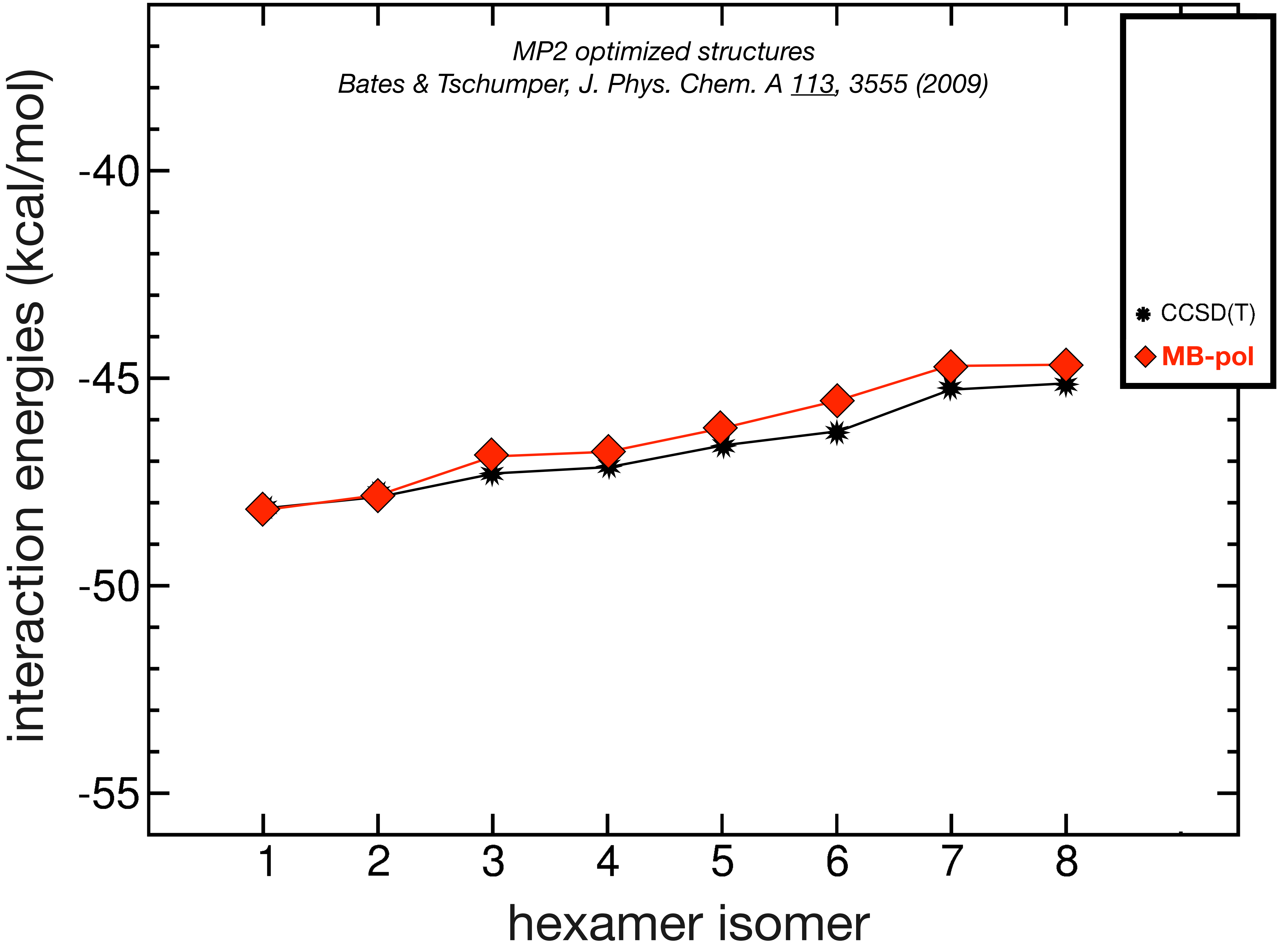

isomer 6


isomer 3


isomer 7


isomer 4

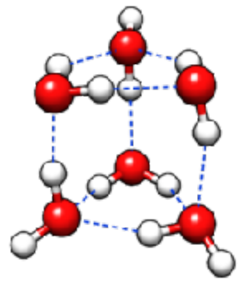

isomer 8

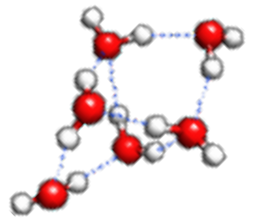

MP2 optimized structures
Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

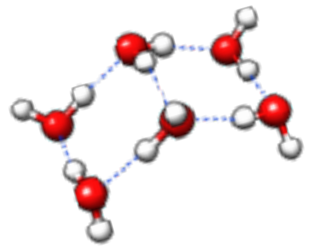
Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

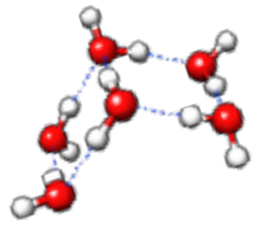
MB-pol in Perspective

MB-pol vs. DFT: Interaction energies for water clusters

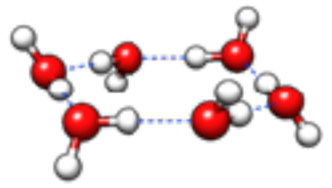

isomer 1

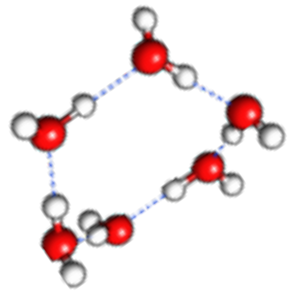

isomer 2

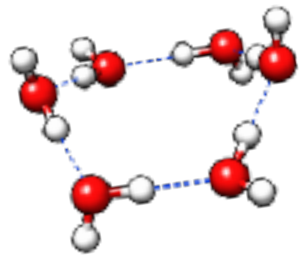

isomer 3

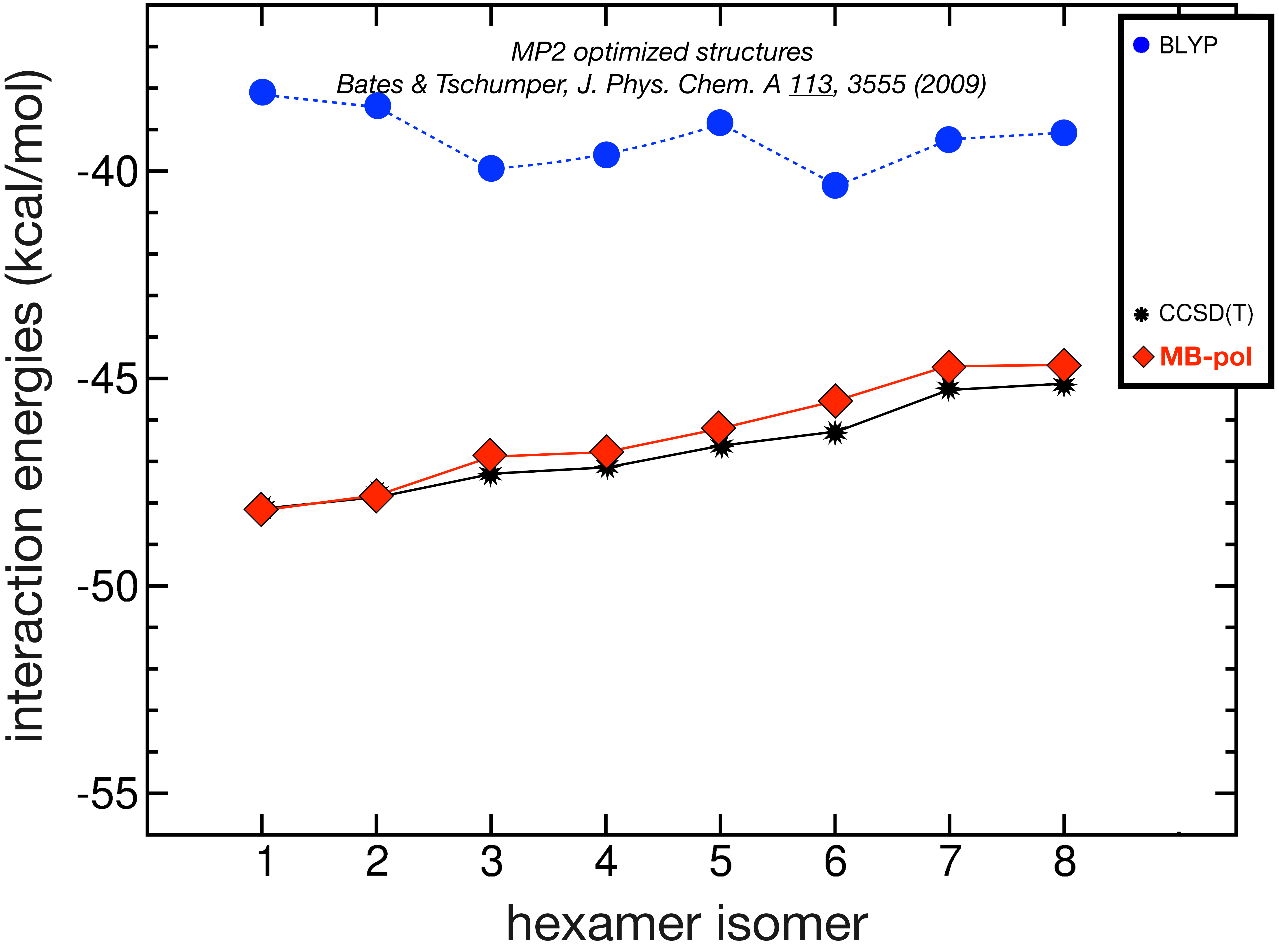

isomer 4


isomer 5


isomer 6


isomer 7


isomer 8

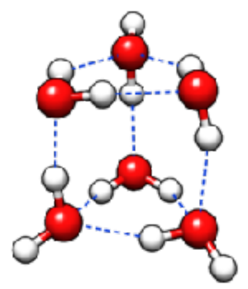


MP2 optimized structures
Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

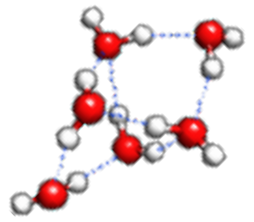
Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

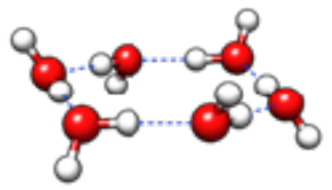
MB-pol in Perspective

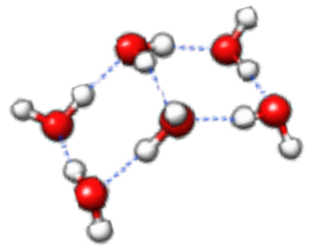
MB-pol vs. DFT: Interaction energies for water clusters

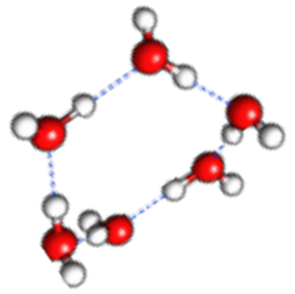

isomer 1

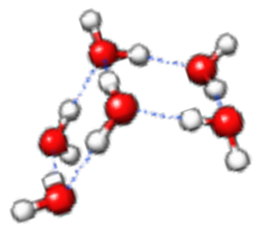

isomer 5

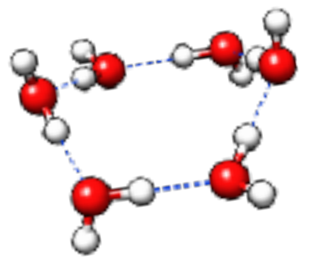

isomer 2

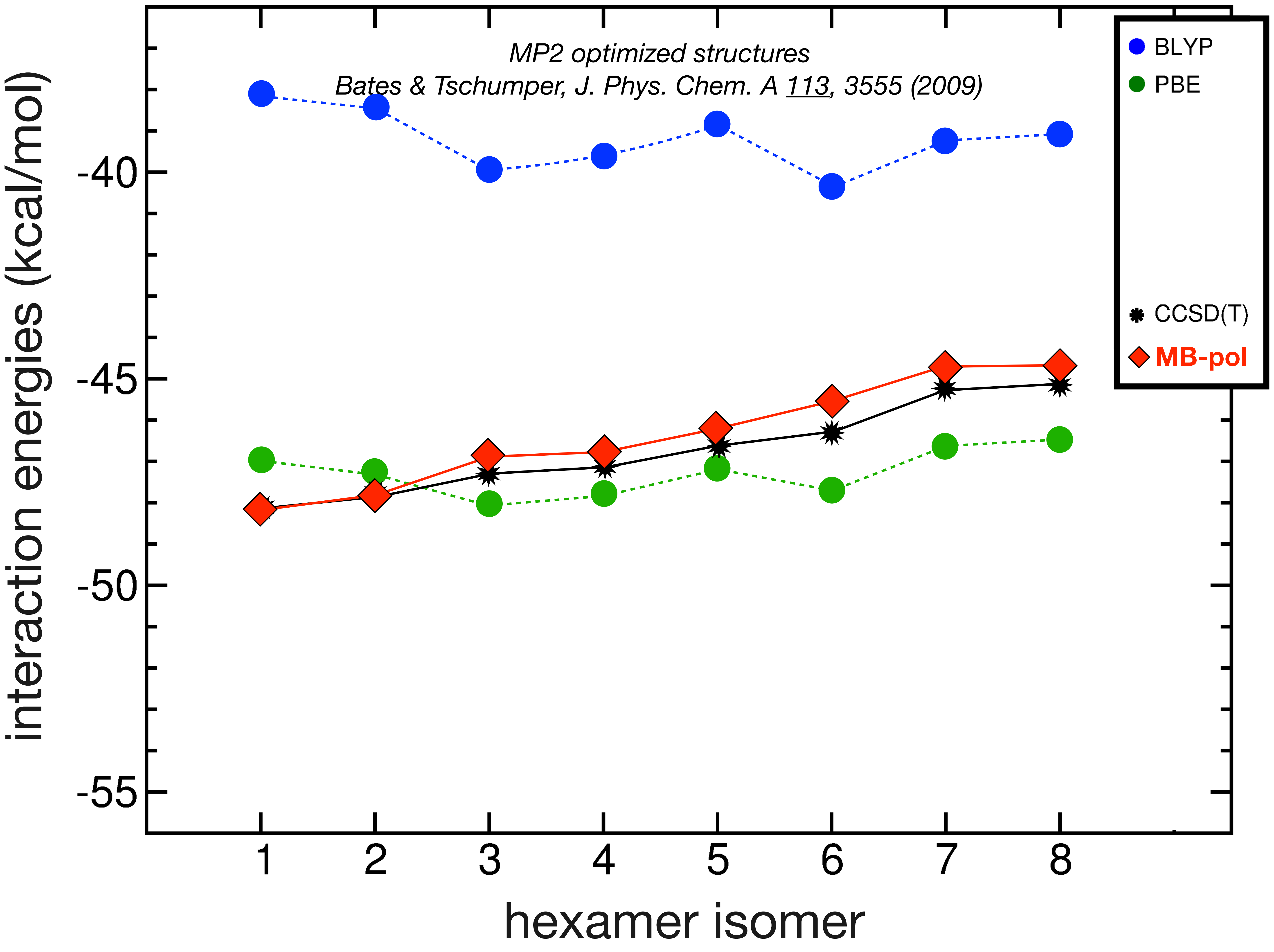

isomer 6


isomer 3


isomer 7


isomer 4


isomer 8

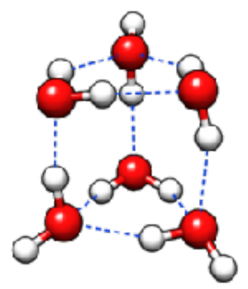

MP2 optimized structures

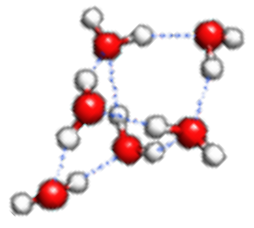
Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

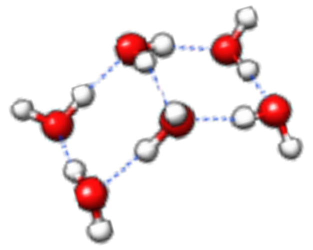
Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

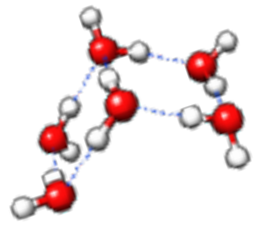
MB-pol in Perspective

MB-pol vs. DFT: Interaction energies for water clusters

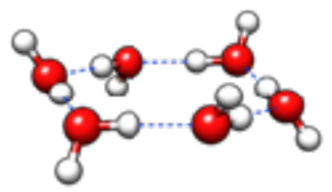

isomer 1

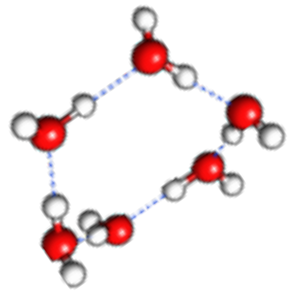

isomer 2

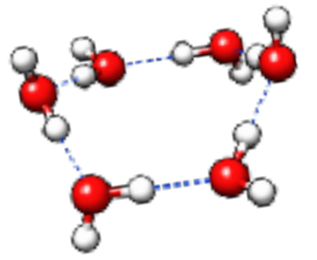

isomer 3

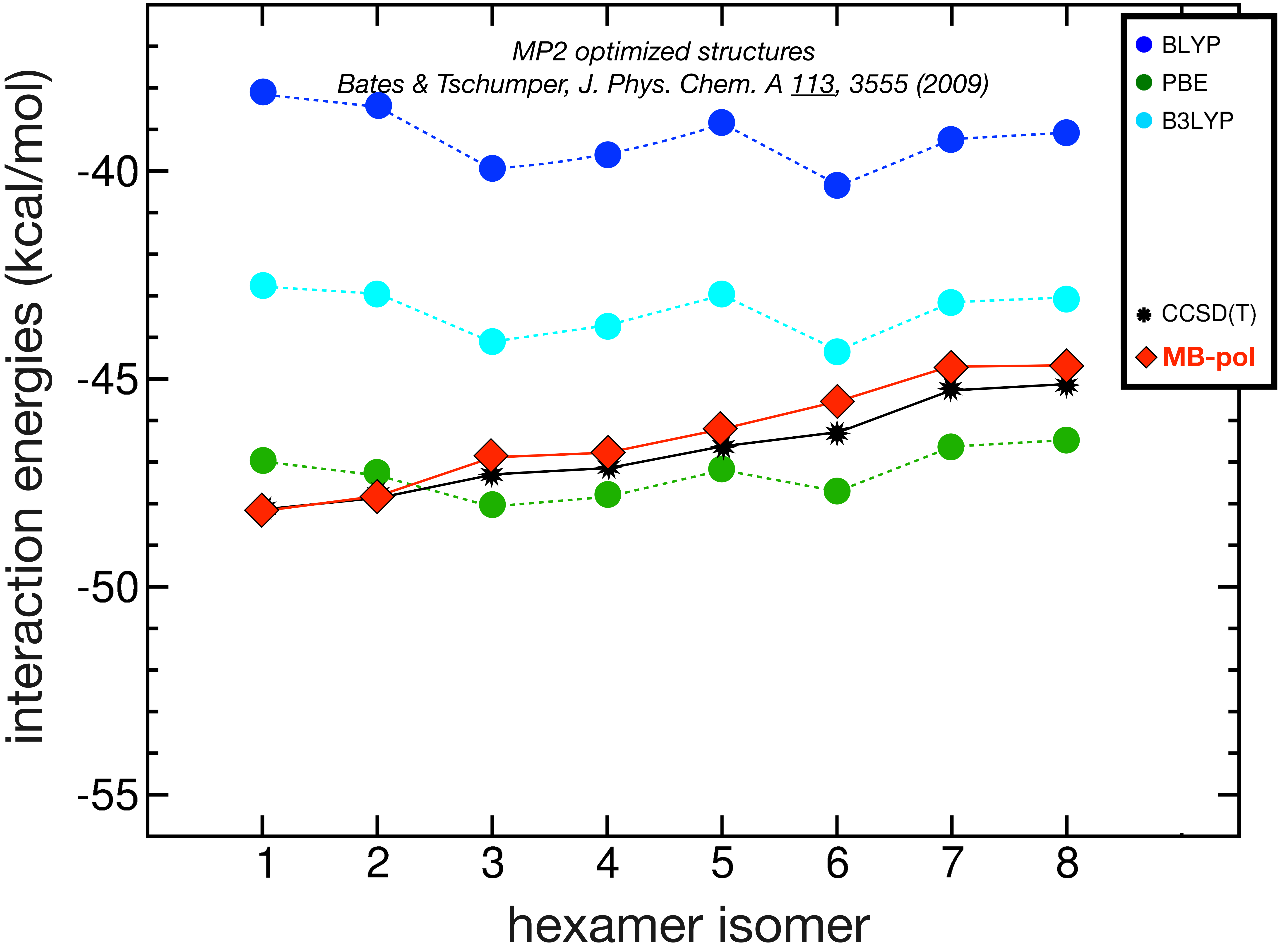

isomer 4


isomer 5


isomer 6


isomer 7


isomer 8

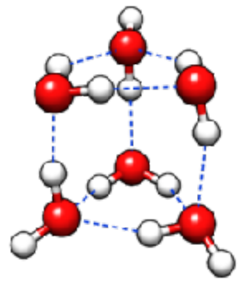


MP2 optimized structures
Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

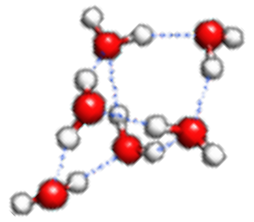
Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

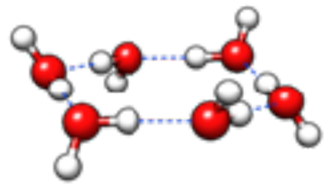
MB-pol in Perspective

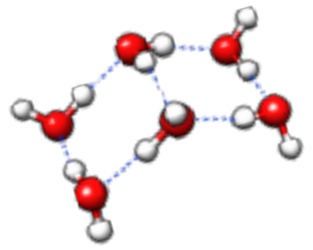
MB-pol vs. DFT: Interaction energies for water clusters

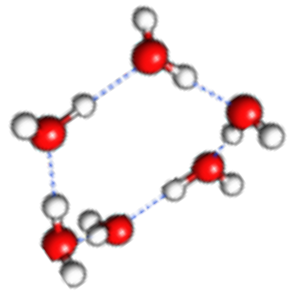

isomer 1

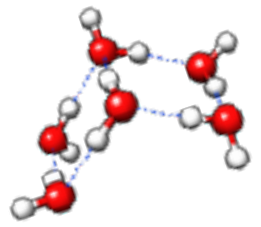

isomer 5

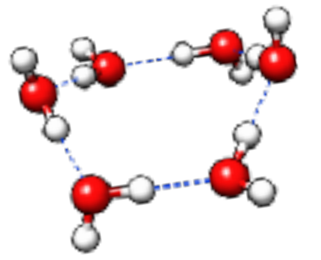

isomer 2

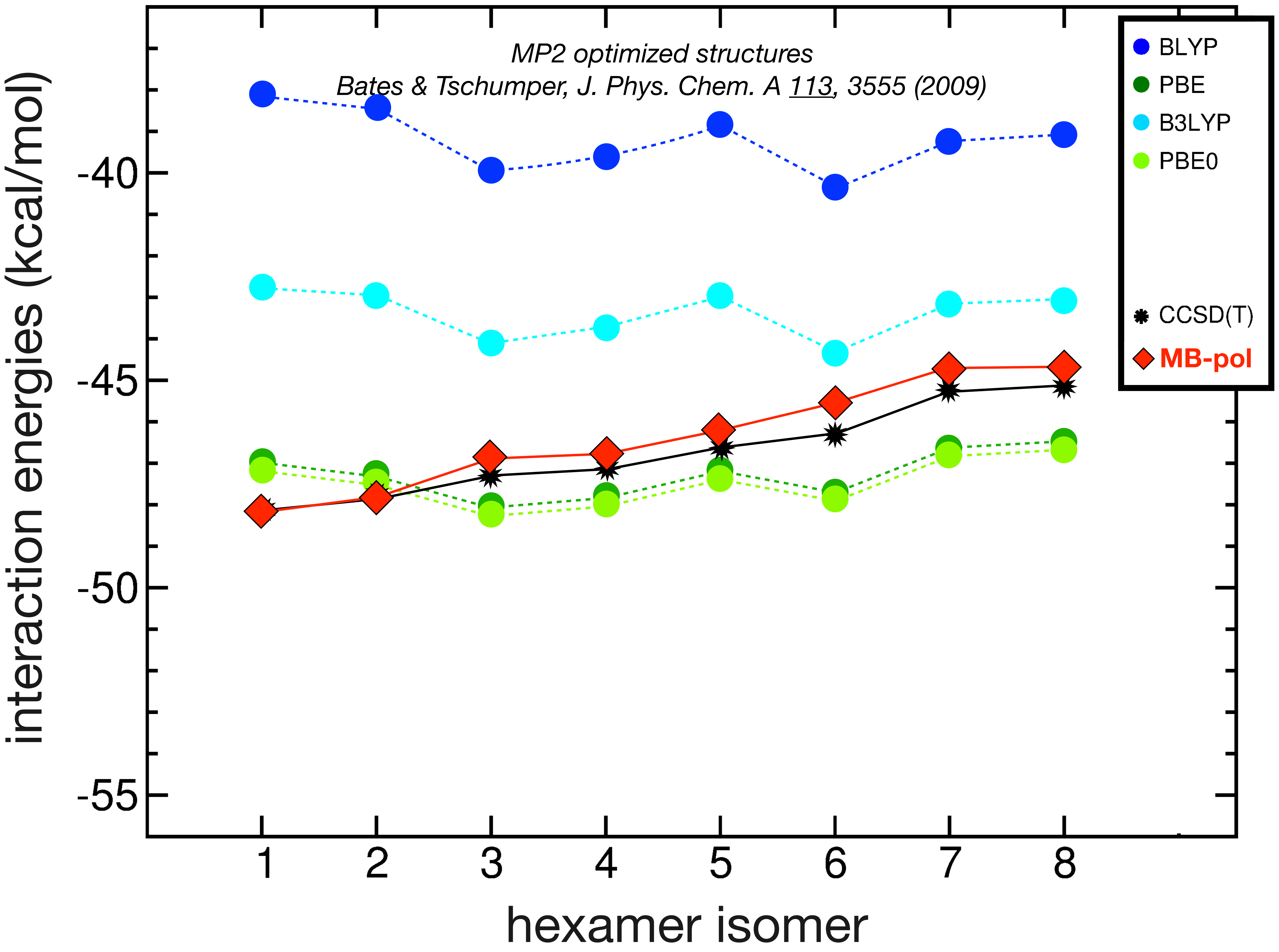

isomer 6


isomer 3


isomer 7


isomer 4


isomer 8

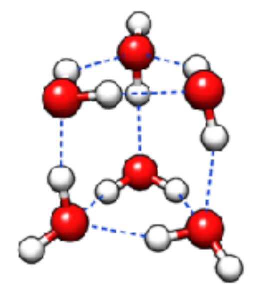

MP2 optimized structures

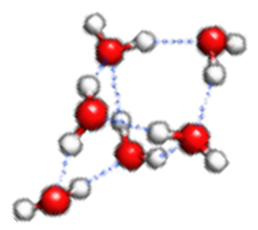
Bates & Tschumper, J. Phys. Chem. A *113*, *3555 (2009)*

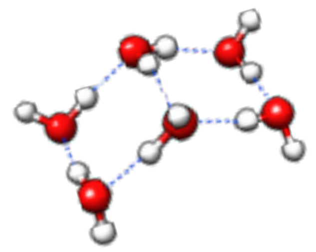
Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. *116*, *7501 (2016)*

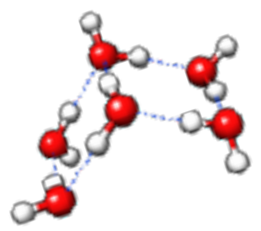
MB-pol in Perspective

MB-pol vs. DFT: Interaction energies for water clusters

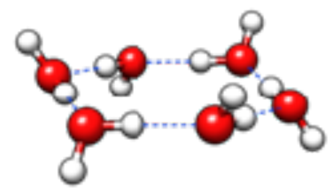

isomer 1


isomer 2

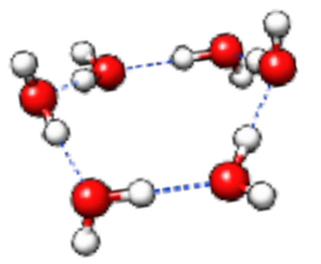

isomer 3

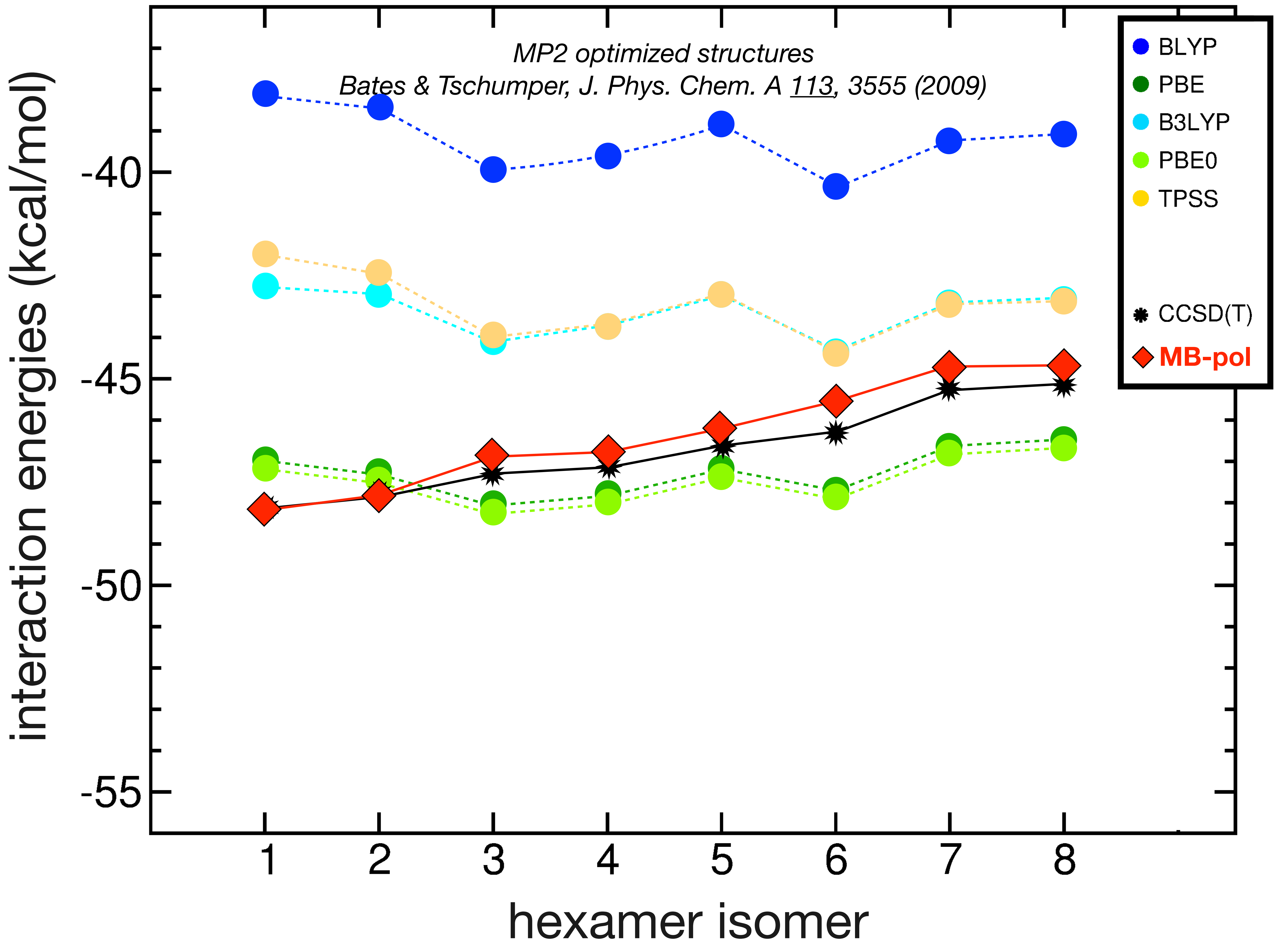

isomer 4


isomer 5


isomer 6


isomer 7


isomer 8

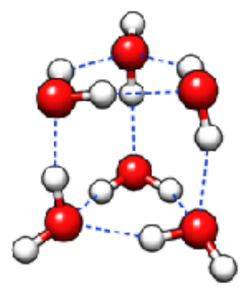

MP2 optimized structures

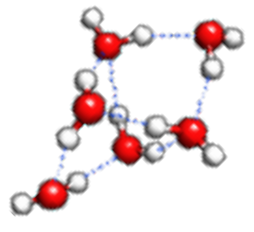
Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

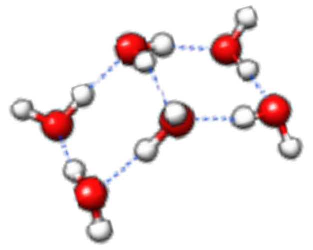
Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

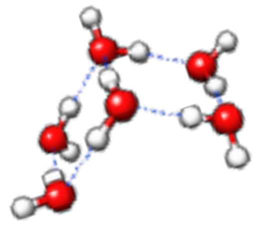
MB-pol in Perspective

MB-pol vs. DFT: Interaction energies for water clusters

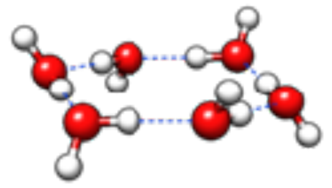

isomer 1

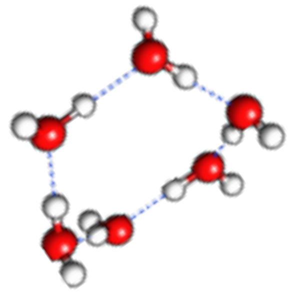

isomer 2

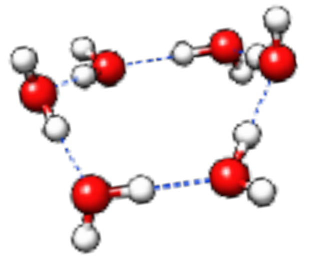

isomer 3

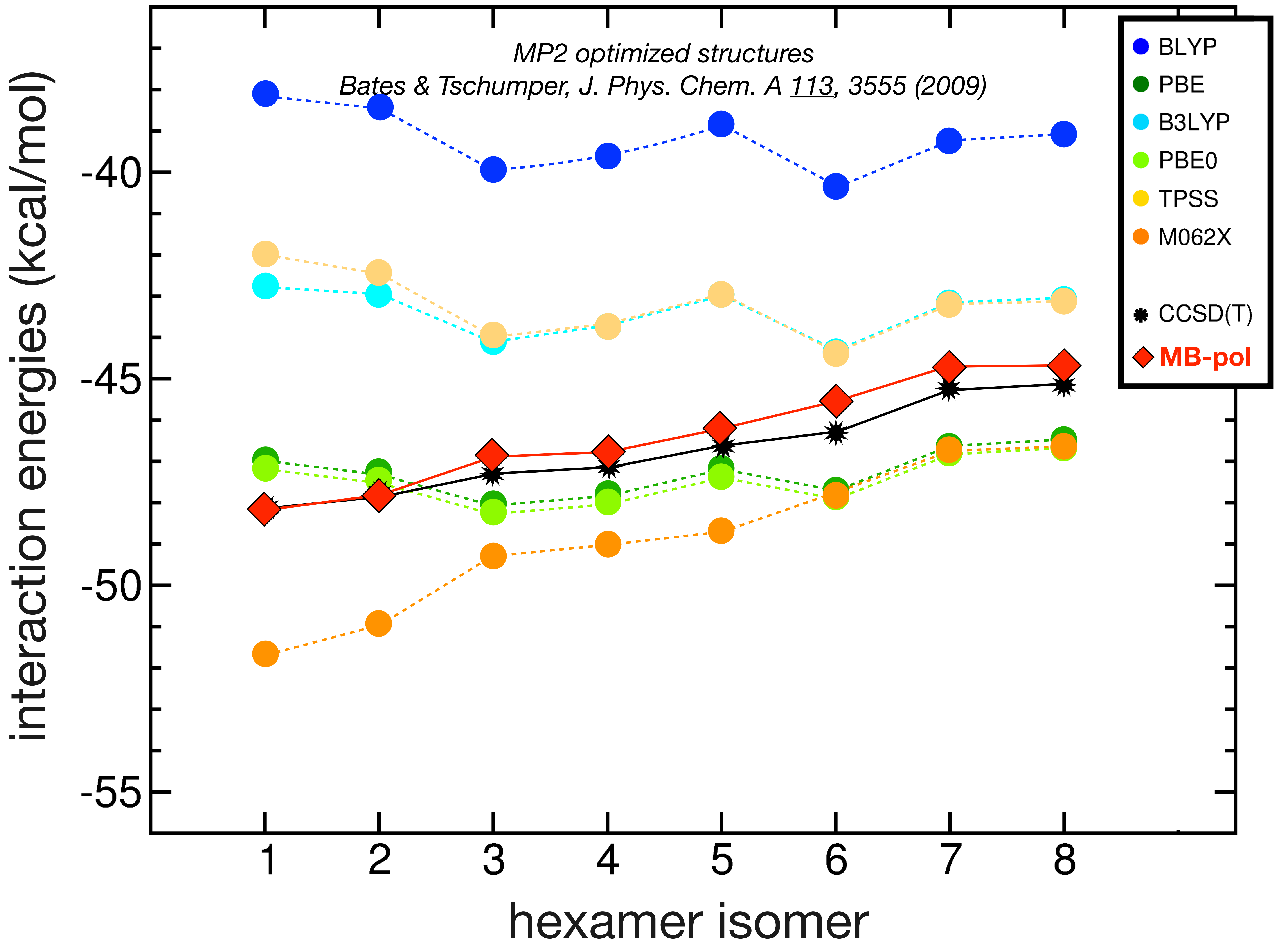

isomer 4


isomer 5


isomer 6


isomer 7


isomer 8

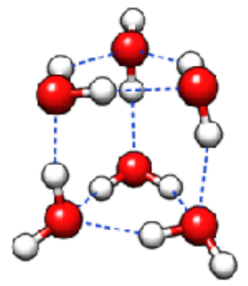

MP2 optimized structures

Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

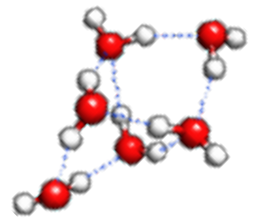
Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

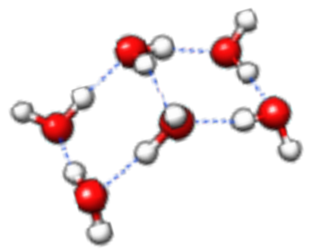
MB-pol vs. DFT: Interaction energies for water clusters



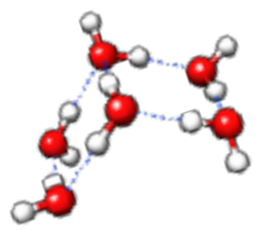
isomer 1



isomer 2

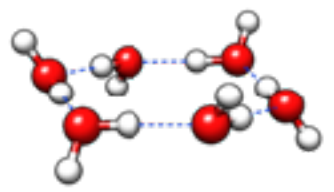


isomer 3

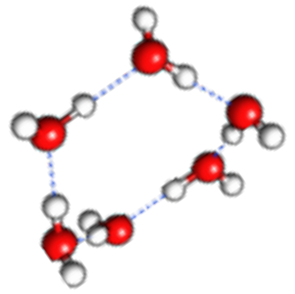


isomer 4

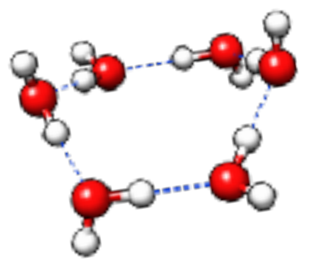
isomer 5



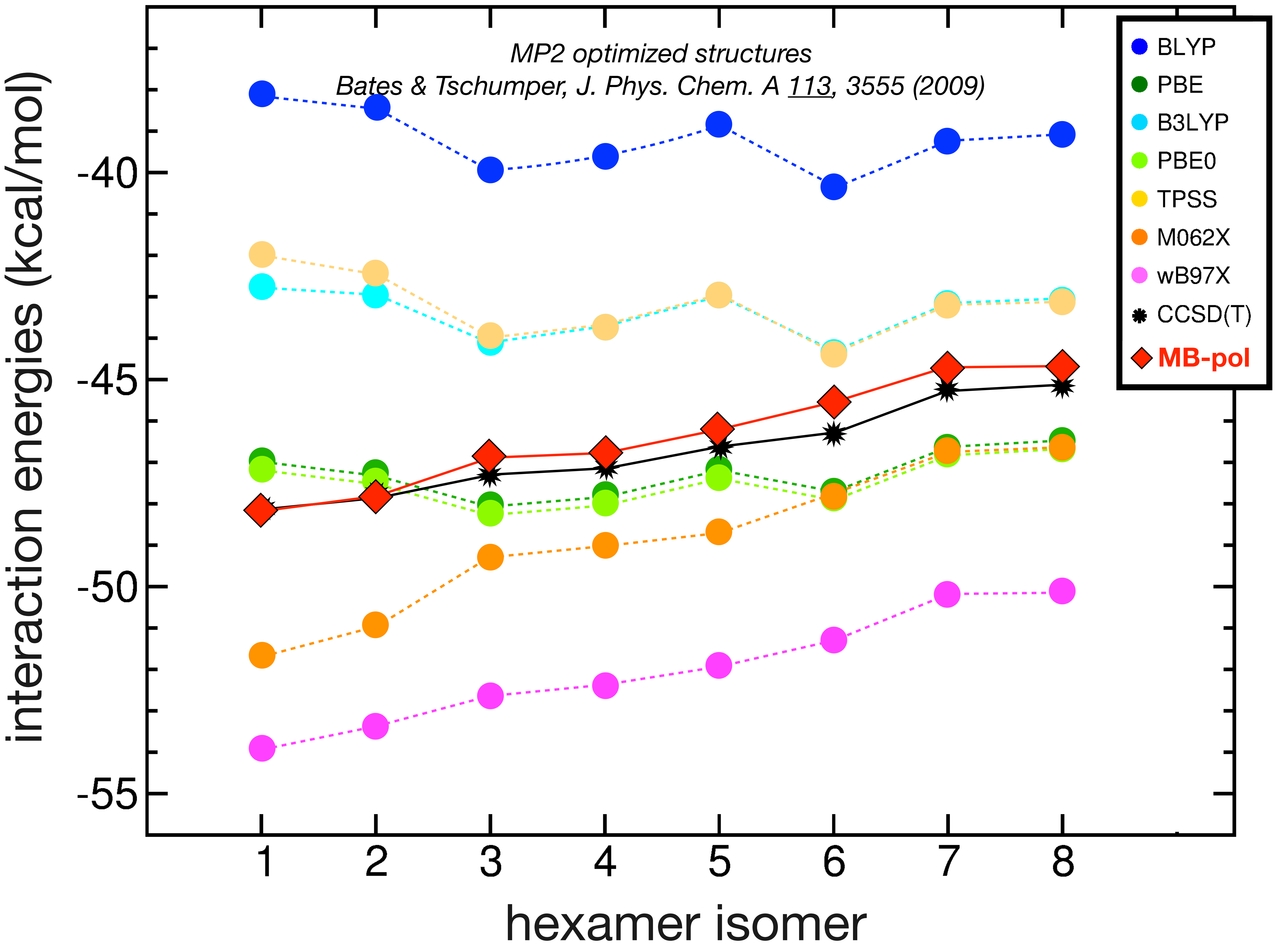
isomer 6



isomer 7



isomer 8



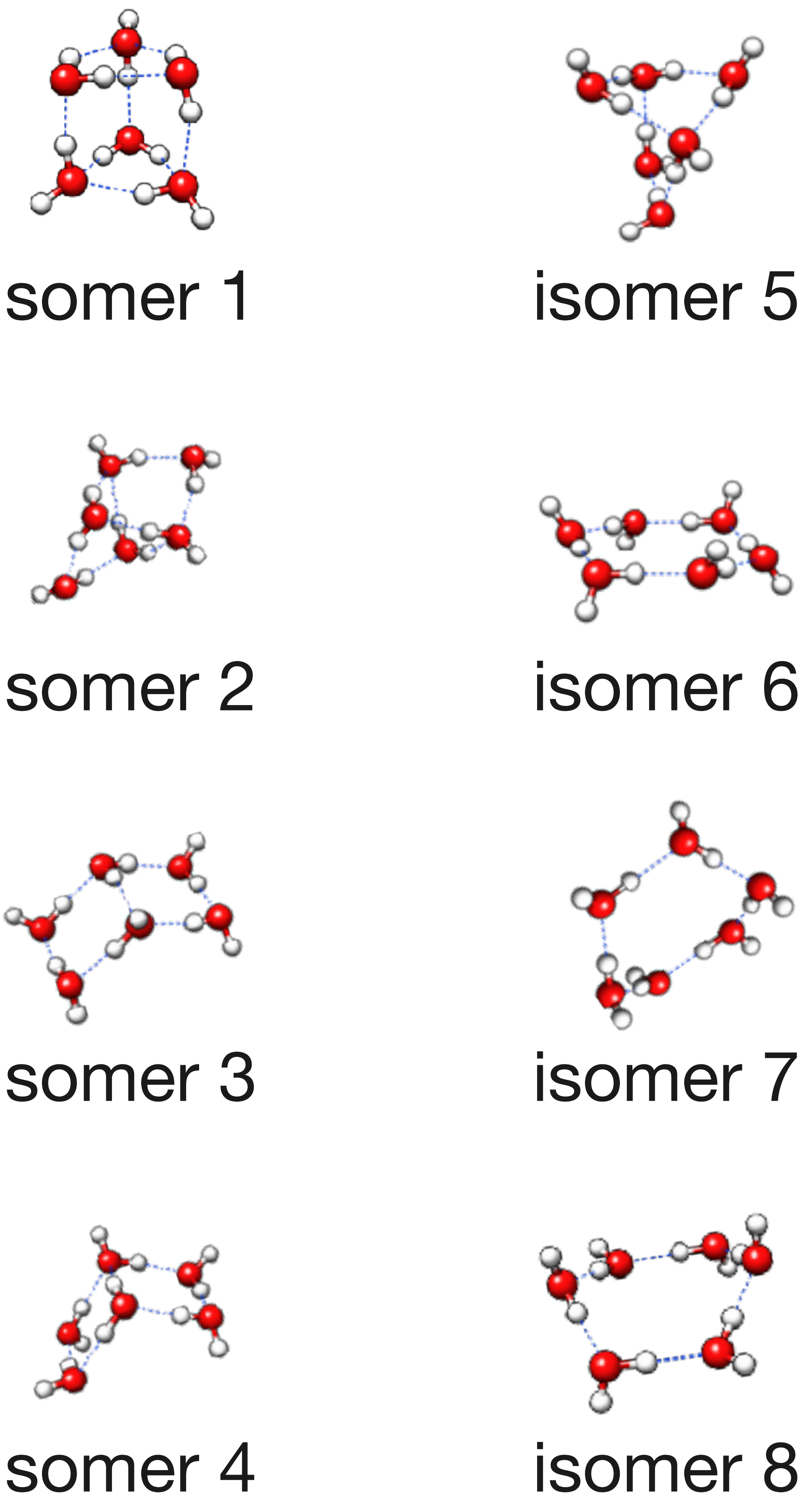
MP2 optimized structures

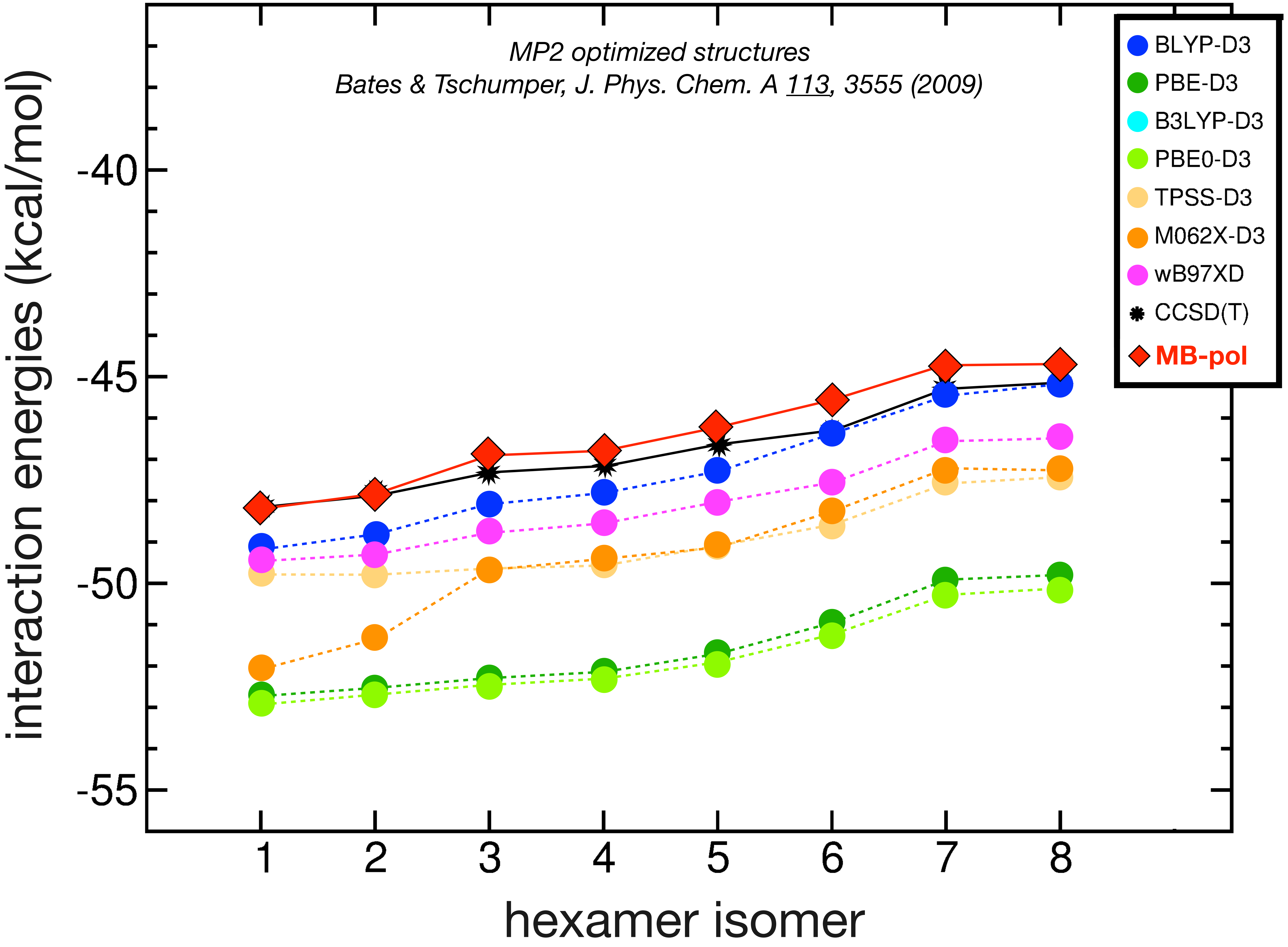
Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

MB-pol vs. DFT-D3: Interaction energies for water clusters

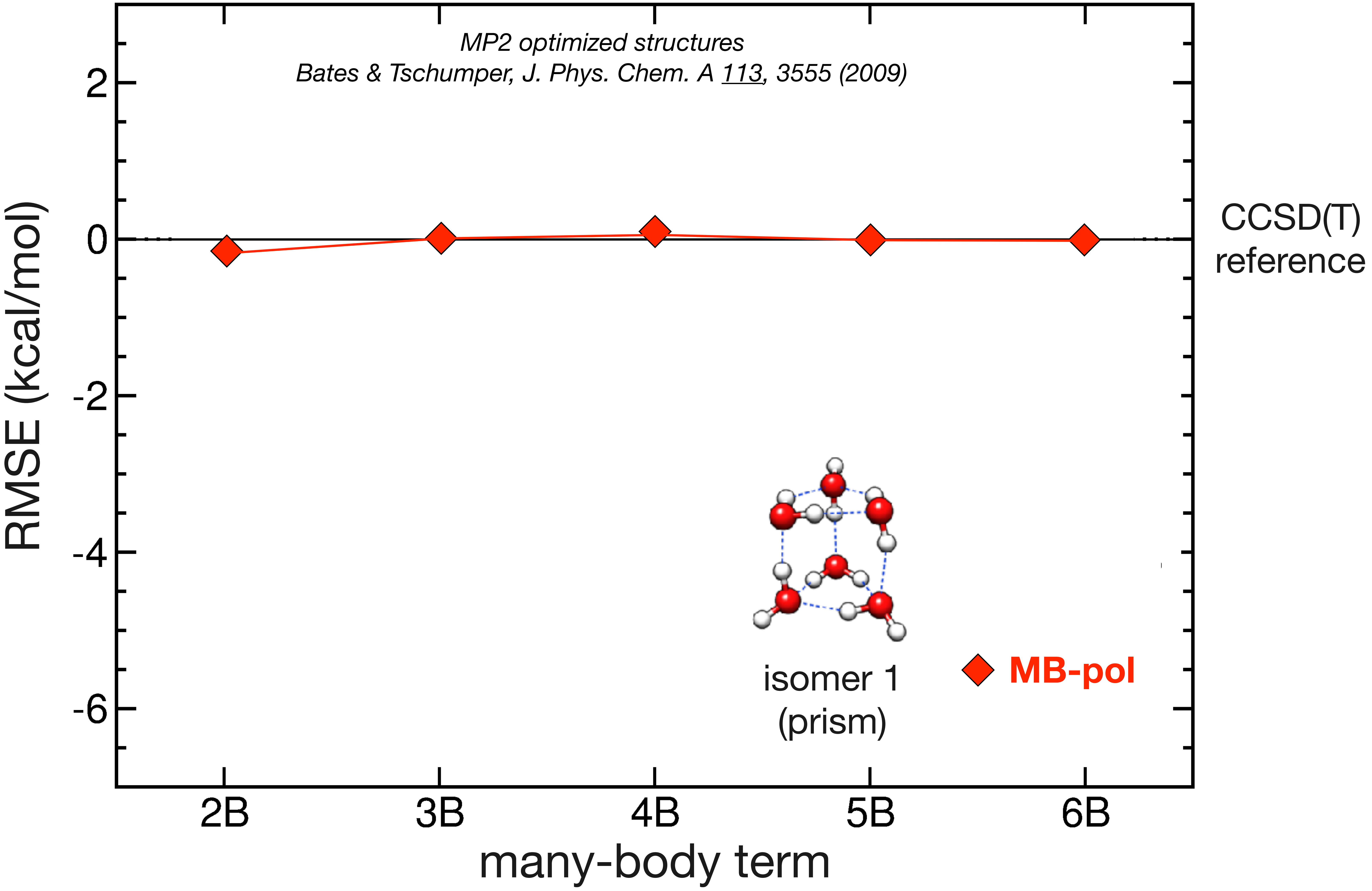




Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

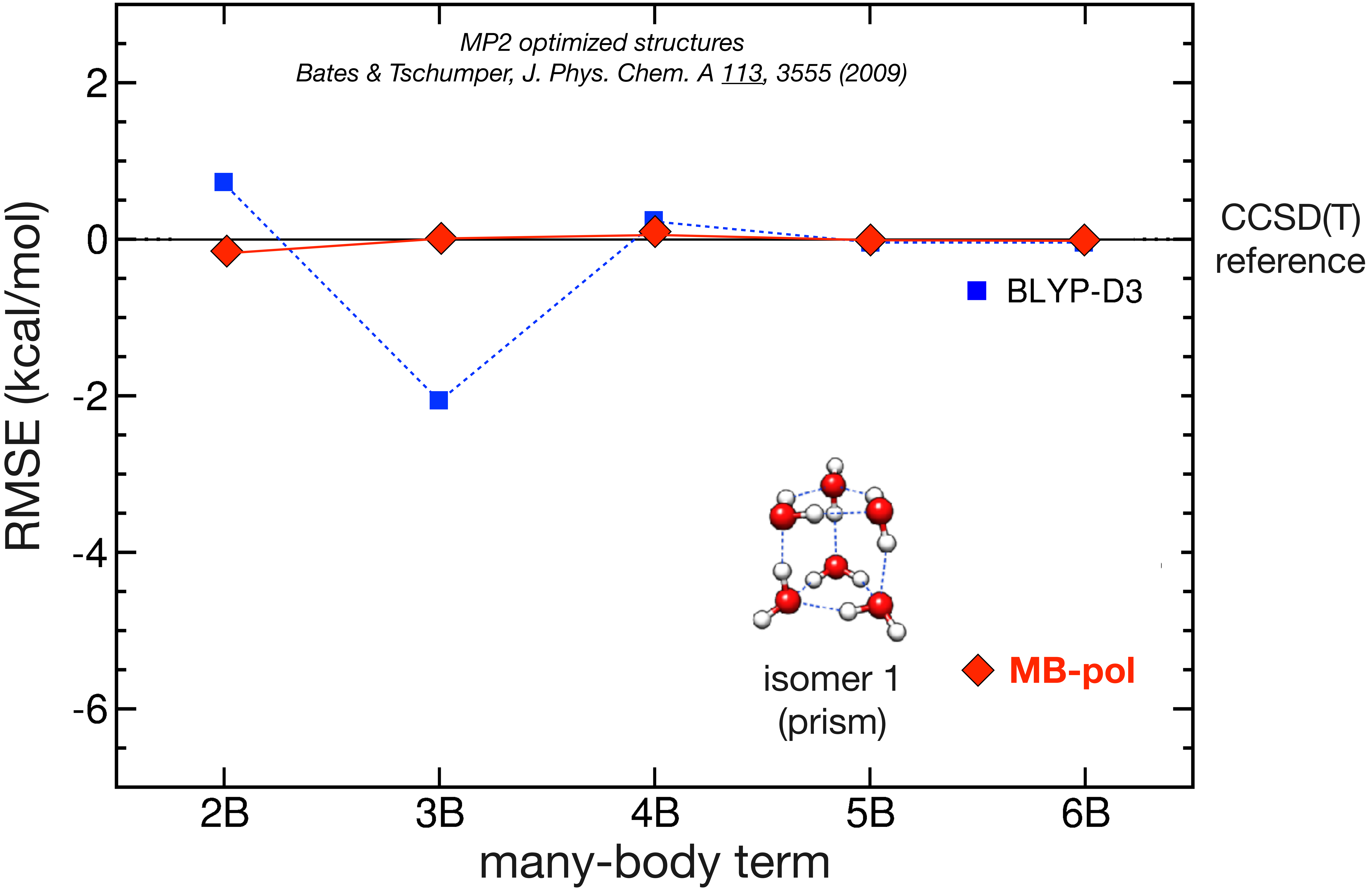
MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

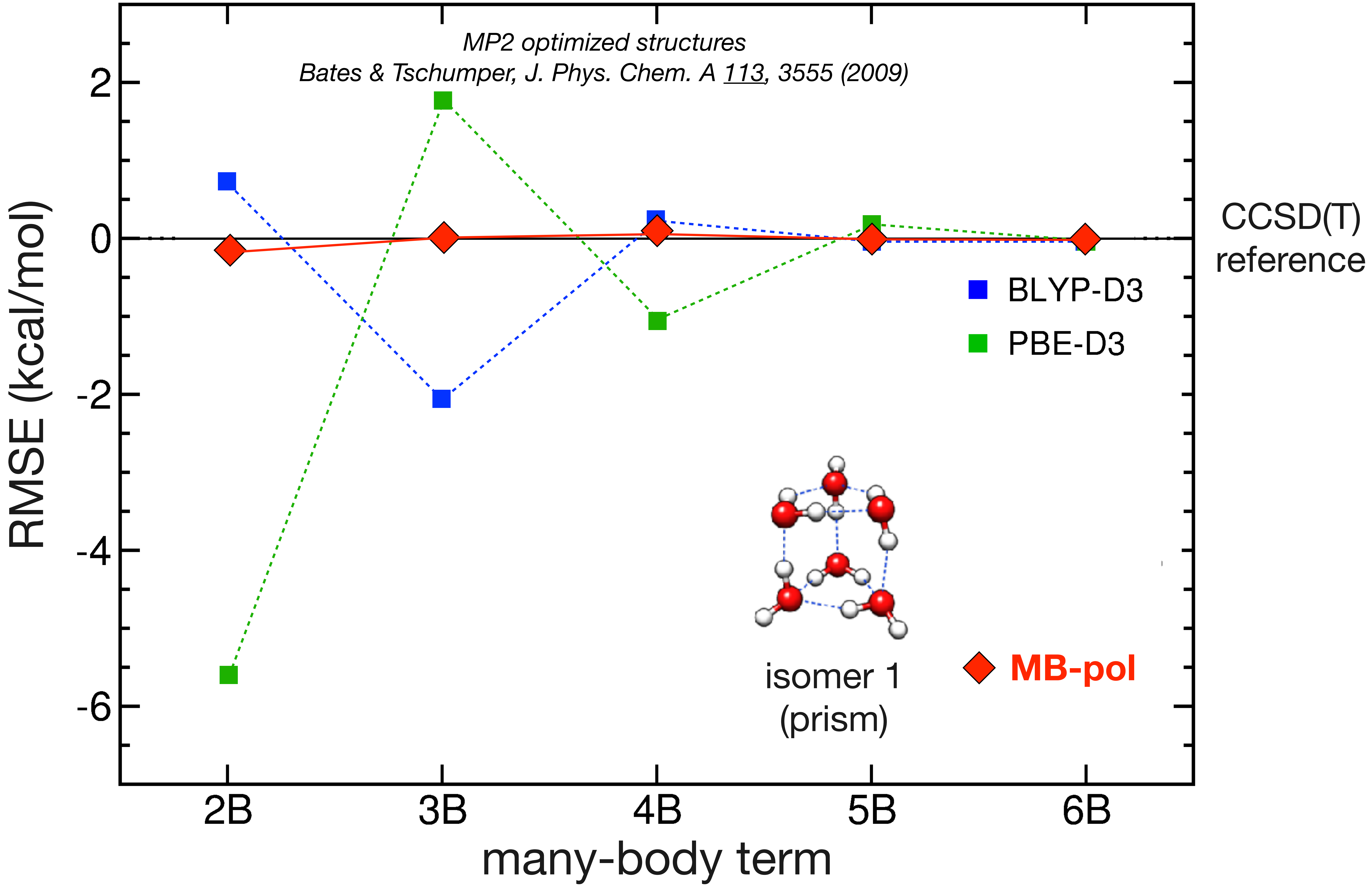
MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



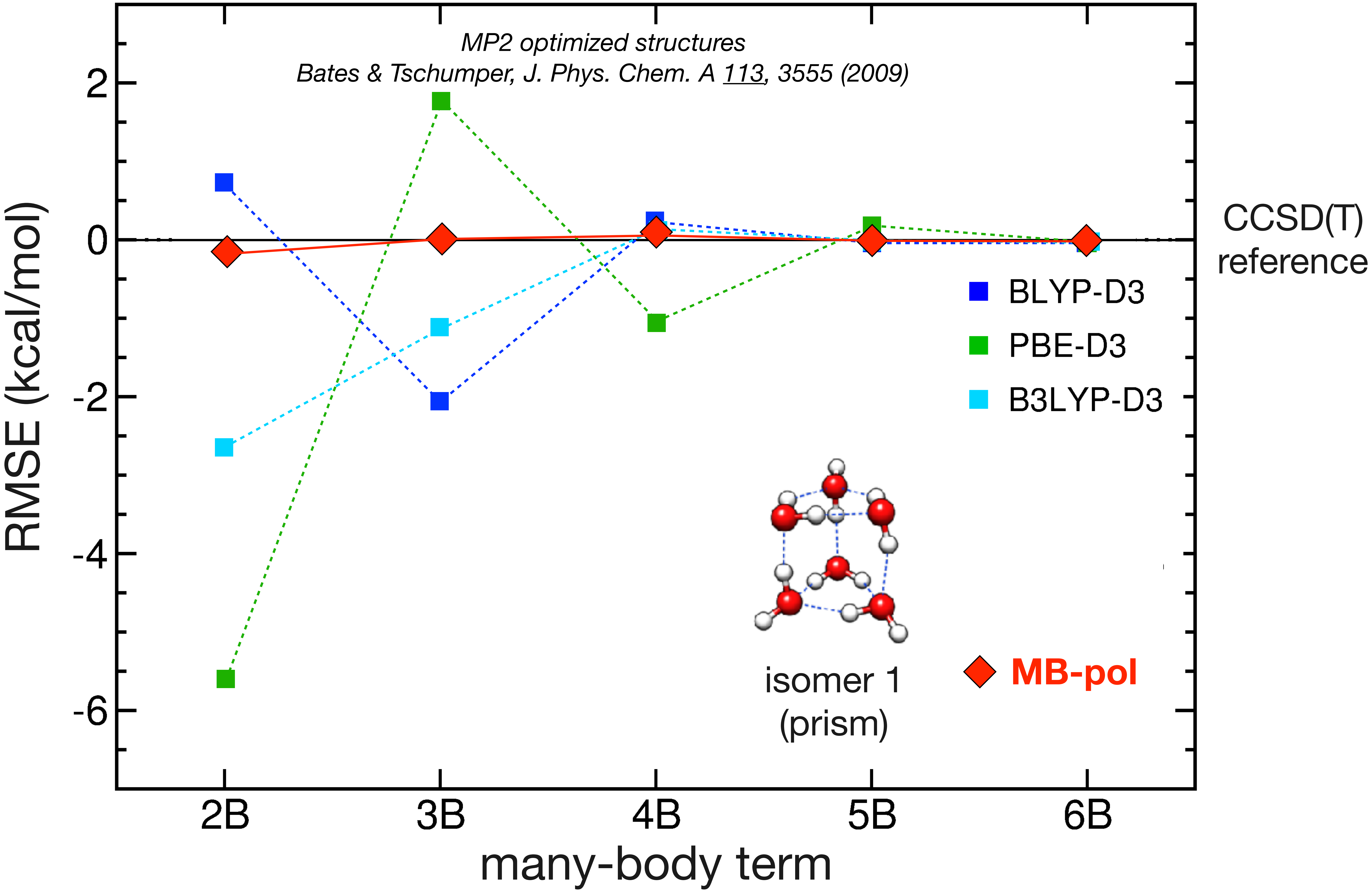
MP2 optimized structures

Bates & Tschumper, J. Phys. Chem. A 113, 3555 (2009)

Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

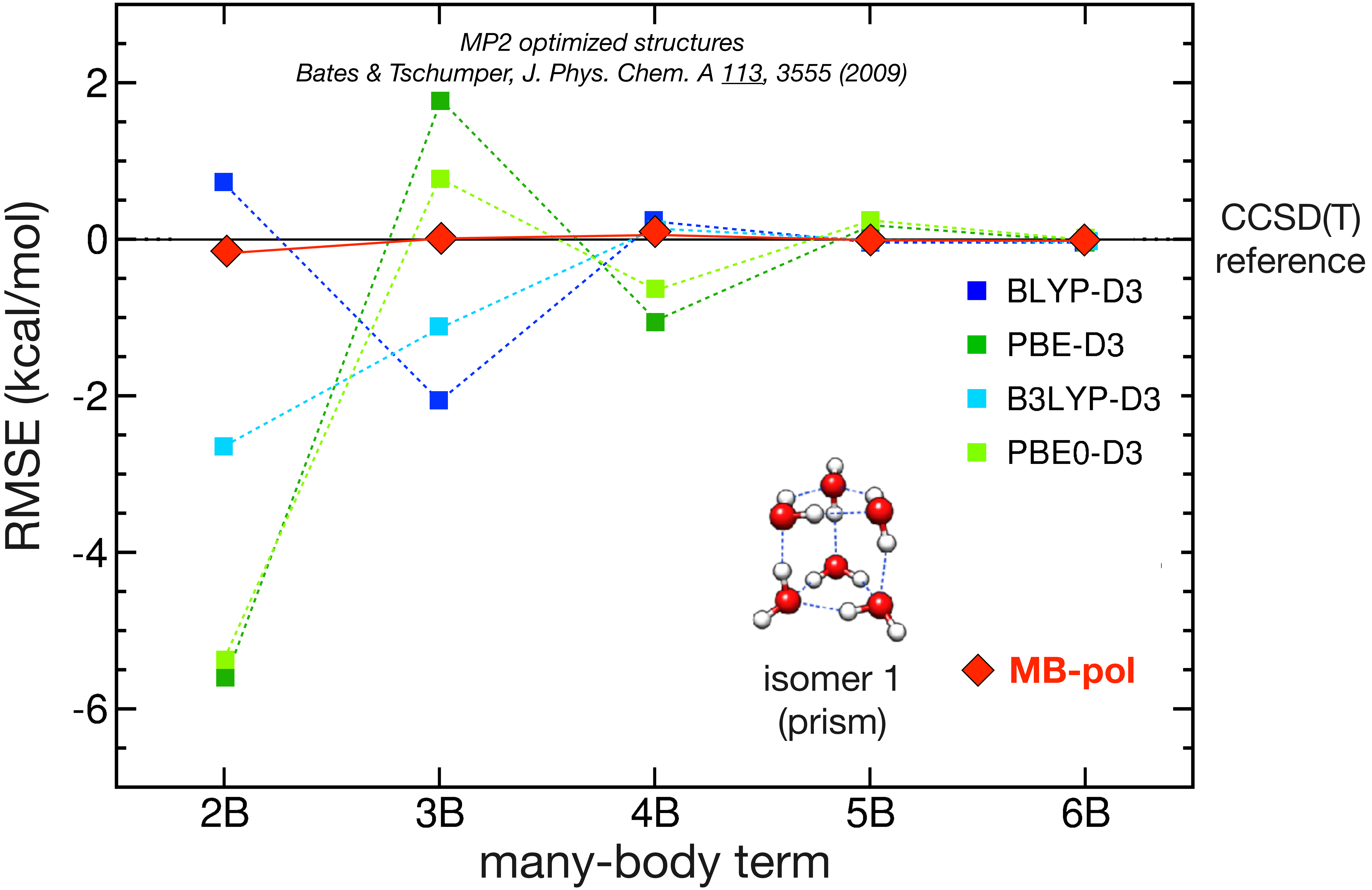
MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

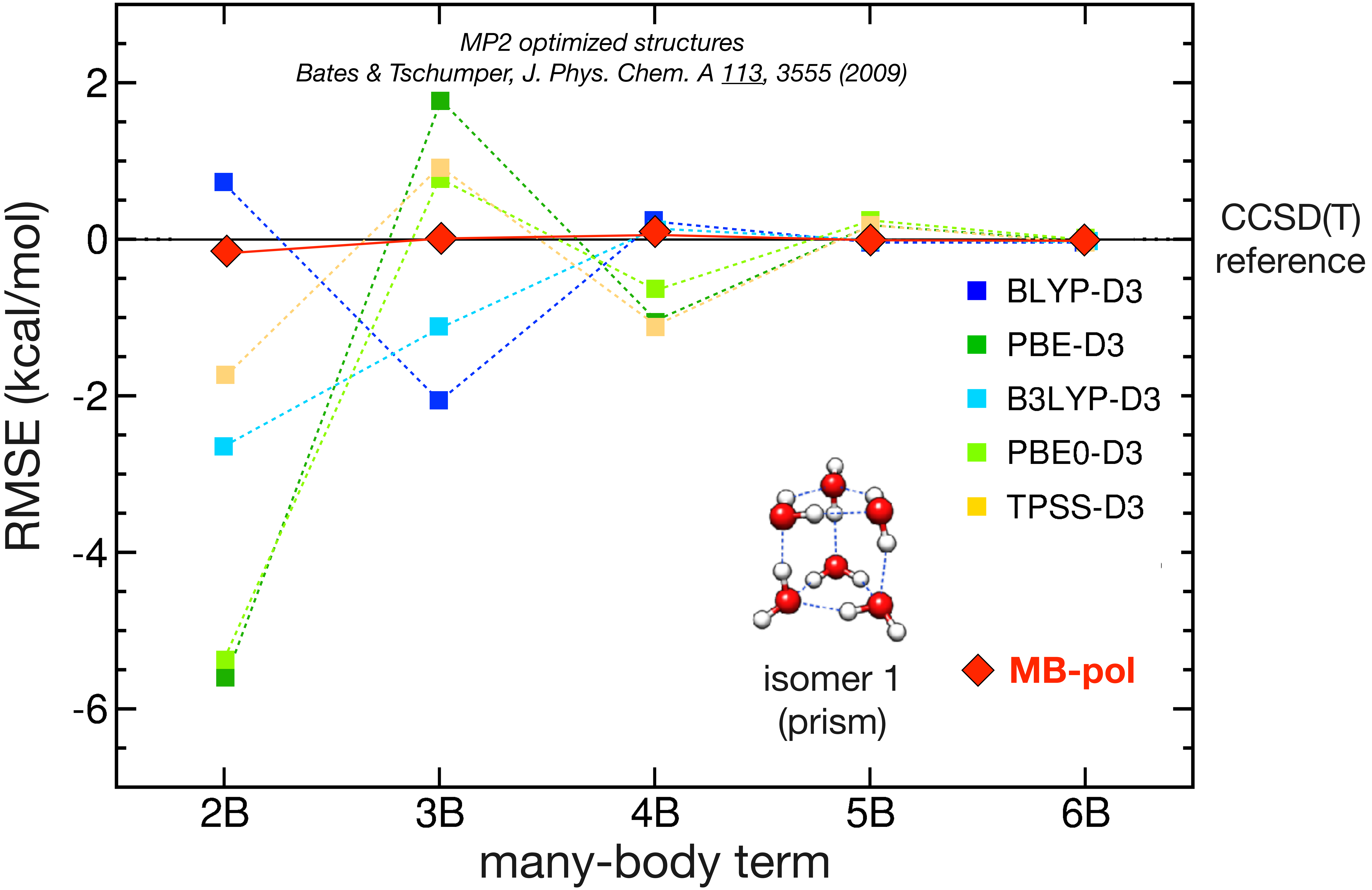
MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

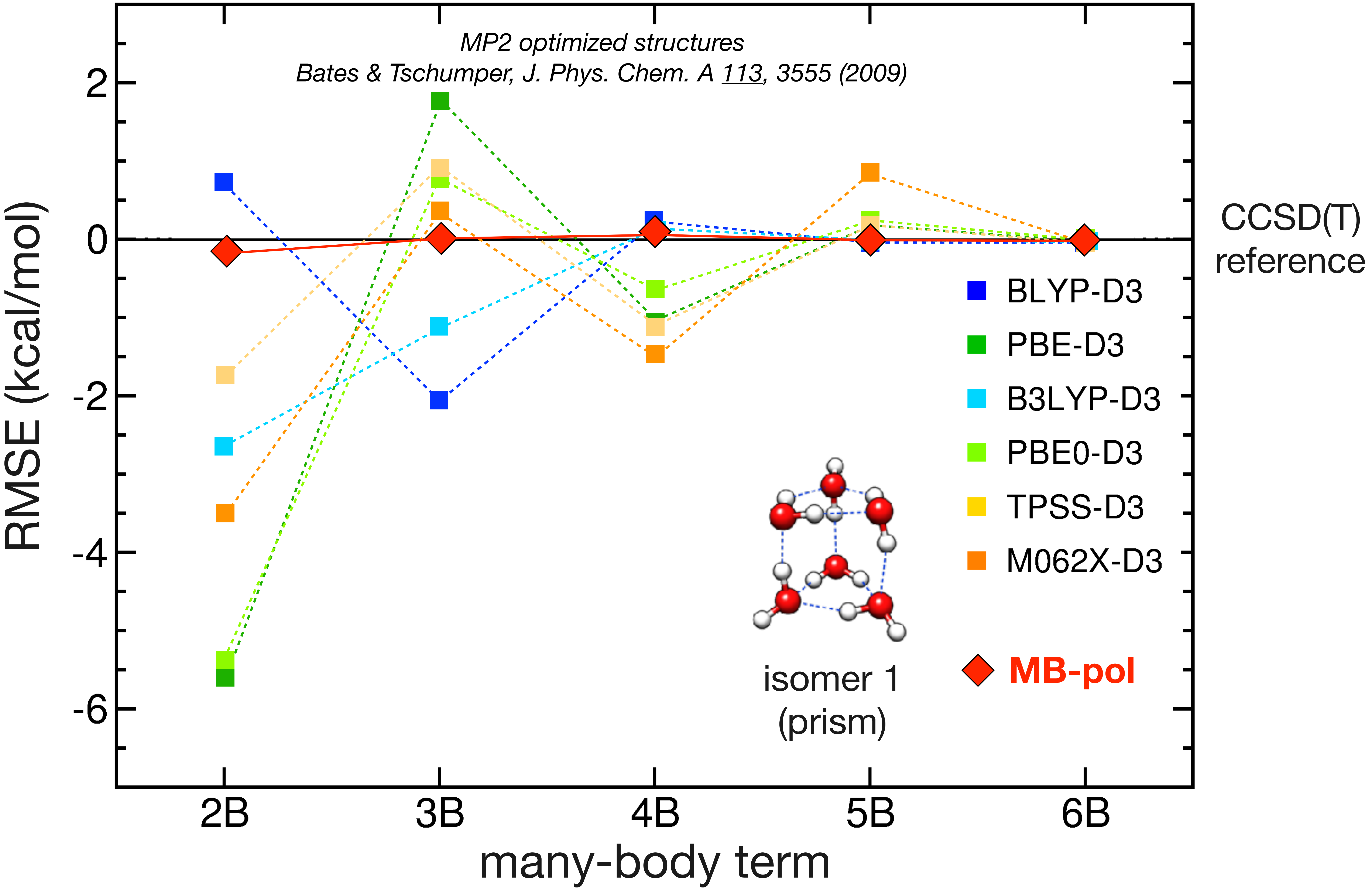
MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

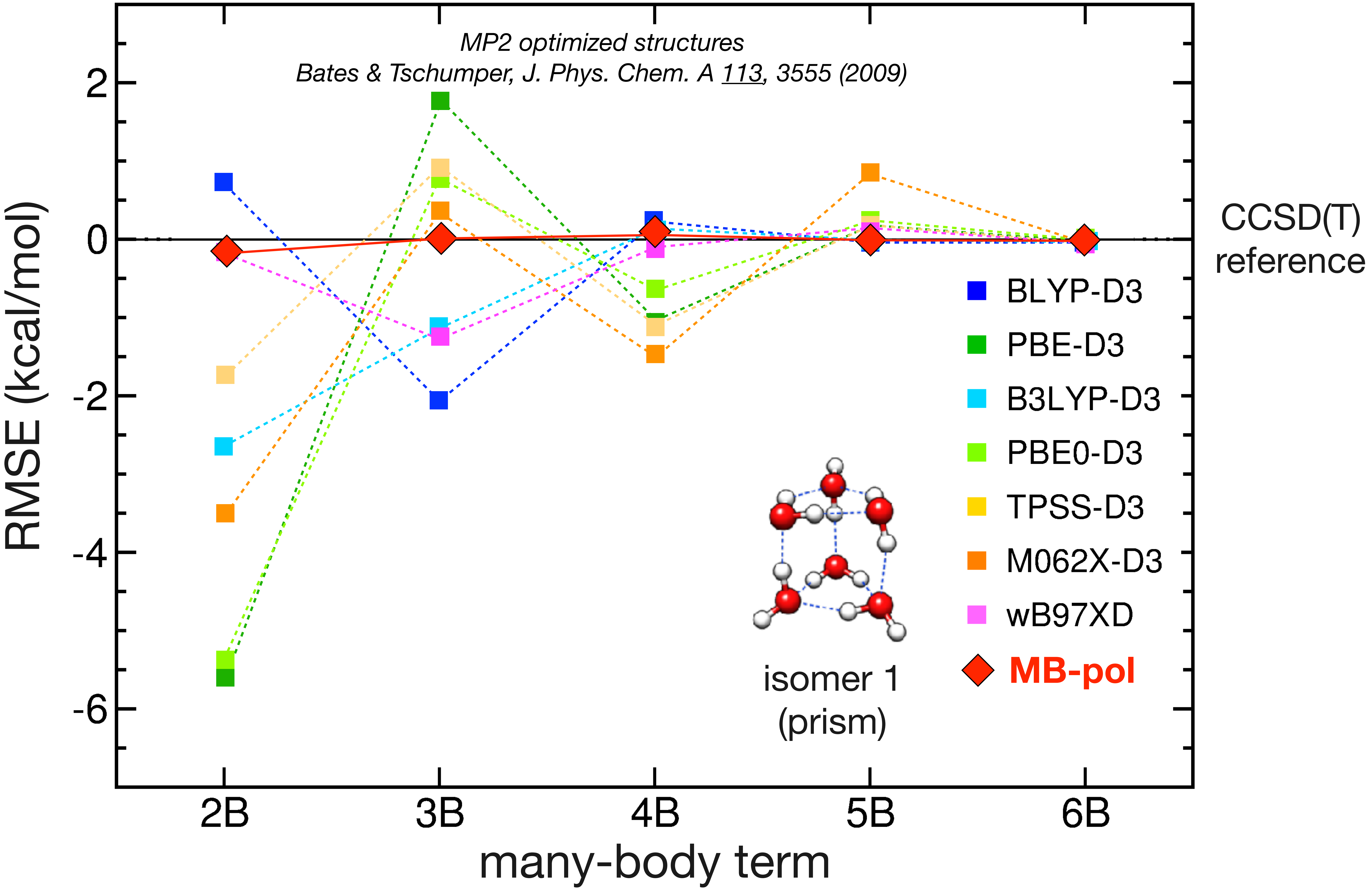
MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

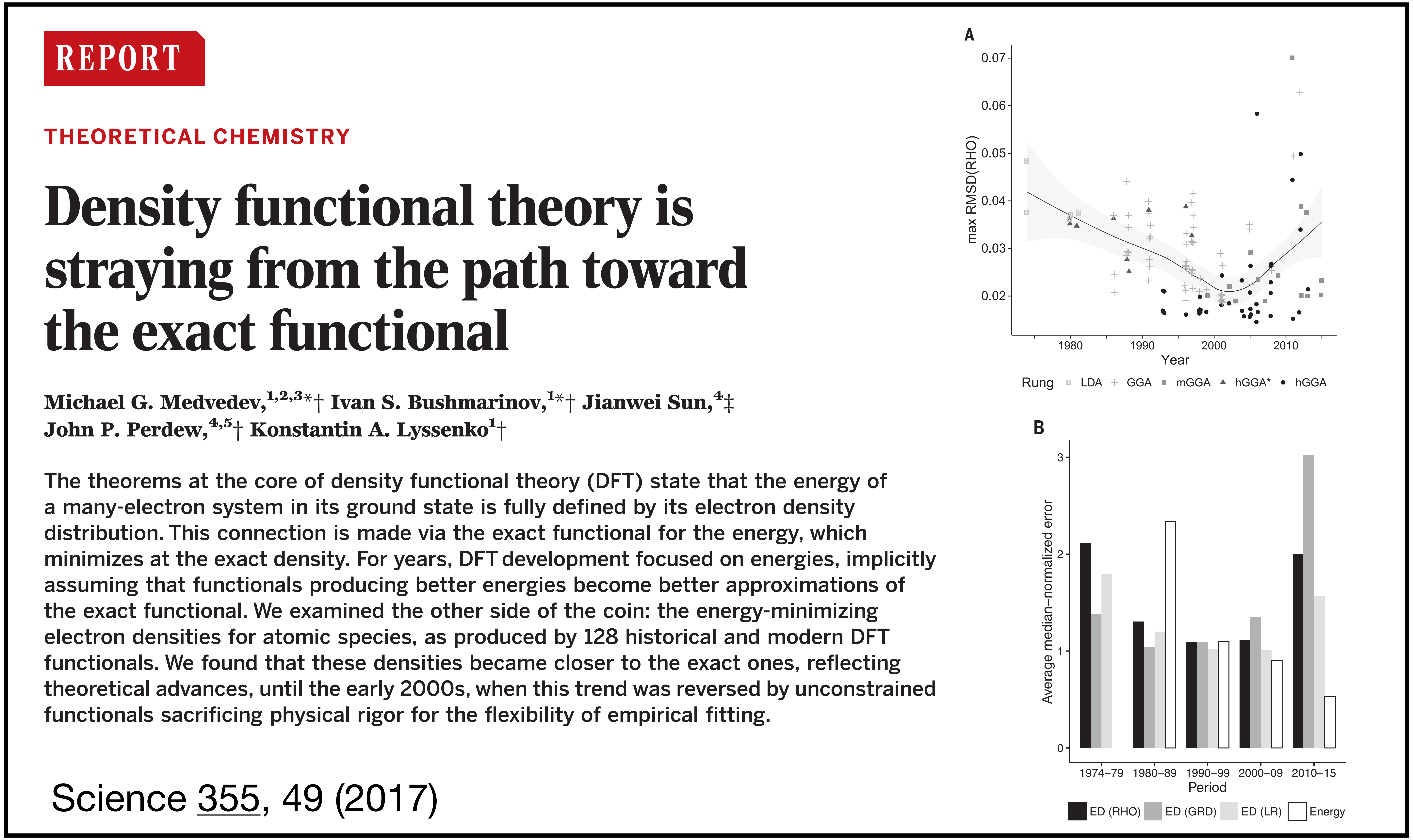
MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



REPORT

THEORETICAL CHEMISTRY

Density functional theory is straying from the path toward the exact functional

Michael G. Medvedev,[1,2,3]*† Ivan S. Bushmarinov,[1]*† Jianwei Sun,[4]‡ John P. Perdew,[4,5]† Konstantin A. Lyssenko[1]†

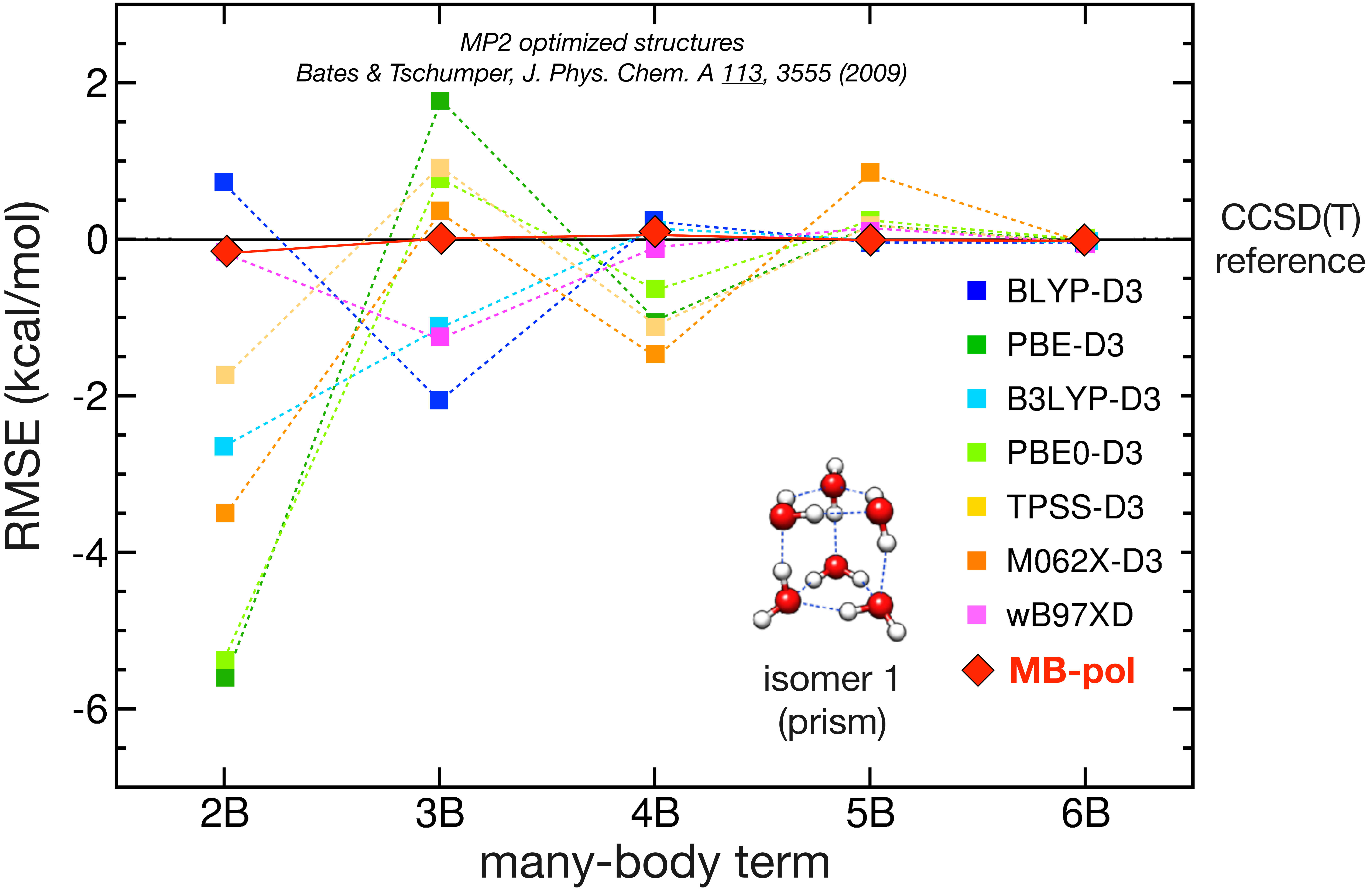
The theorems at the core of density functional theory (DFT) state that the energy of a many-electron system in its ground state is fully defined by its electron density distribution. This connection is made via the exact functional for the energy, which minimizes at the exact density. For years, DFT development focused on energies, implicitly assuming that functionals producing better energies become better approximations of the exact functional. We examined the other side of the coin: the energy-minimizing electron densities for atomic species, as produced by 128 historical and modern DFT functionals. We found that these densities became closer to the exact ones, reflecting theoretical advances, until the early 2000s, when this trend was reversed by unconstrained functionals sacrificing physical rigor for the flexibility of empirical fitting.

Science 355, 49 (2017)

Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

MB-pol in Perspective

MB-pol vs. DFT-D3: Many-body energy decomposition for water clusters



Cisneros, Wikfeldt, Ojamäe, Lu, Xu, Torabifard, Bartók-Pártay, Csányi, Molinero & FP, Chem. Rev. 116, 7501 (2016)

Many-Body Effects in Water

Climbing Jacob’s ladder of DFT approximations

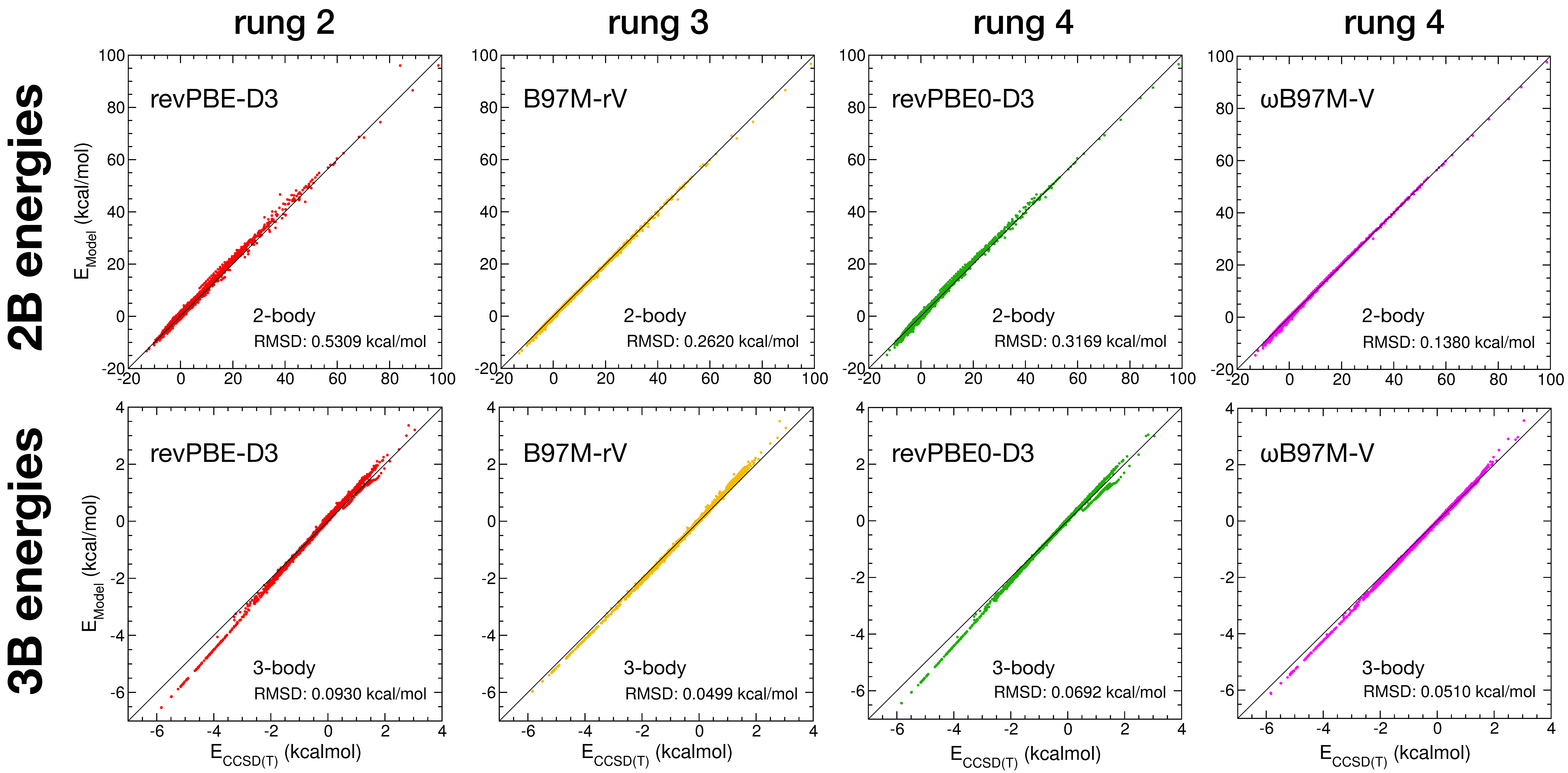
2B and 3B energies

Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

Many-Body Effects in Water

Climbing Jacob’s ladder of DFT approximations

2B and 3B energies

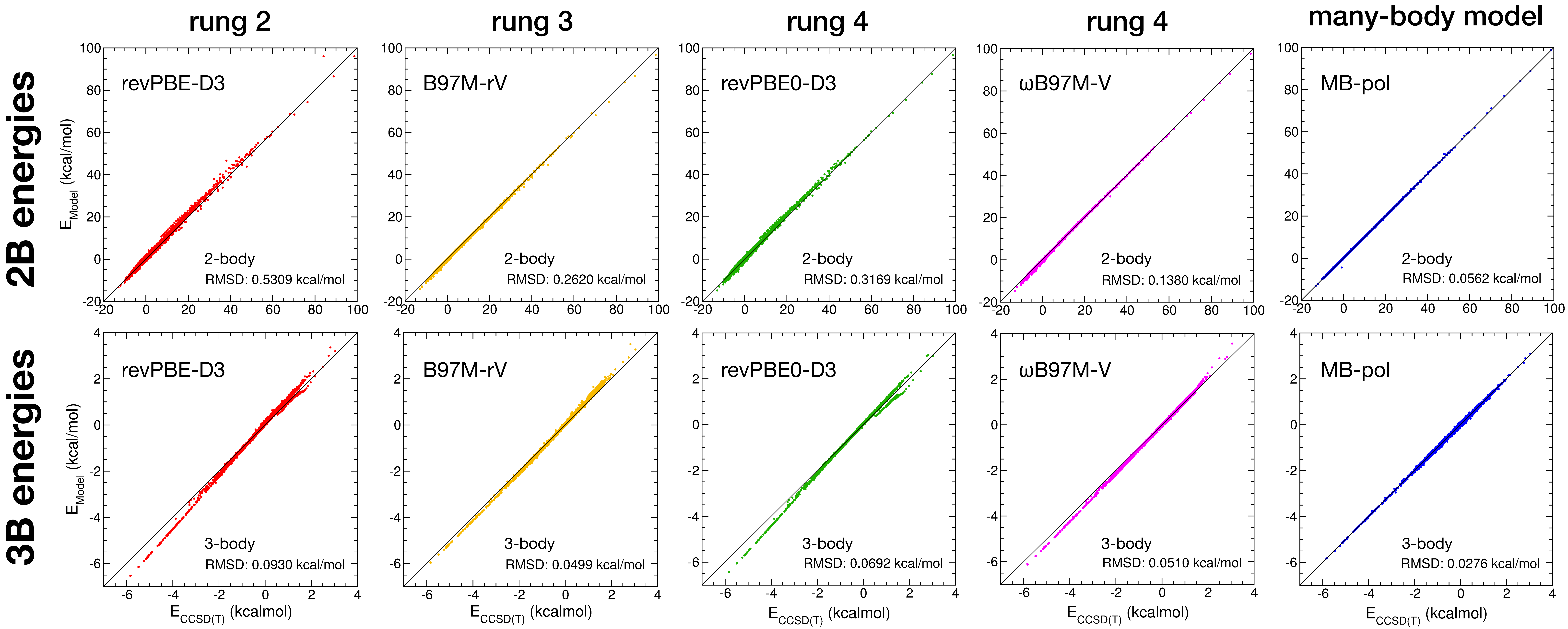


Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

Many-Body Effects in Water

Climbing Jacob's ladder of DFT approximations

2B and 3B energies



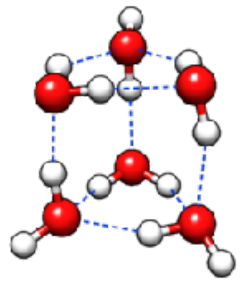
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

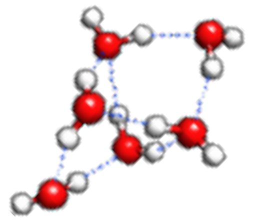
Many-Body Effects in Water

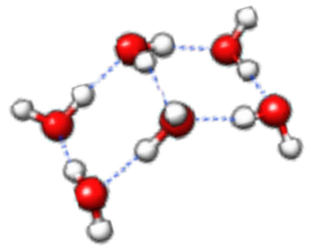
Climbing Jacob's ladder of DFT approximations

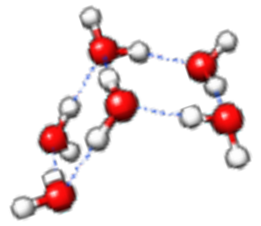
GGA: revPBE-D3

water hexamer

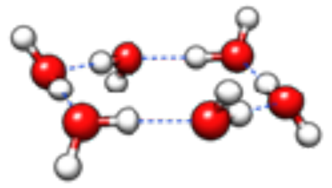

isomer 1

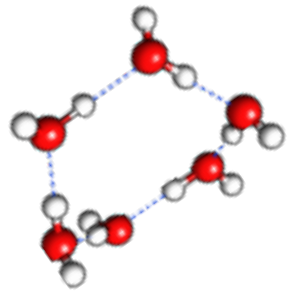

isomer 2

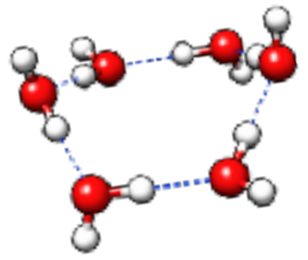

isomer 3


isomer 4

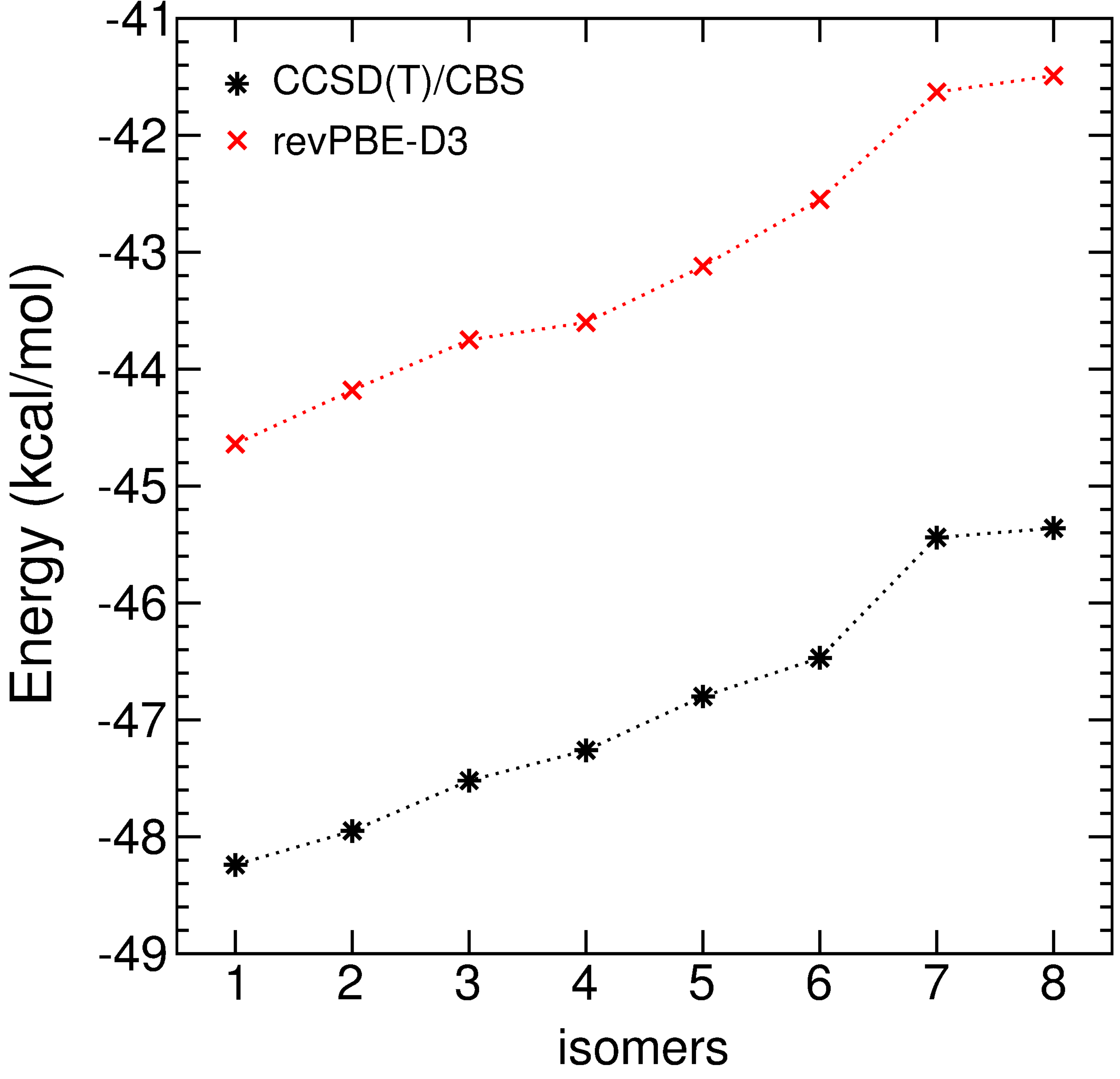

isomer 5


isomer 6

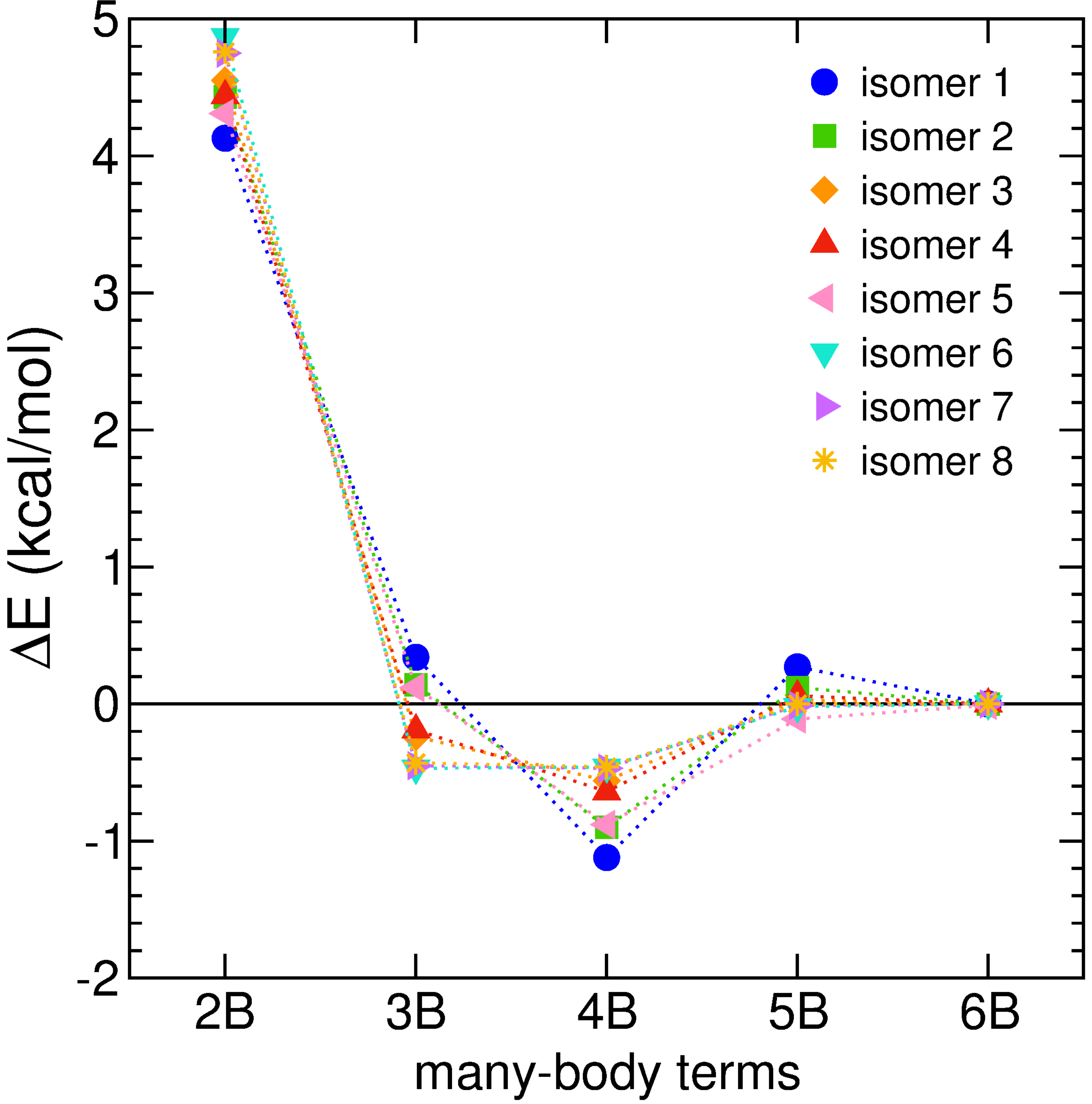

isomer 7


isomer 8

interaction energies



many-body contributions



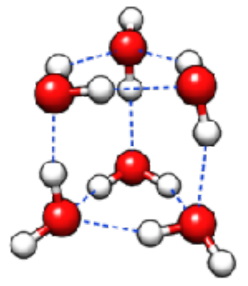
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

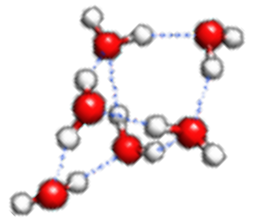
Many-Body Effects in Water

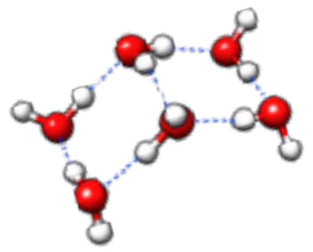
Climbing Jacob’s ladder of DFT approximations

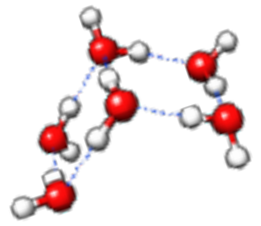
Meta-GGA: B97M-rV

water hexamer

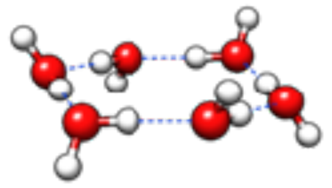

isomer 1


isomer 2

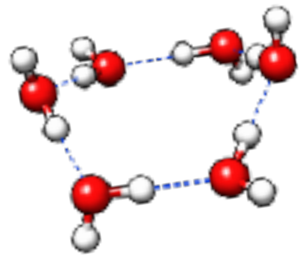

isomer 3

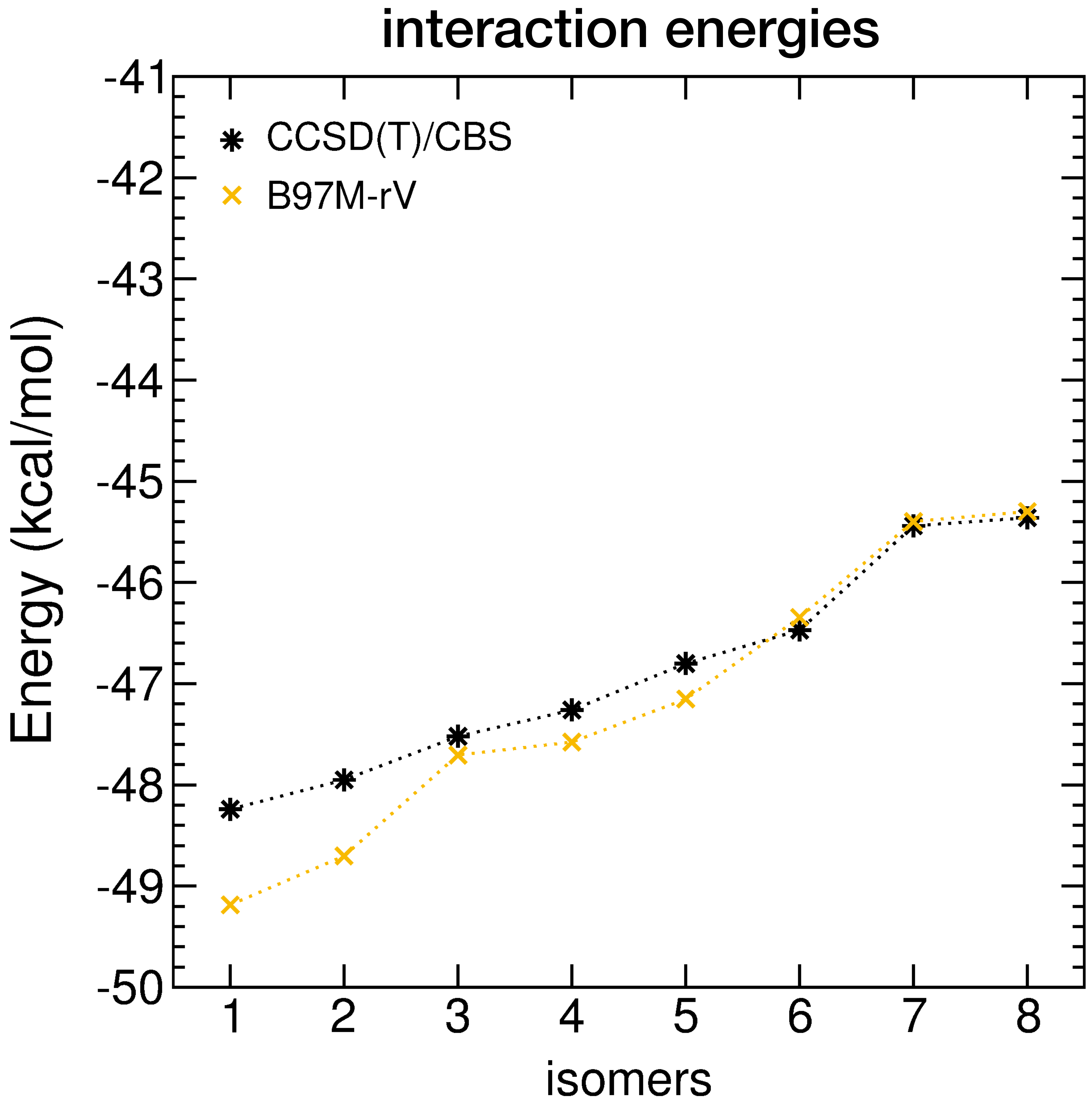

isomer 4

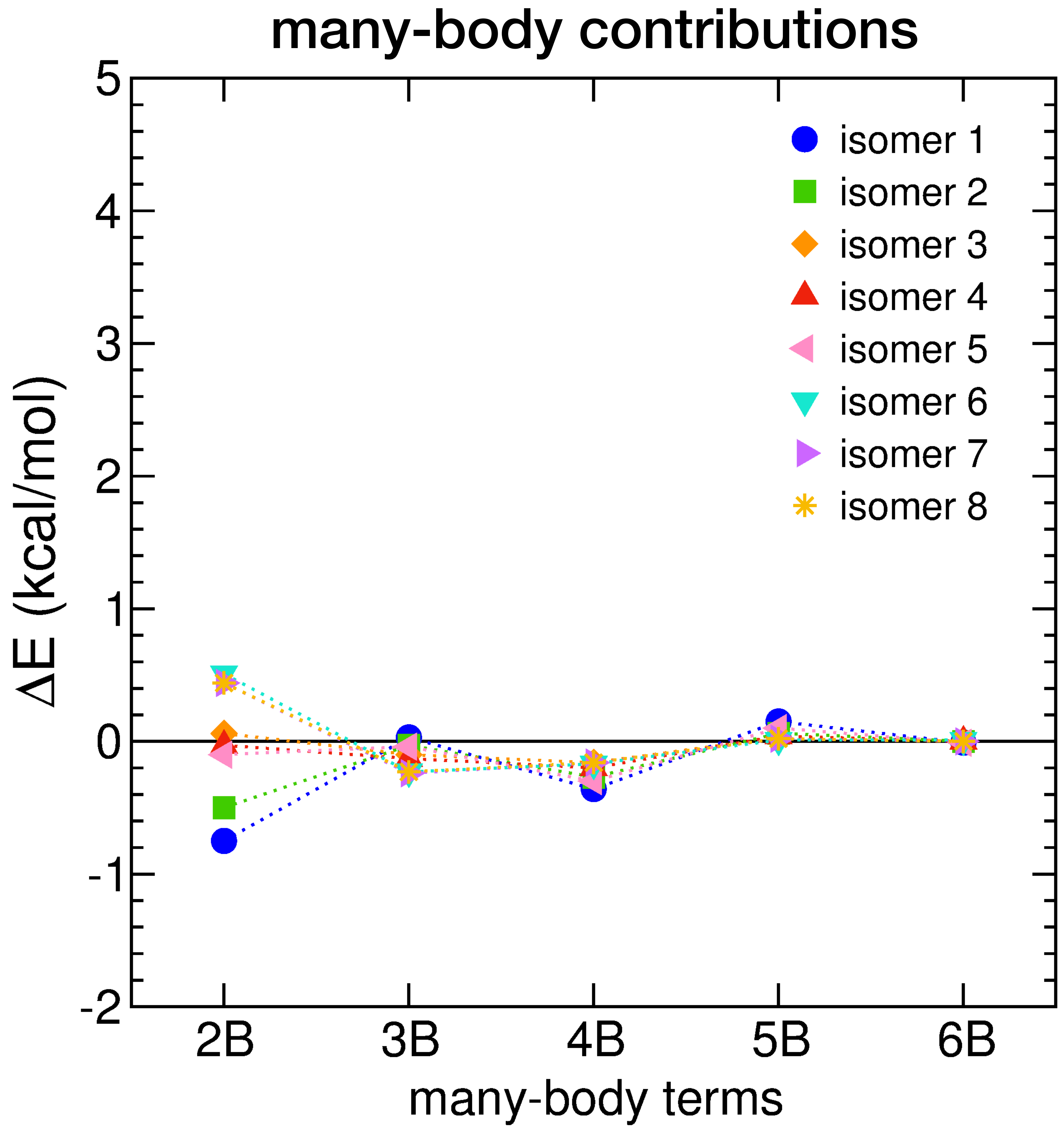

isomer 5


isomer 6


isomer 7


isomer 8


interaction energies


many-body contributions

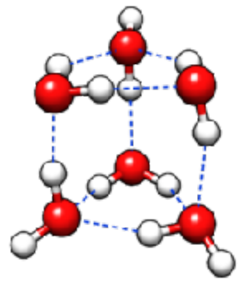
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

Many-Body Effects in Water

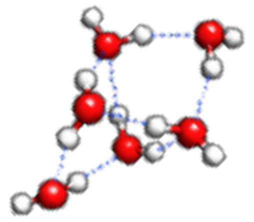
Climbing Jacob's ladder of DFT approximations

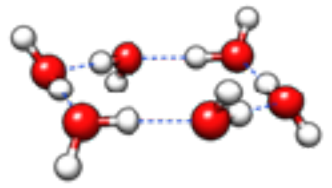
Meta-GGA: SCAN

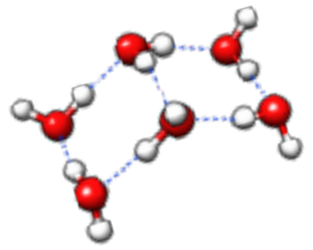
water hexamer

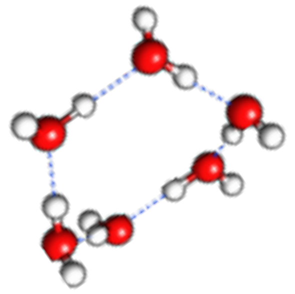

isomer 1

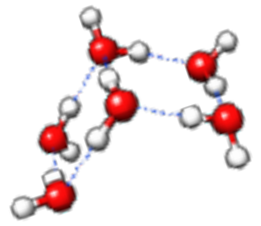

isomer 5

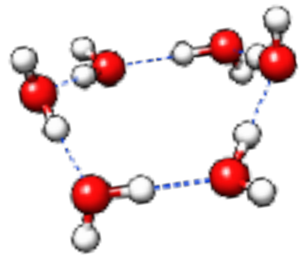

isomer 2

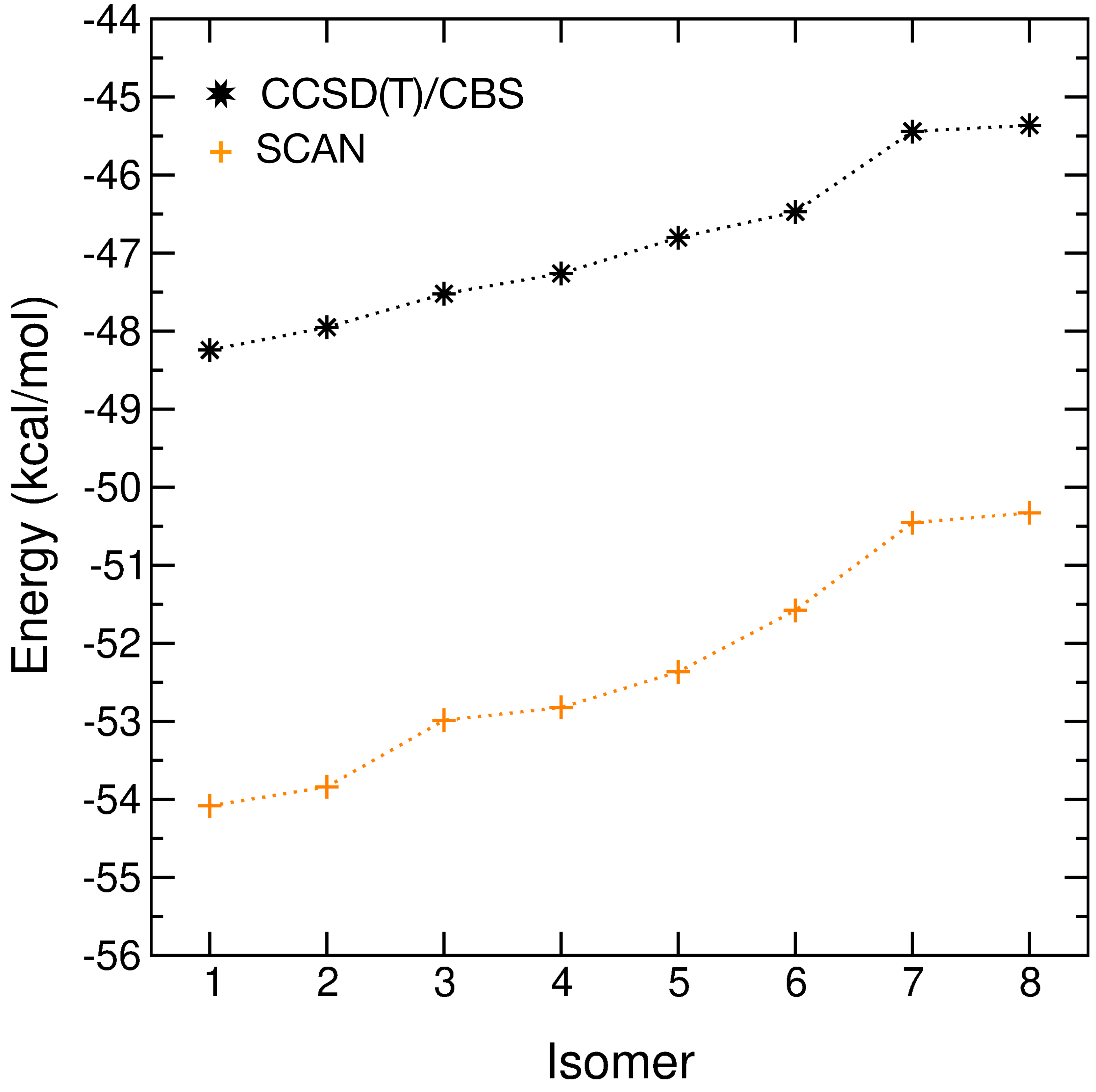

isomer 6

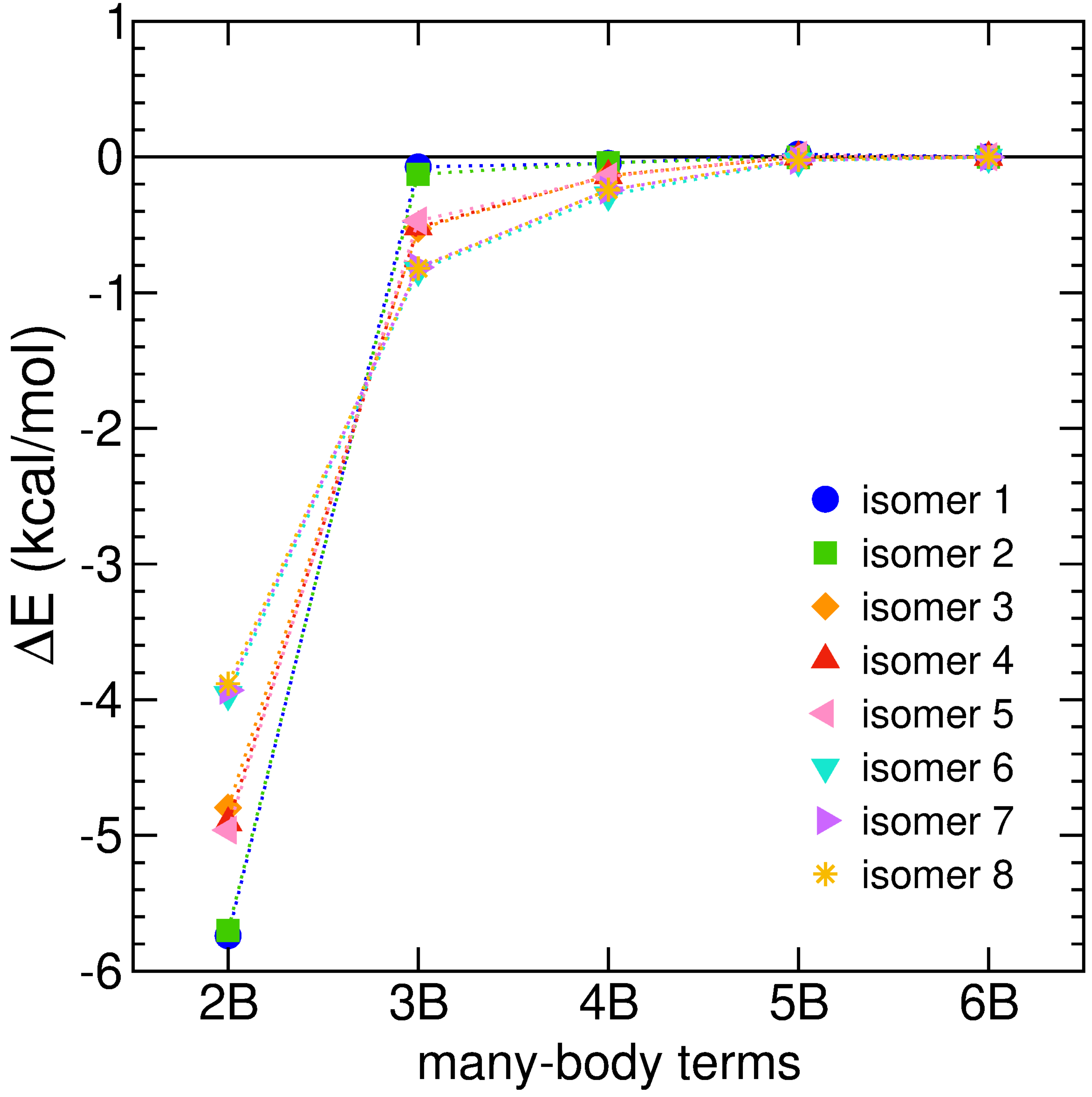

isomer 3


isomer 7


isomer 4


isomer 8


interaction energies


many-body contributions

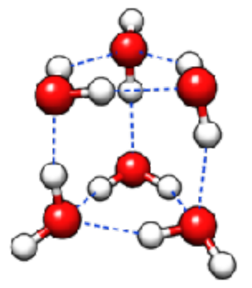
Lambros, Hu & FP, in progress

Many-Body Effects in Water

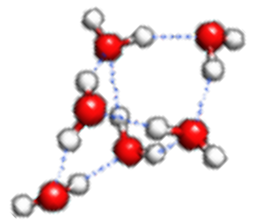
Climbing Jacob’s ladder of DFT approximations

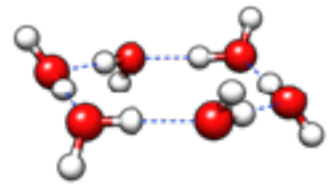
Hybrid: revPBE0-D3

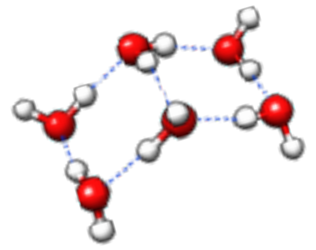
water hexamer

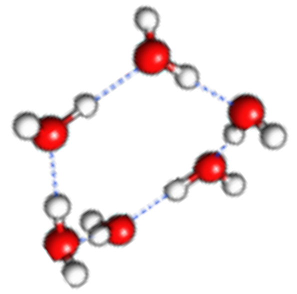

isomer 1

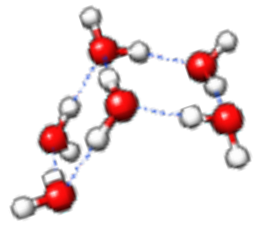

isomer 5

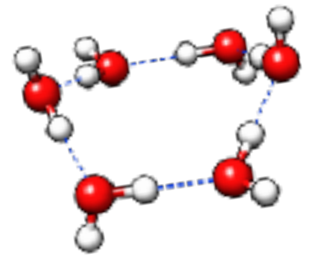

isomer 2


isomer 6

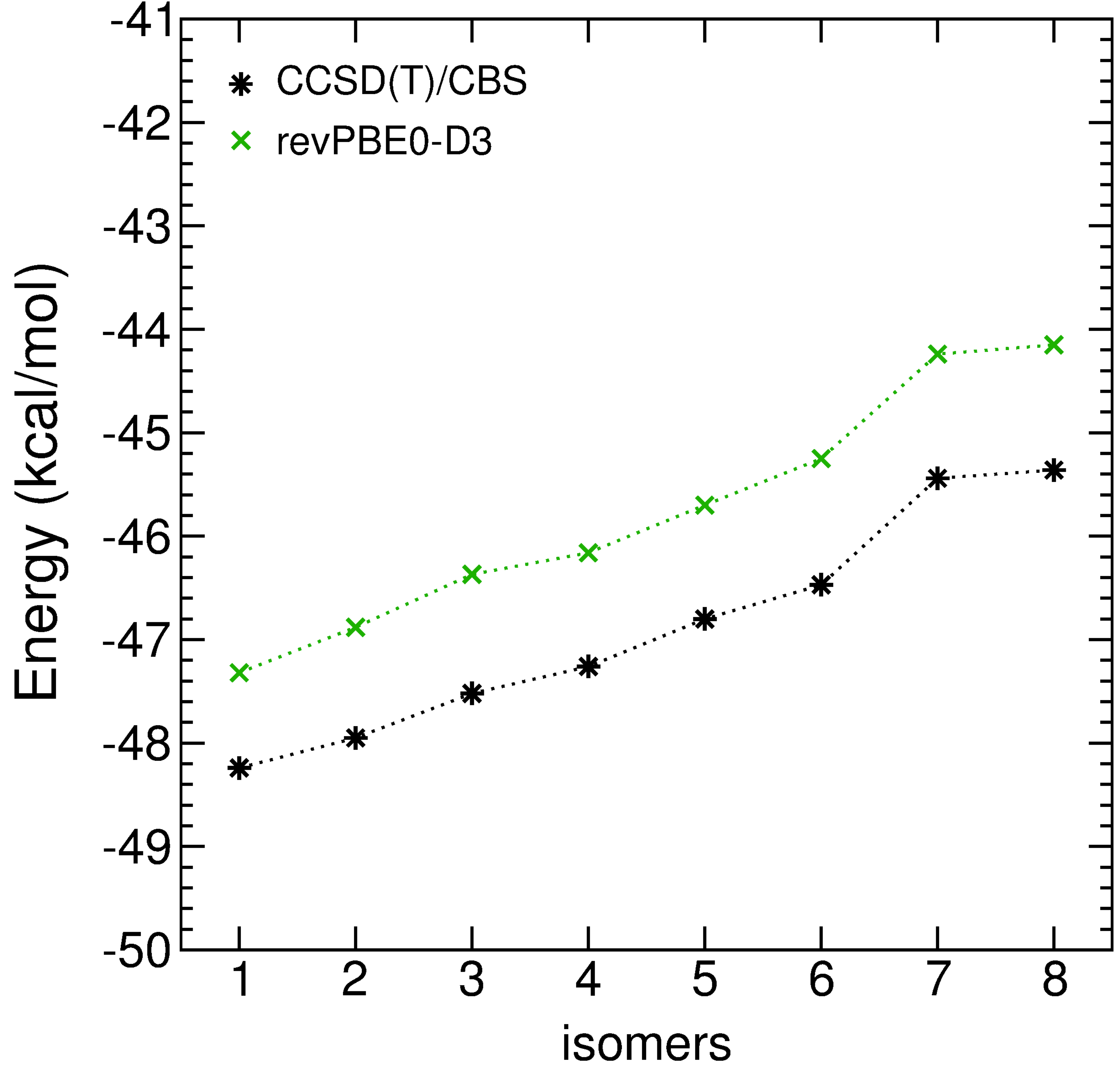

isomer 3


isomer 7

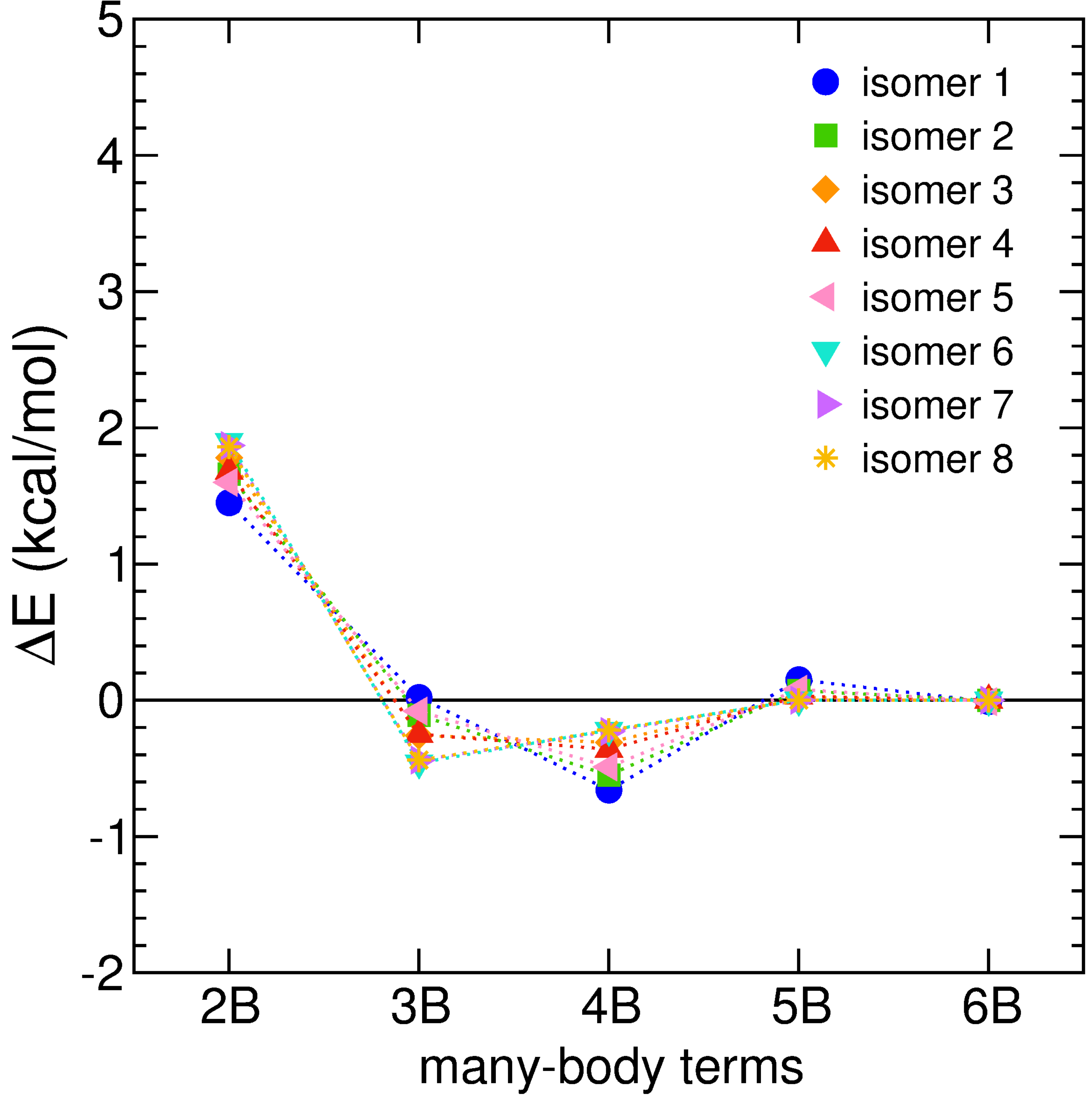

isomer 4


isomer 8

interaction energies



many-body contributions



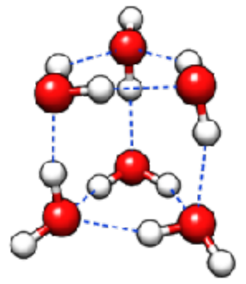
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

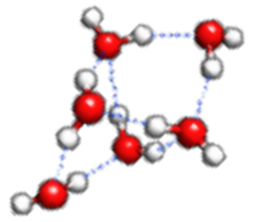
Many-Body Effects in Water

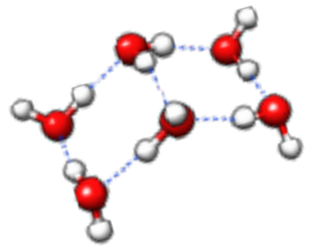
Climbing Jacob’s ladder of DFT approximations

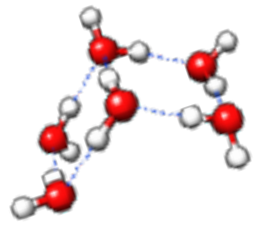
Range-separated, hybrid: ωB97M-V

water hexamer

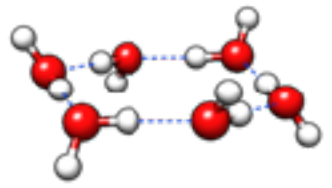

isomer 1

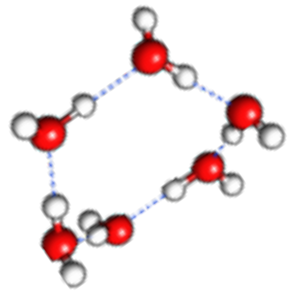

isomer 2

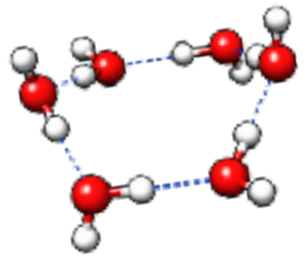

isomer 3


isomer 4

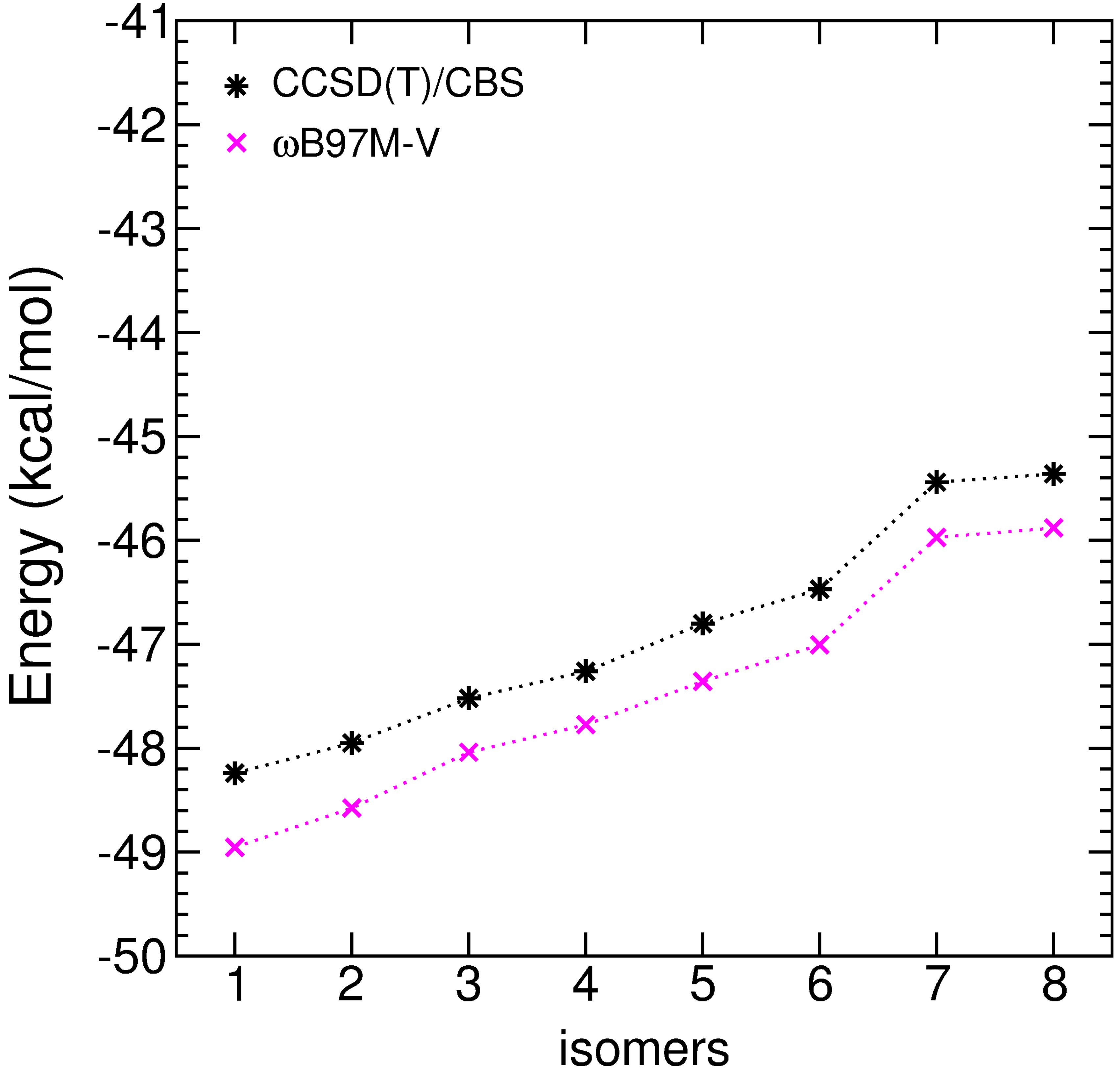

isomer 5


isomer 6

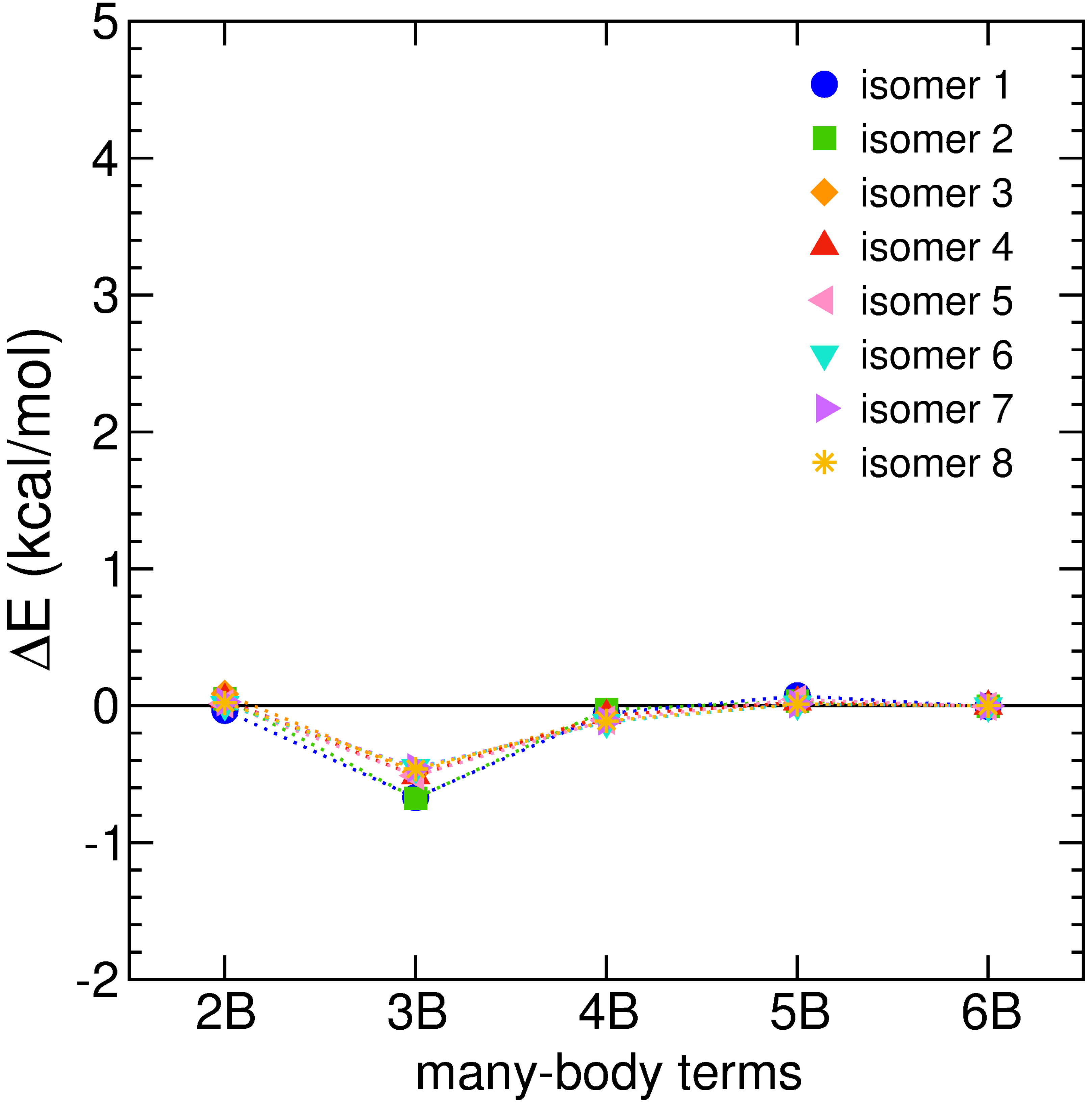

isomer 7


isomer 8

interaction energies



many-body contributions



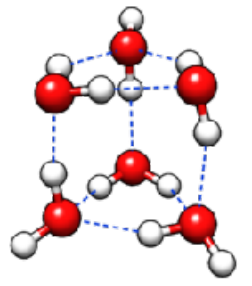
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

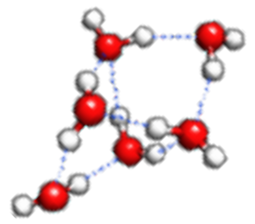
Many-Body Effects in Water

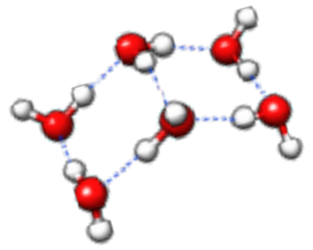
Climbing Jacob's ladder of molecular models

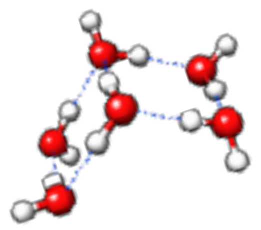
MB-pol

water hexamer

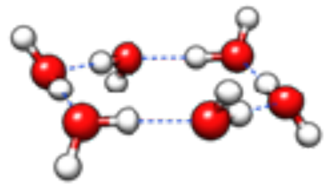

isomer 1

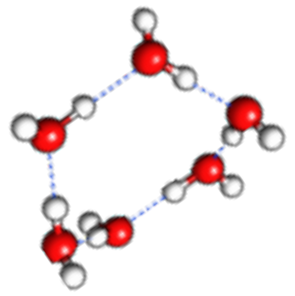

isomer 2

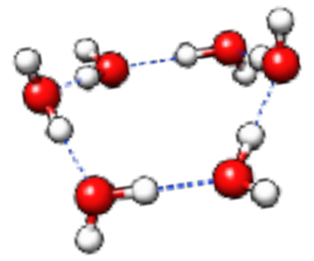

isomer 3


isomer 4

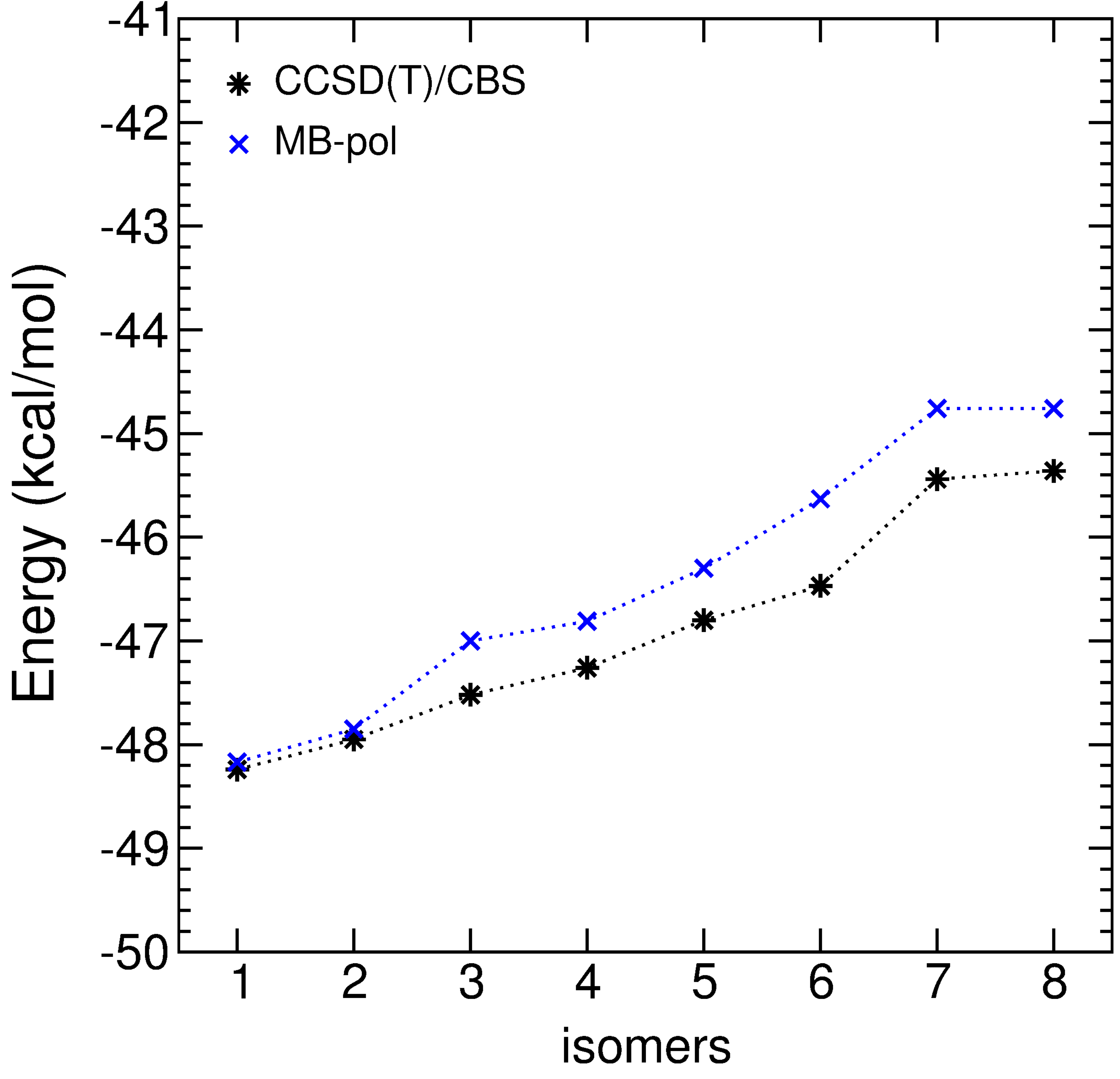

isomer 5


isomer 6

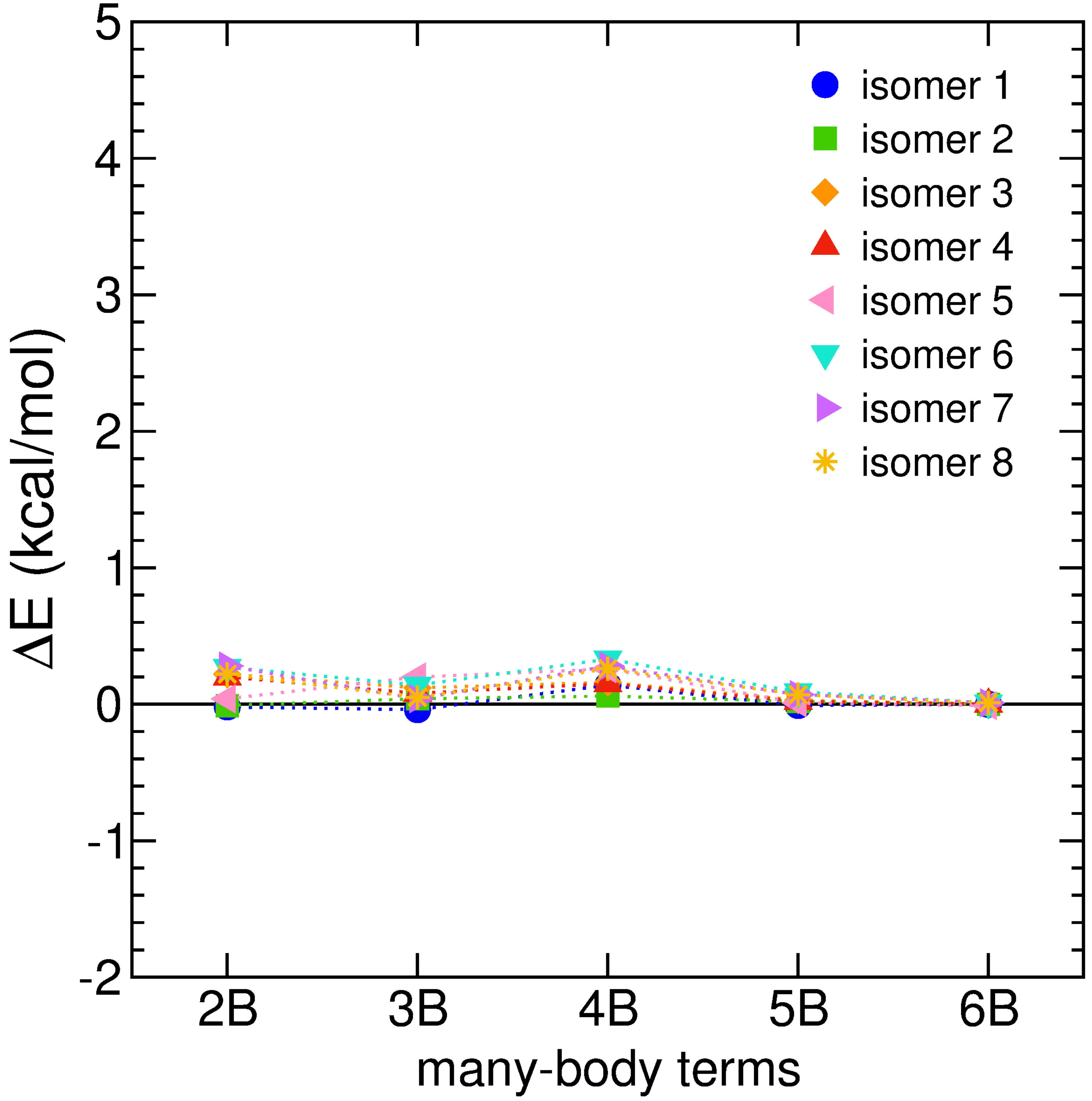

isomer 7


isomer 8

interaction energies



many-body contributions



Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. *10*, *8211 (2019)*

MB-pol: CCSD(T) Accuracy at the Cost of a Force Field

$$E_N = \boxed{\sum_{i}^{N} V^{1B}(i) \qquad \text{1-body}}$$

$$+ \sum_{i<j}^{N} V^{2B}(i,j) \qquad \text{2-body}$$

$$+ \sum_{i<j<k}^{N} V^{3B}(i,j,k) \qquad \text{3-body}$$

$$+ \ldots \qquad \text{4-body} \quad \cdots$$

$$+ V^{NB}(1,\ldots,N) \qquad \text{N-body}$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

flexible monomers

N-body polarization

Explicit 2B and 3B corrections from CCSD(T)/CBS

- *J. Phys. Chem. Lett. 3, 3765 (2012)*
- *J. Chem. Theory Comput. 9, 1103 (2013)*
- *J. Chem. Theory Comput. 9, 5395 (2013)*
- *J. Chem. Theory Comput. 10, 1599 (2014)*
- *J. Chem. Theory Comput. 10, 2906 (2014)*

FP, Acc. Chem. Res. 49, 1844 (2016)

MB-DFT: Probing Many-Body Effects in Water

$$
\begin{aligned}
E_N = & \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\
& + \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\
& + \sum_{i<j<k}^{N} V^{3B}(i,j,k) && \text{3-body} \\
& + \ldots && \text{4-body} \\
& && \cdots \\
& + V^{NB}(1,\ldots,N) && \text{N-body}
\end{aligned}
$$

flexible monomers

N-body polarization

Explicit 2B and 3B corrections from DFT

3 MB PEFs for each XC functional
- (2B+3B)-XC

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

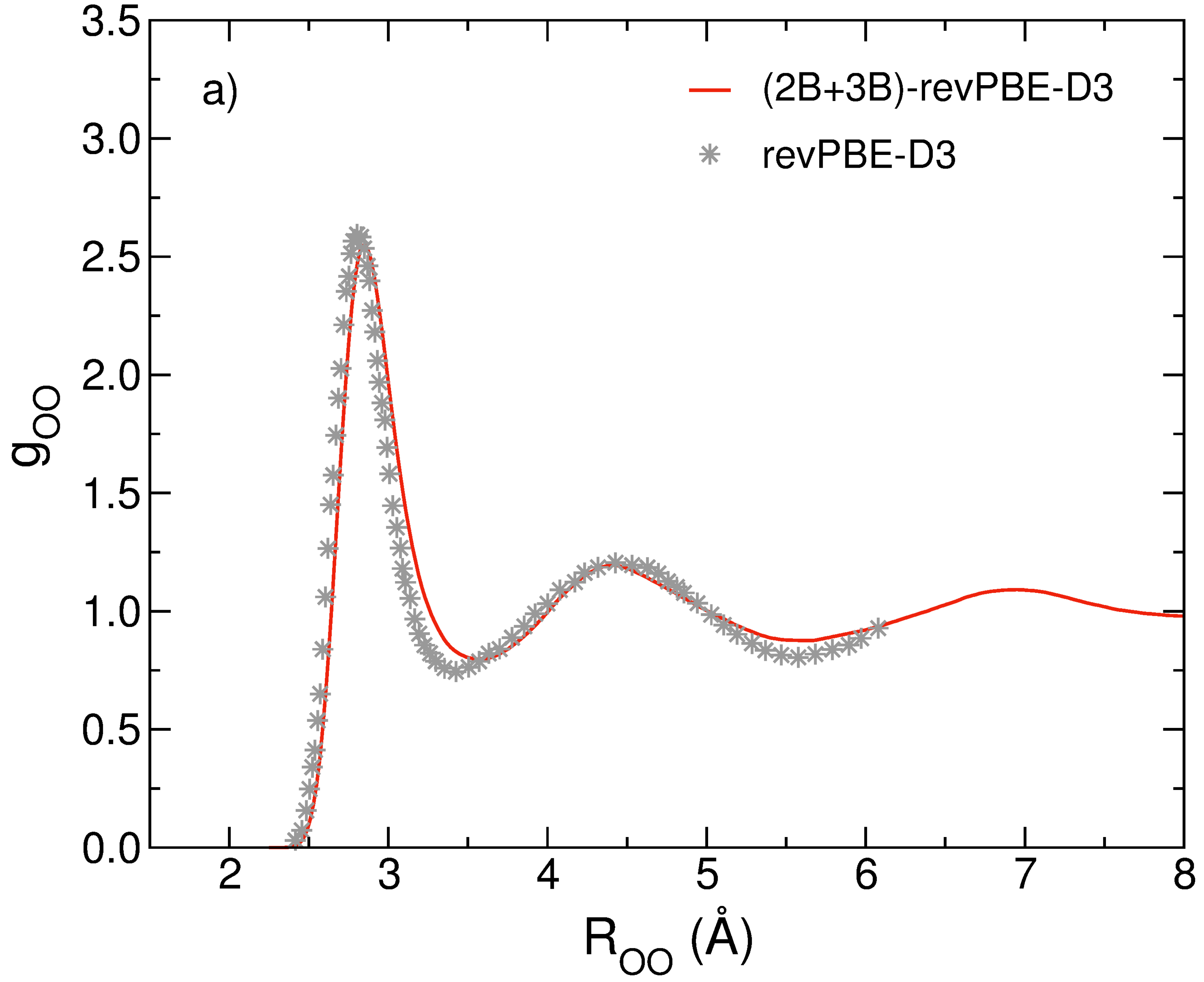
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT a.k.a. “DFT à la Carte”: MB Models with DFT Accuracy

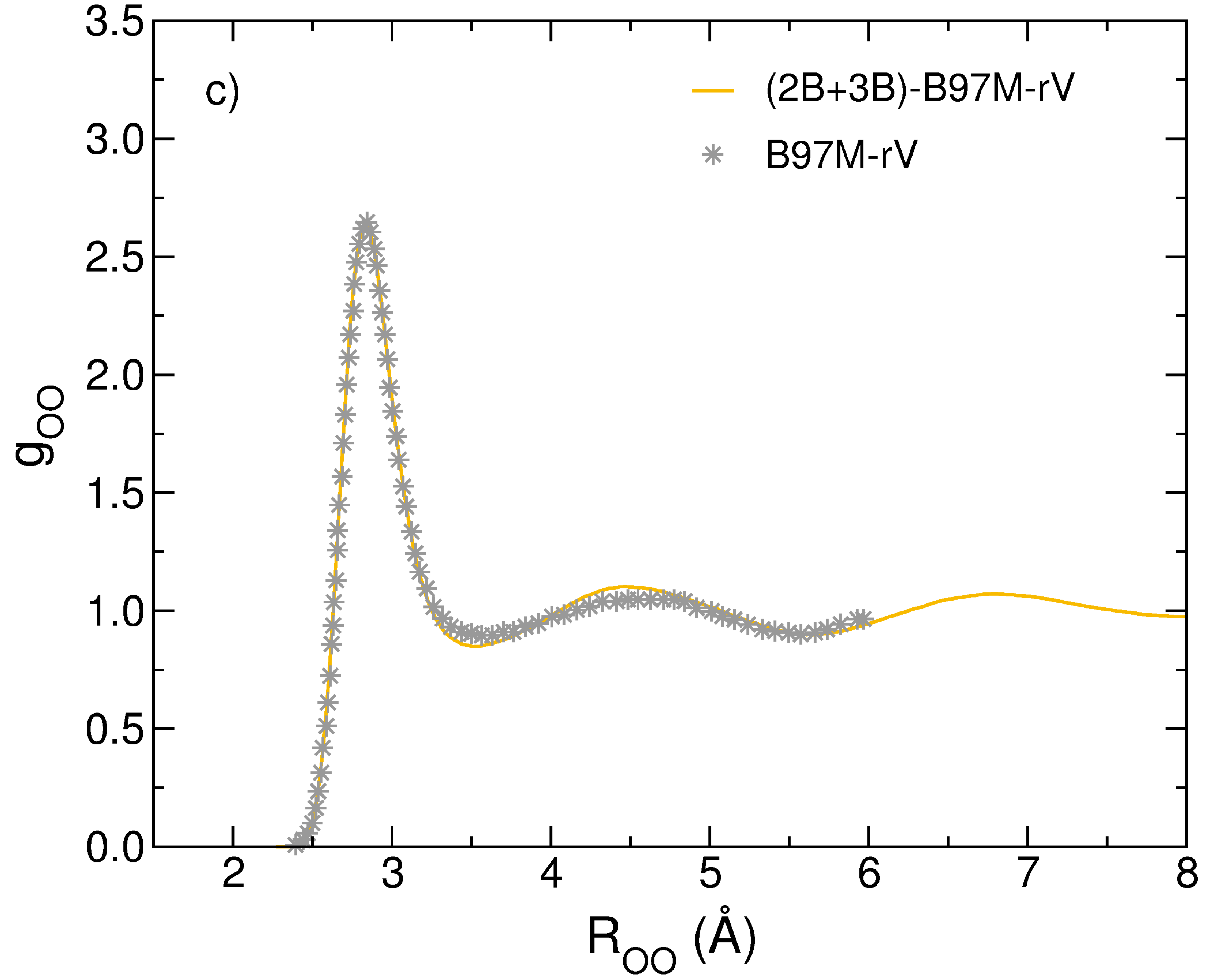
MB-DFT vs. DFT

256 molecules vs. 64 molecules

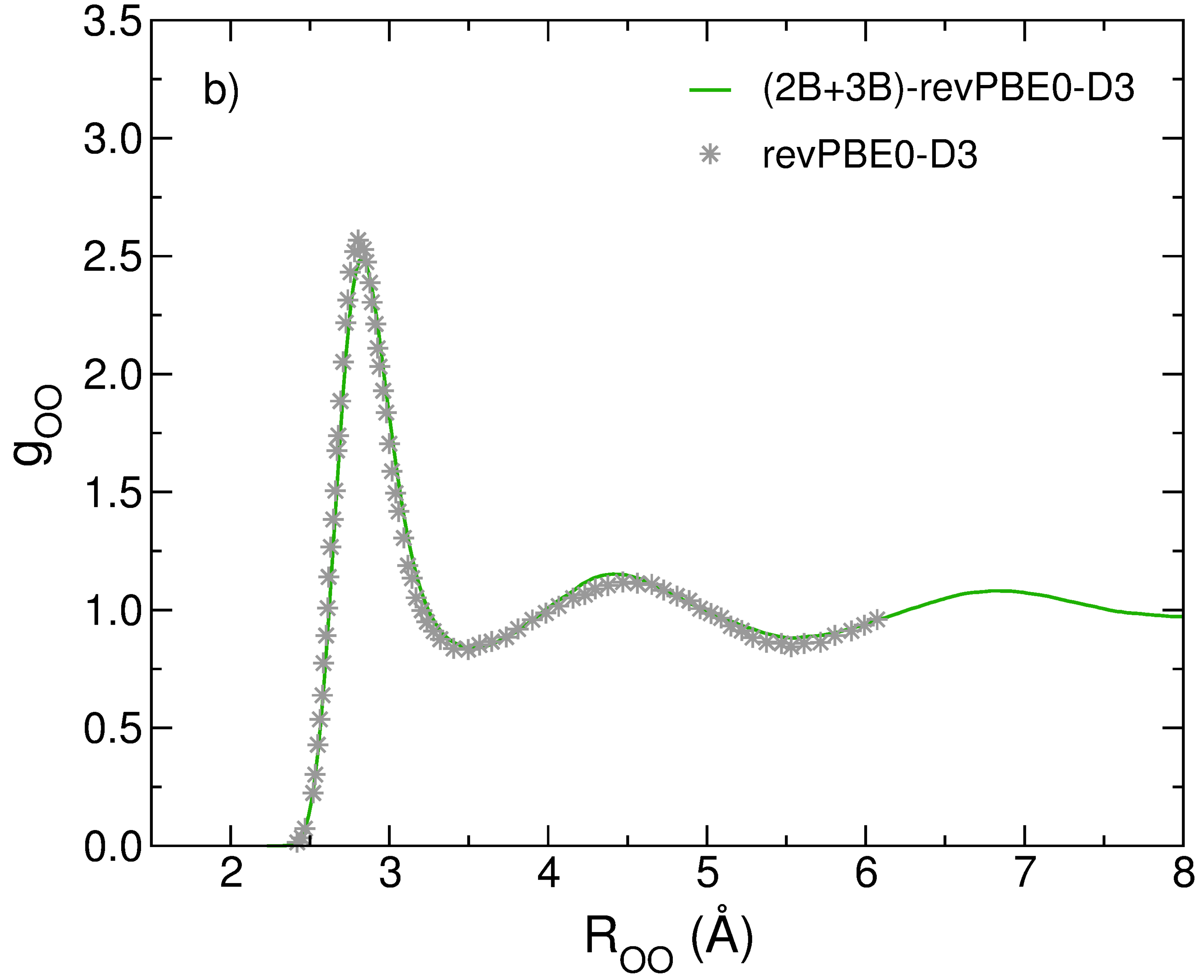
revPBE-D3



B97M-rV



revPBE0-D3



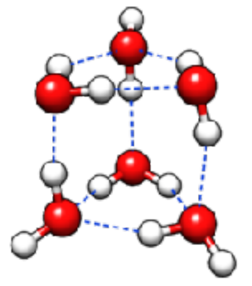
✷ *PIMD DFT: Marsalek & Markland, J. Phys. Chem. Lett. 8, 1545 (2017)*

Pestana, Marsalek, Markland & Head-Gordon, J. Phys. Chem. Lett. 9, 5009 (2018)

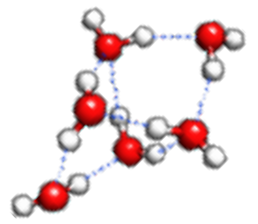
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

Many-Body Effects in Water

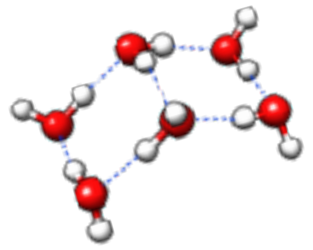
Climbing Jacob's ladder of DFT approximations



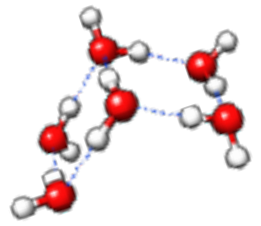
isomer 1



isomer 2

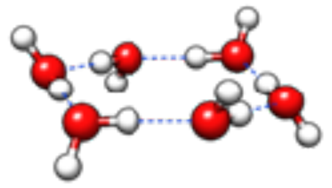


isomer 3

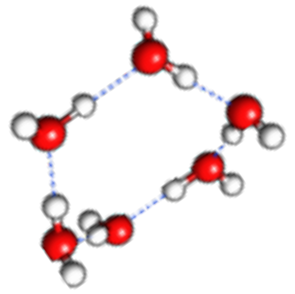


isomer 4

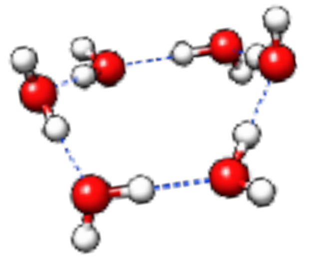
isomer 5



isomer 6



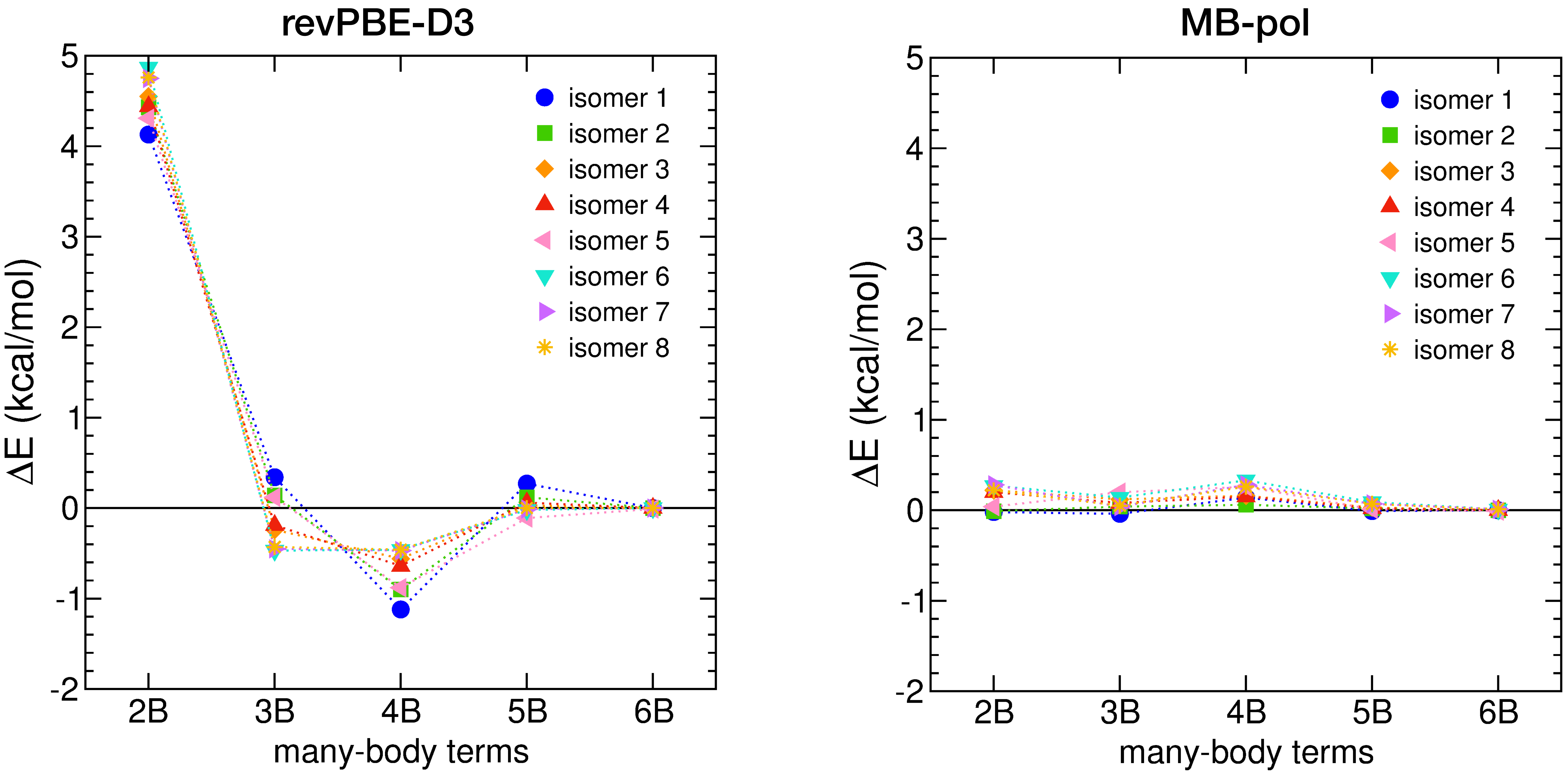
isomer 7



isomer 8

many-body contributions

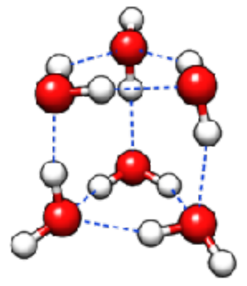
water hexamer

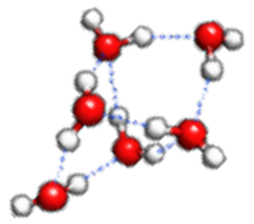


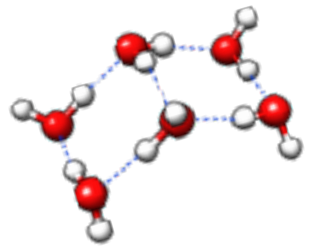
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

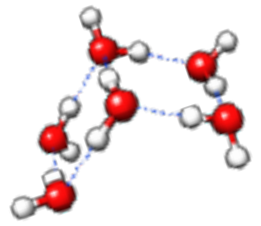
Many-Body Effects in Water

Climbing Jacob’s ladder of DFT approximations

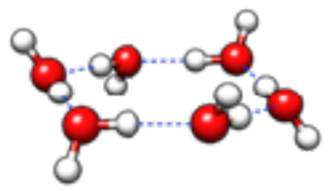

isomer 1

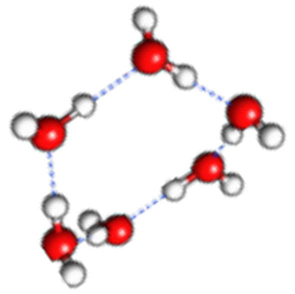

isomer 2

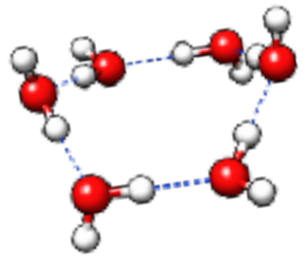

isomer 3


isomer 4


isomer 5

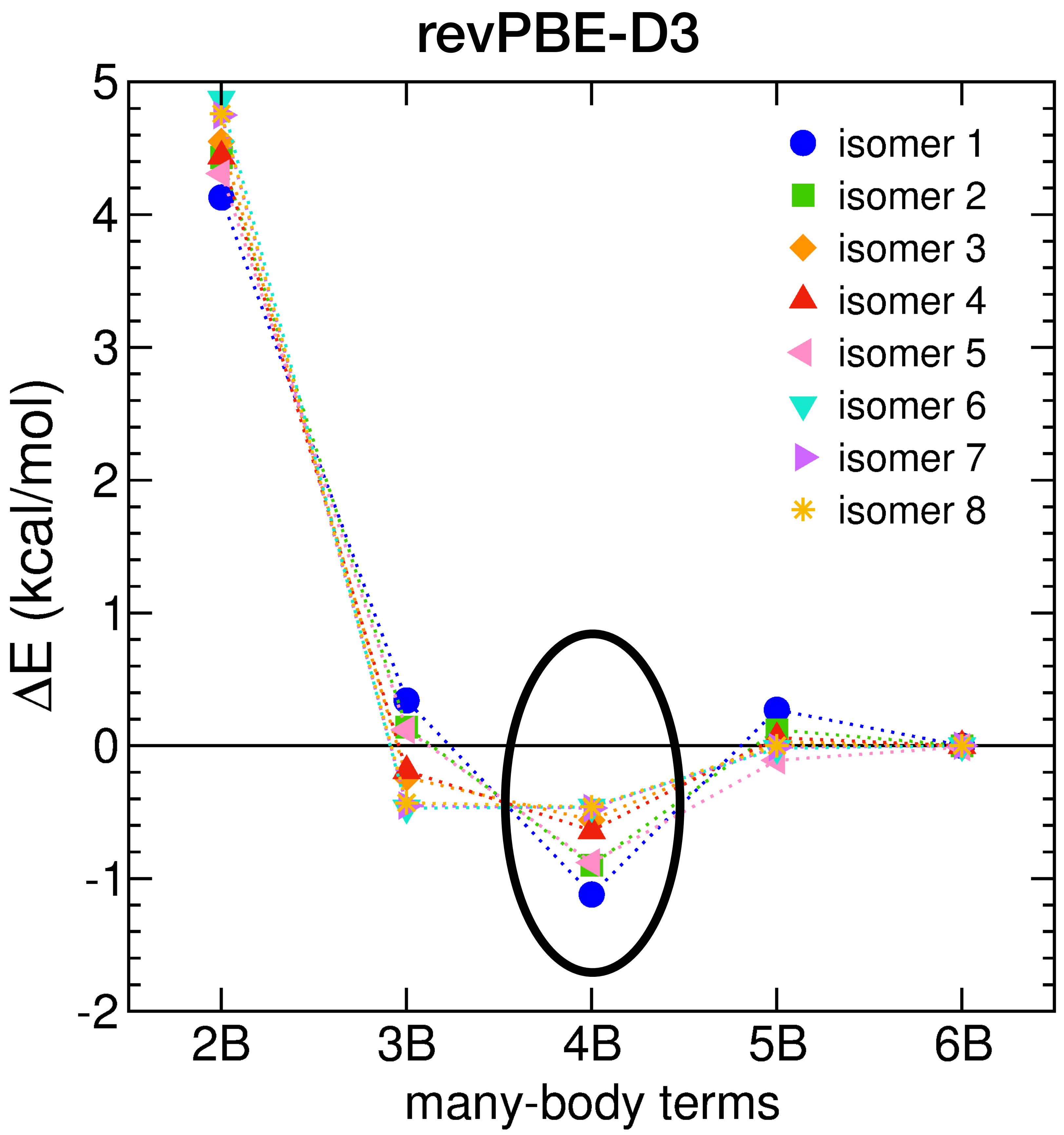

isomer 6

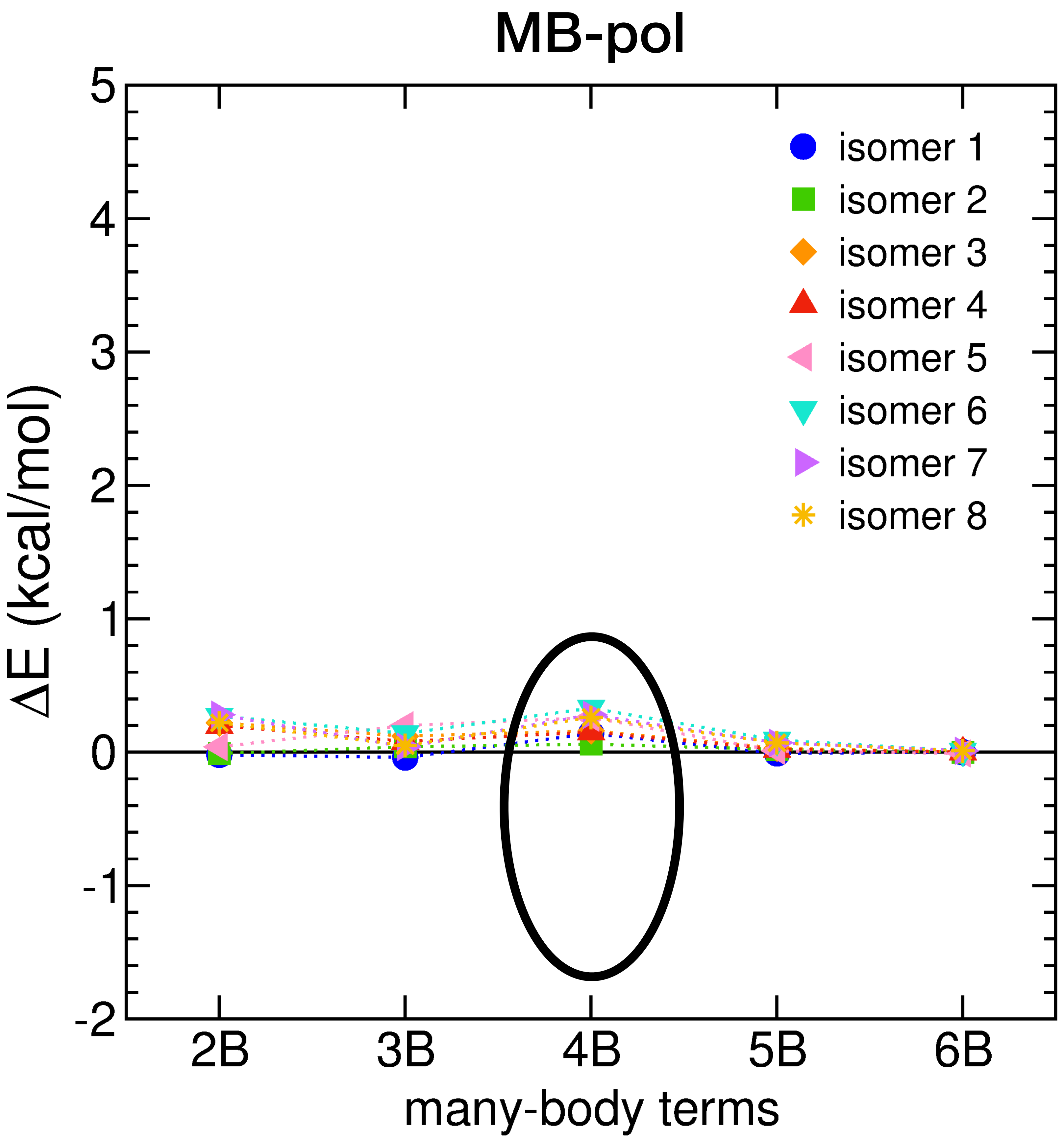

isomer 7


isomer 8

many-body contributions

water hexamer


revPBE-D3


MB-pol

Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

$$\begin{aligned} E_N = & \sum_{i}^{N} V^{1B}(i) && \text{1-body} \\ & + \sum_{i<j}^{N} V^{2B}(i,j) && \text{2-body} \\ & + \sum_{i<j<k}^{N} V^{3B}(i,j,k) && \text{3-body} \\ & + \ldots && \text{4-body} \\ & && \cdots \\ & + V^{NB}(1,\ldots,N) && \text{N-body} \end{aligned}$$

flexible monomers

N-body polarization

Explicit 2B and 3B corrections from DFT

3 MB PEFs for each XC functional
- (2B+3B)-XC

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

$$E_N = \sum_{i}^{N} V^{1B}(i)$$
1-body

$$+ \sum_{i<j}^{N} V^{2B}(i,j)$$
2-body

$$+ \sum_{i<j<k}^{N} V^{3B}(i,j,k)$$
3-body

$$+ \dots$$
4-body

…

$$+ V^{NB}(1,\dots,N)$$
N-body

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

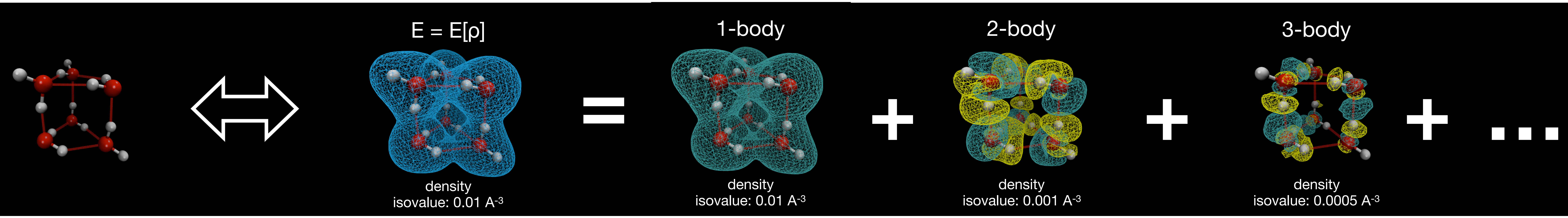
flexible monomers

N-body polarization

Explicit 2B corrections from DFT

Explicit 3B corrections from CCSD(T)/CBS

3 MB PEFs for each XC functional
- (2B+3B)-XC
- (2B)-XC



Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

$$E_N = \sum_{i}^{N} V^{1B}(i) \qquad \text{1-body}$$

$$+ \sum_{i<j}^{N} V^{2B}(i,j) \qquad \text{2-body}$$

$$+ \sum_{i<j<k}^{N} V^{3B}(i,j,k) \qquad \text{3-body}$$

$$+ \ldots \qquad \text{4-body}$$

$$\cdots$$

$$+ V^{NB}(1,\ldots,N) \qquad \text{N-body}$$

Hankins, Moskowitz & Stillinger, J. Chem. Phys. 53, 4544 (1970)

Popkie, Kistenmacher & Clementi, J. Chem. Phys. 59, 1325 (1973)

flexible monomers

N-body polarization

Explicit 2B corrections from CCSD(T)/CBS

Explicit 3B corrections from DFT

3 MB PEFs for each XC functional
- (2B+3B)-XC
- (2B)-XC
- (3B)-XC

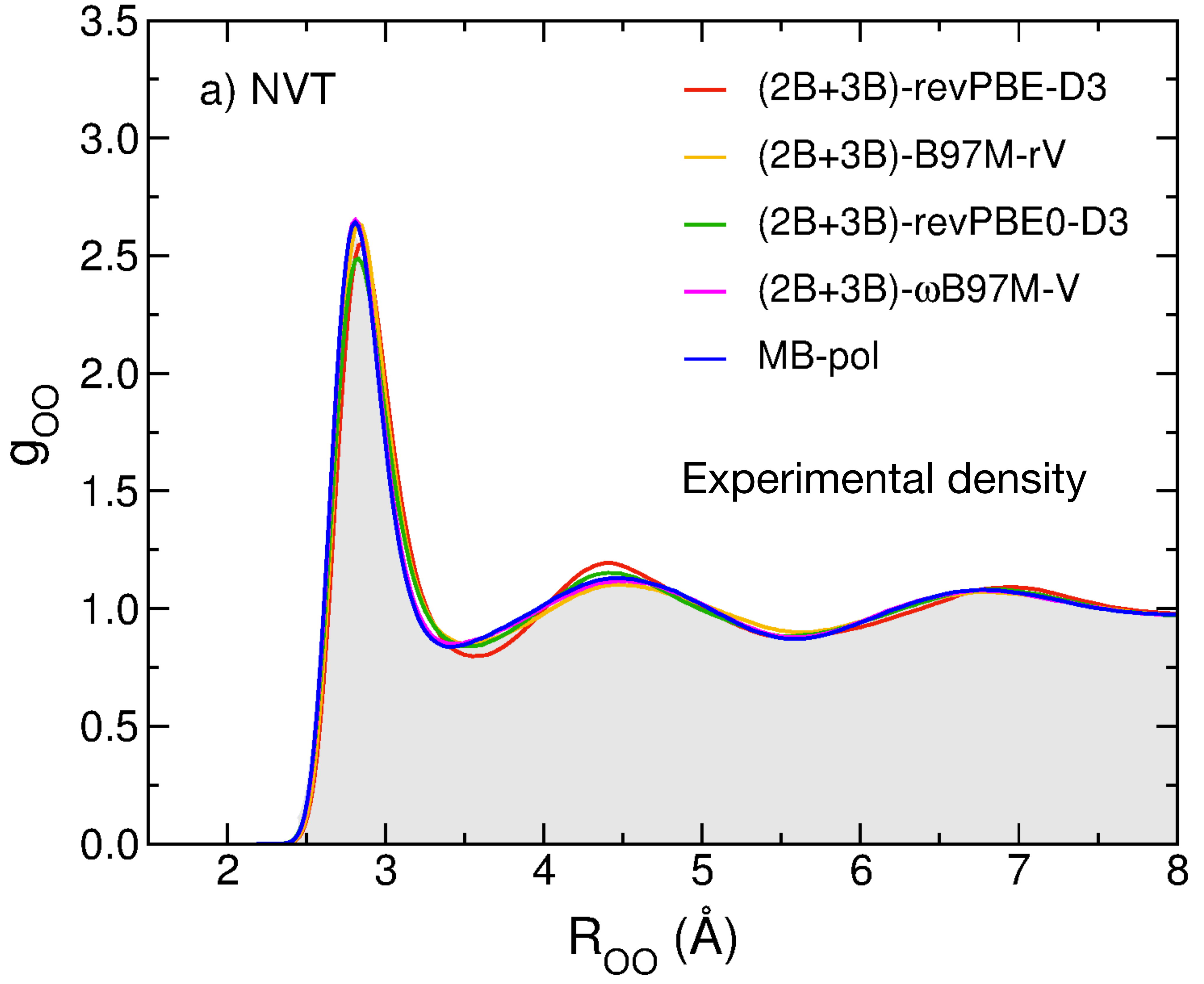
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

Climbing Jacob's ladder of DFT approximations

2B DFT + 3B DFT + NB MB-pol

water structure: O-O RDF



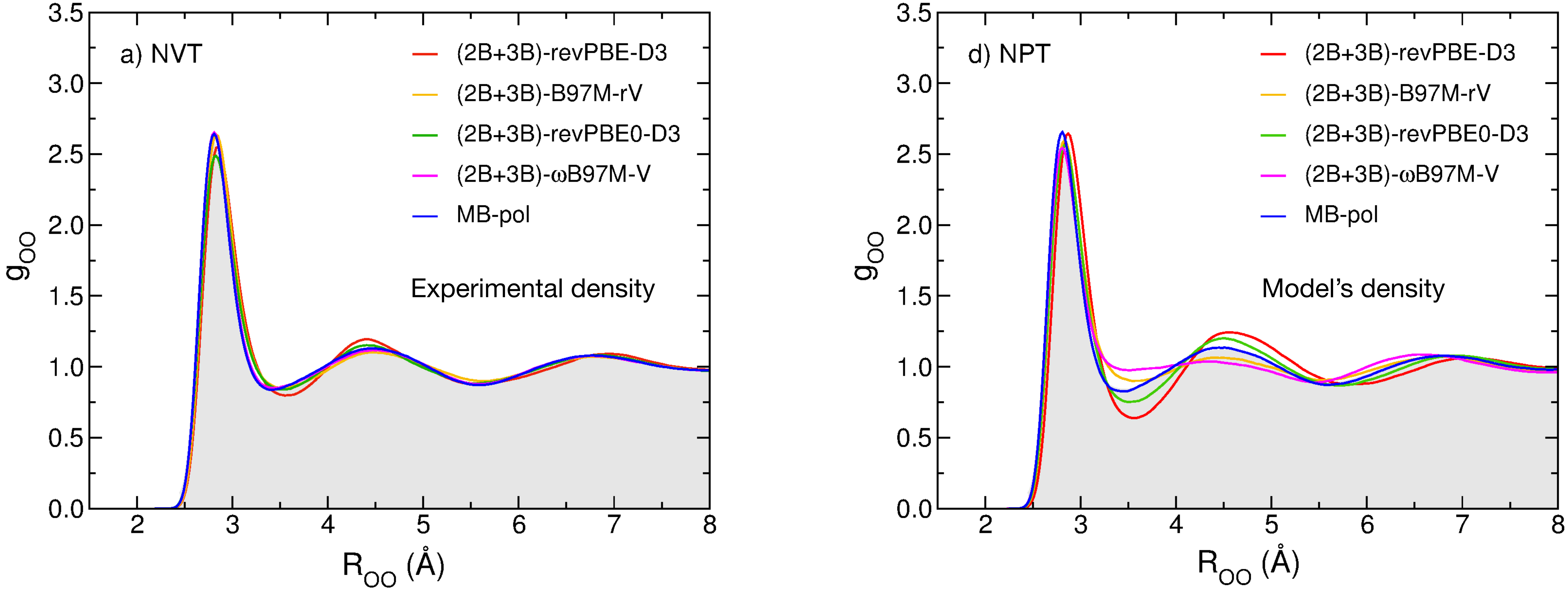
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

Climbing Jacob's ladder of DFT approximations

2B DFT + 3B DFT + NB MB-pol

water structure: O-O RDF



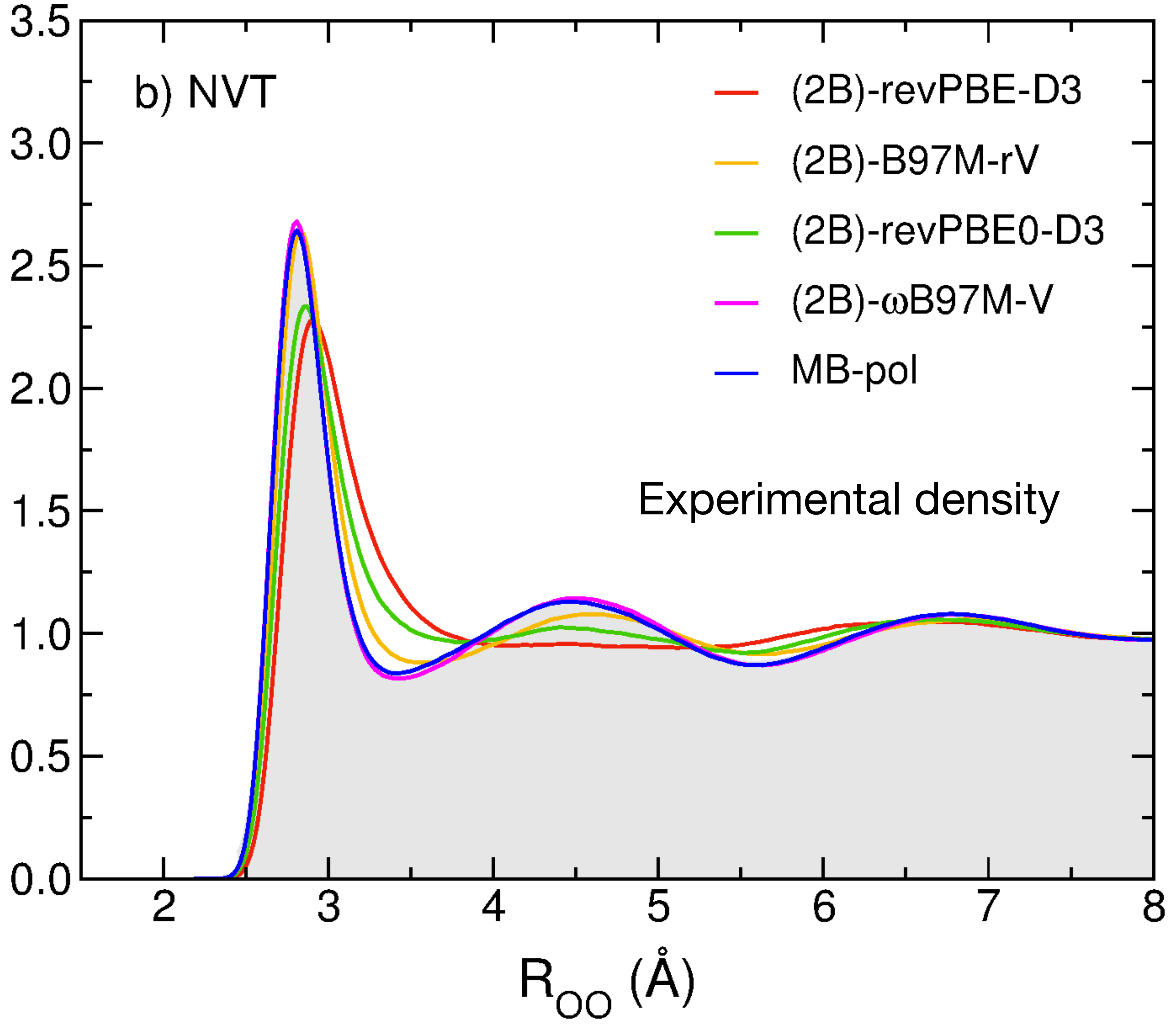
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

Climbing Jacob's ladder of DFT approximations

2B DFT + 3B CCSD(T) + NB MB-pol

water structure: O-O RDF



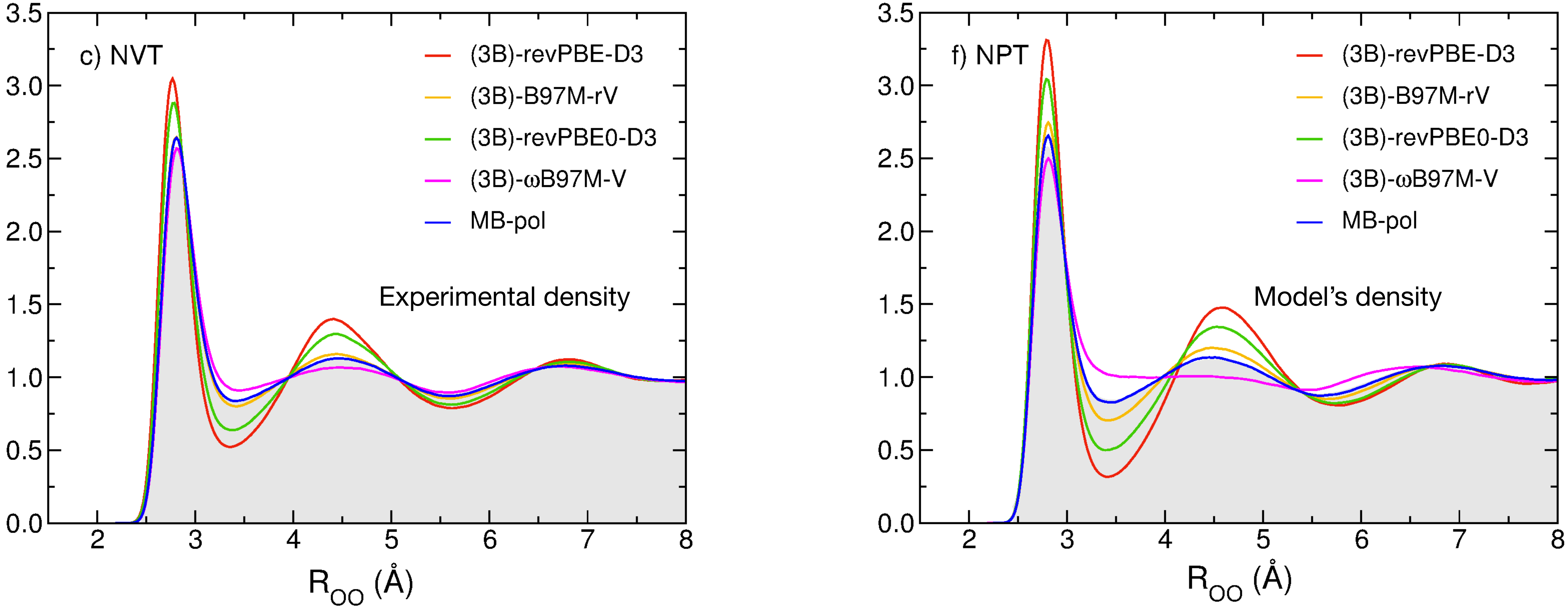
Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

Climbing Jacob's ladder of DFT approximations

2B CCSD(T) + 3B DFT + NB MB-pol

water structure: O-O RDF

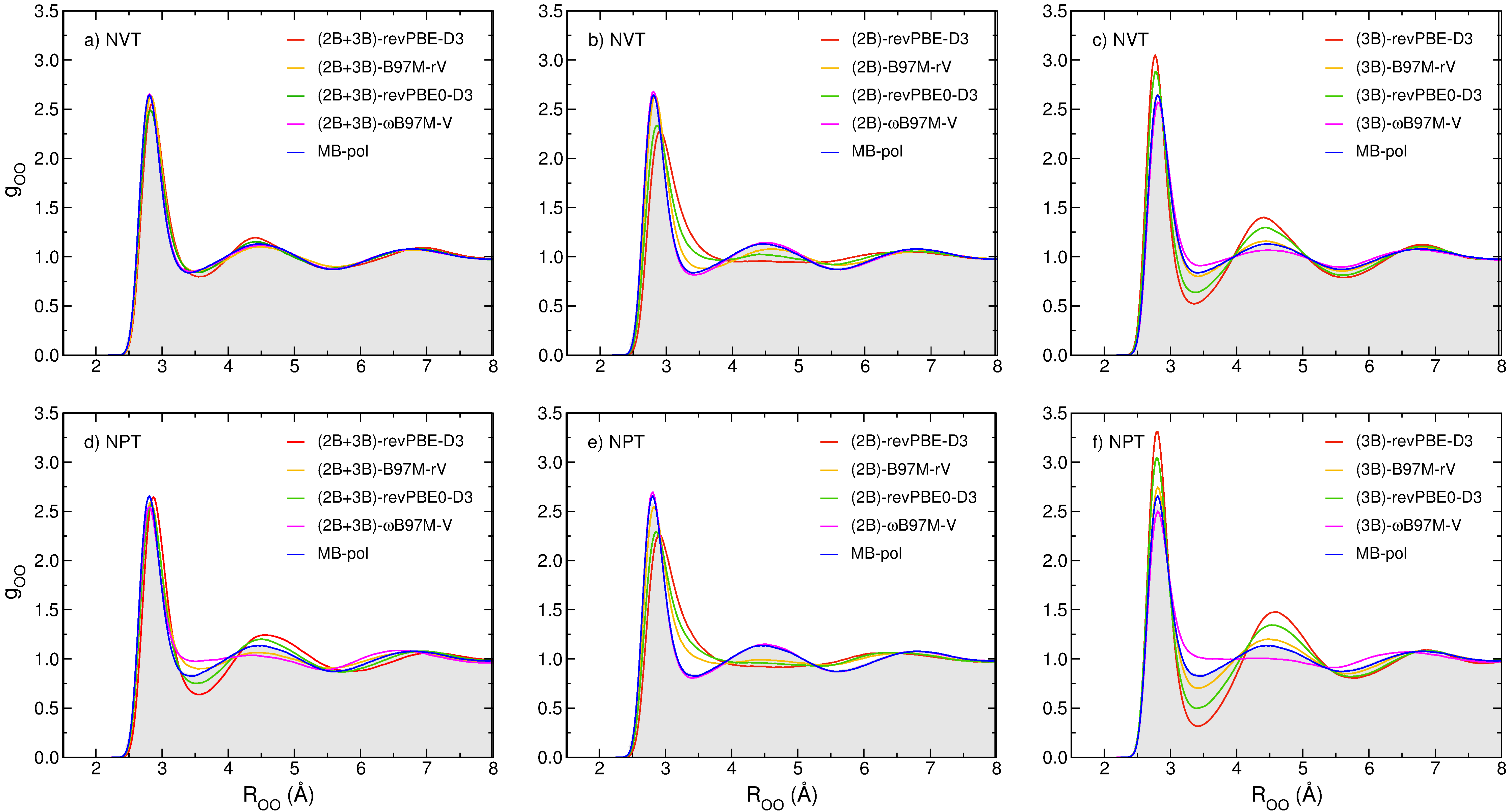


Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

Climbing Jacob's ladder of DFT approximations

water structure: O-O RDF

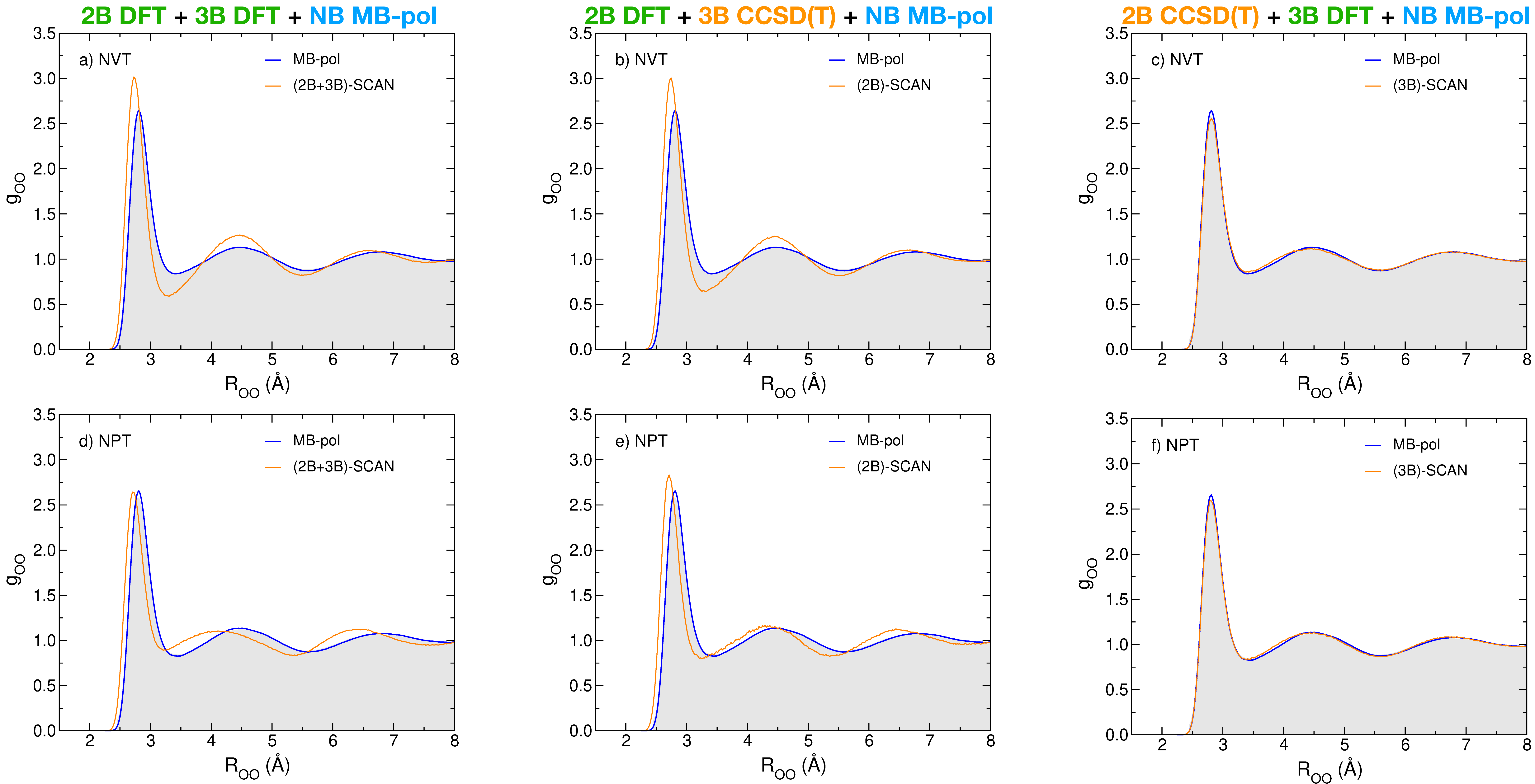


Riera, Lambros, Nguyen, Götz & FP, Chem. Sci. 10, 8211 (2019)

MB-DFT: Probing Many-Body Effects in Water

Climbing Jacob's ladder of DFT approximations: SCAN

water structure: O-O RDF



2B DFT + 3B DFT + NB MB-pol

2B DFT + 3B CCSD(T) + NB MB-pol

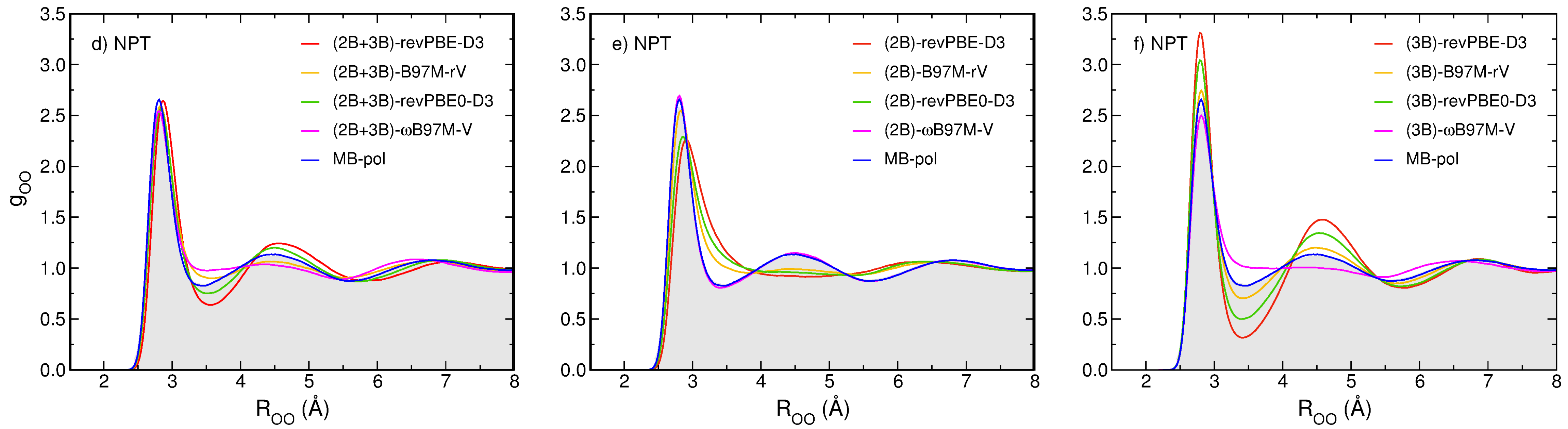
2B CCSD(T) + 3B DFT + NB MB-pol

Lambros, Hu & FP, in progress

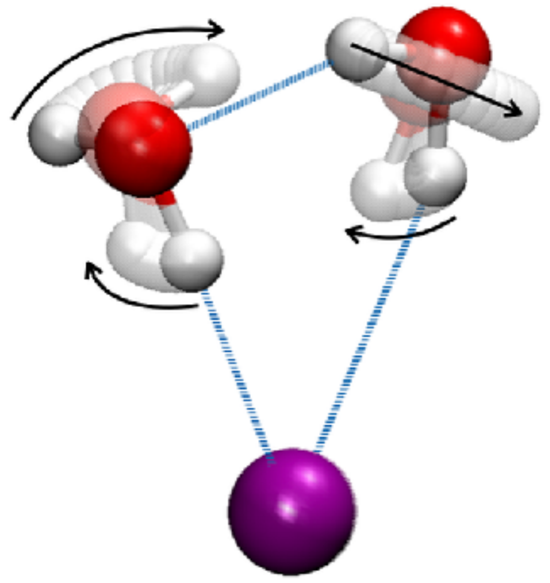
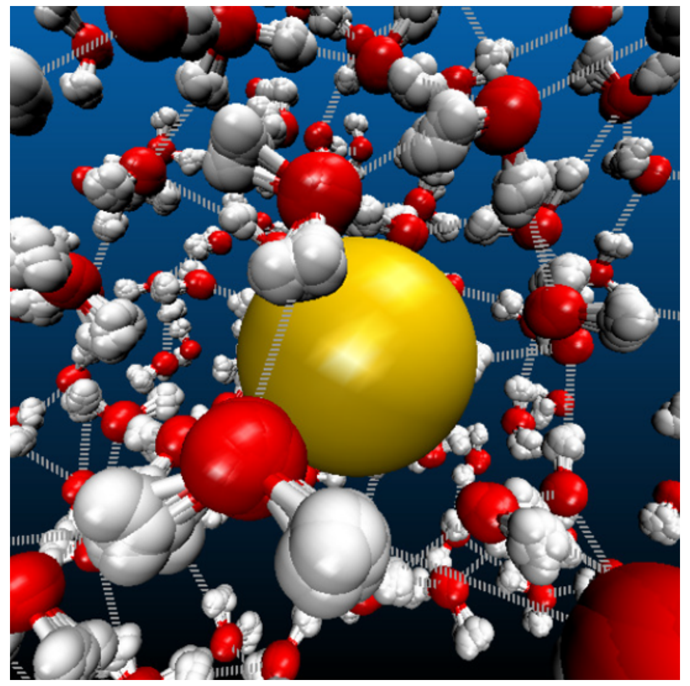
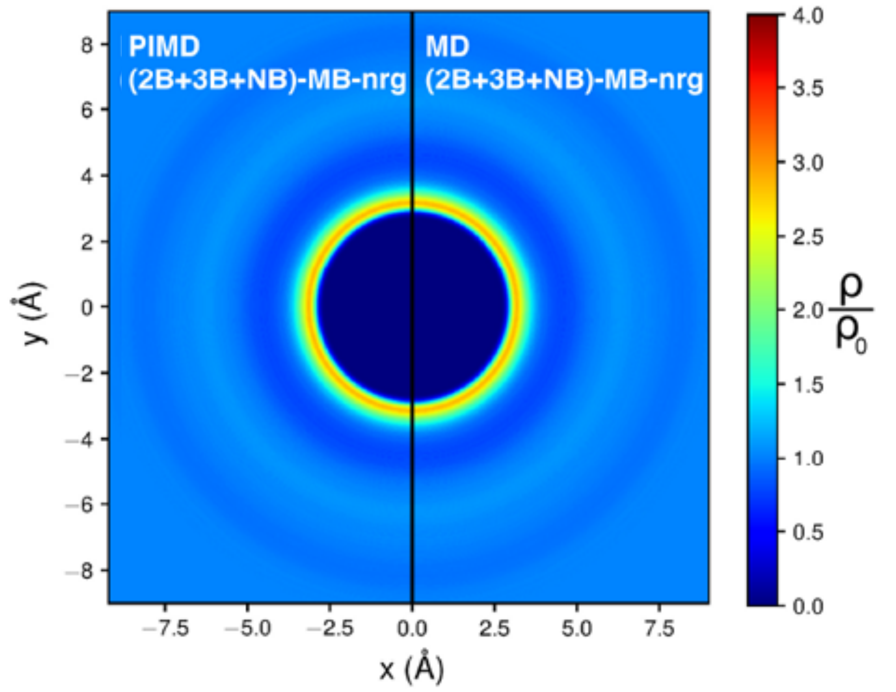
Learning Physics from Many-Body Models

What makes water water?

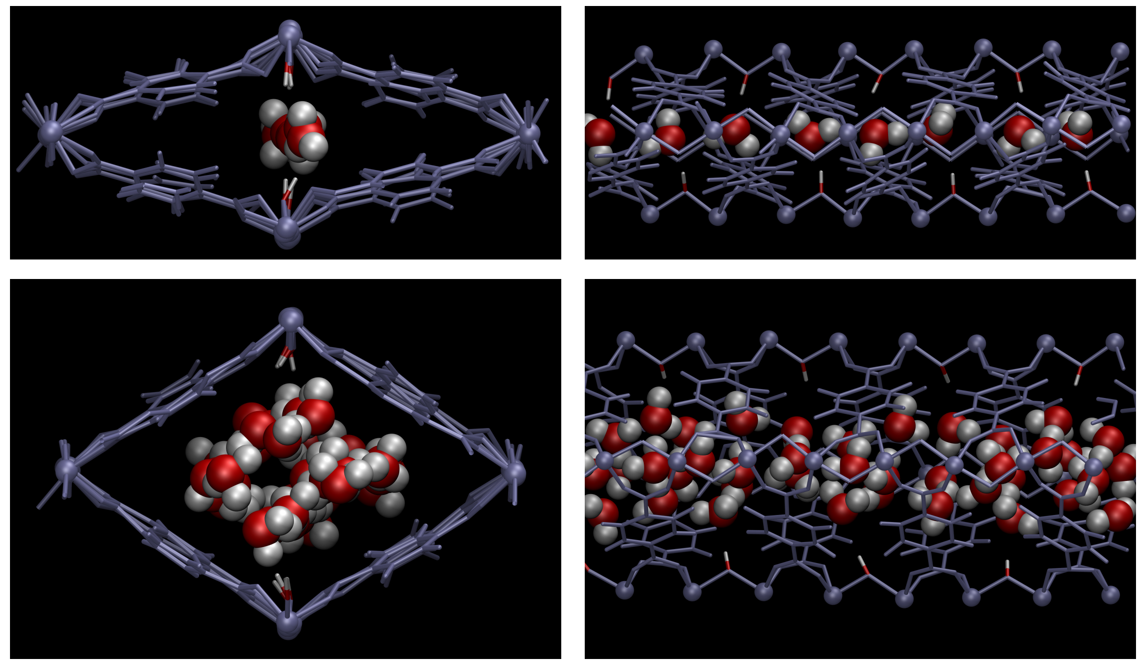
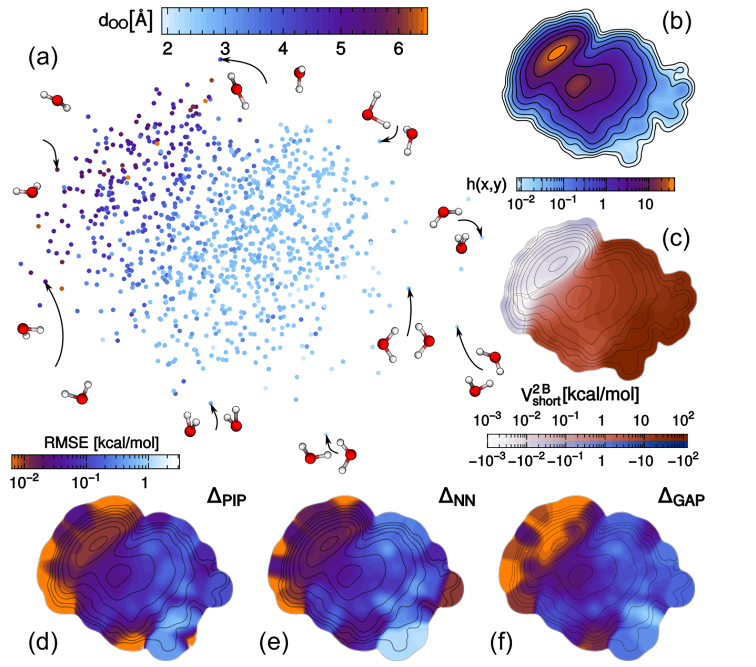
- Low-order many-body interactions determine the local structure of water
- Delocalization errors make current DFT models unreliable for water



How good are simulations of hydration?

Data-driven many-body models: Quo vadis?

http://paesanigroup.ucsd.edu

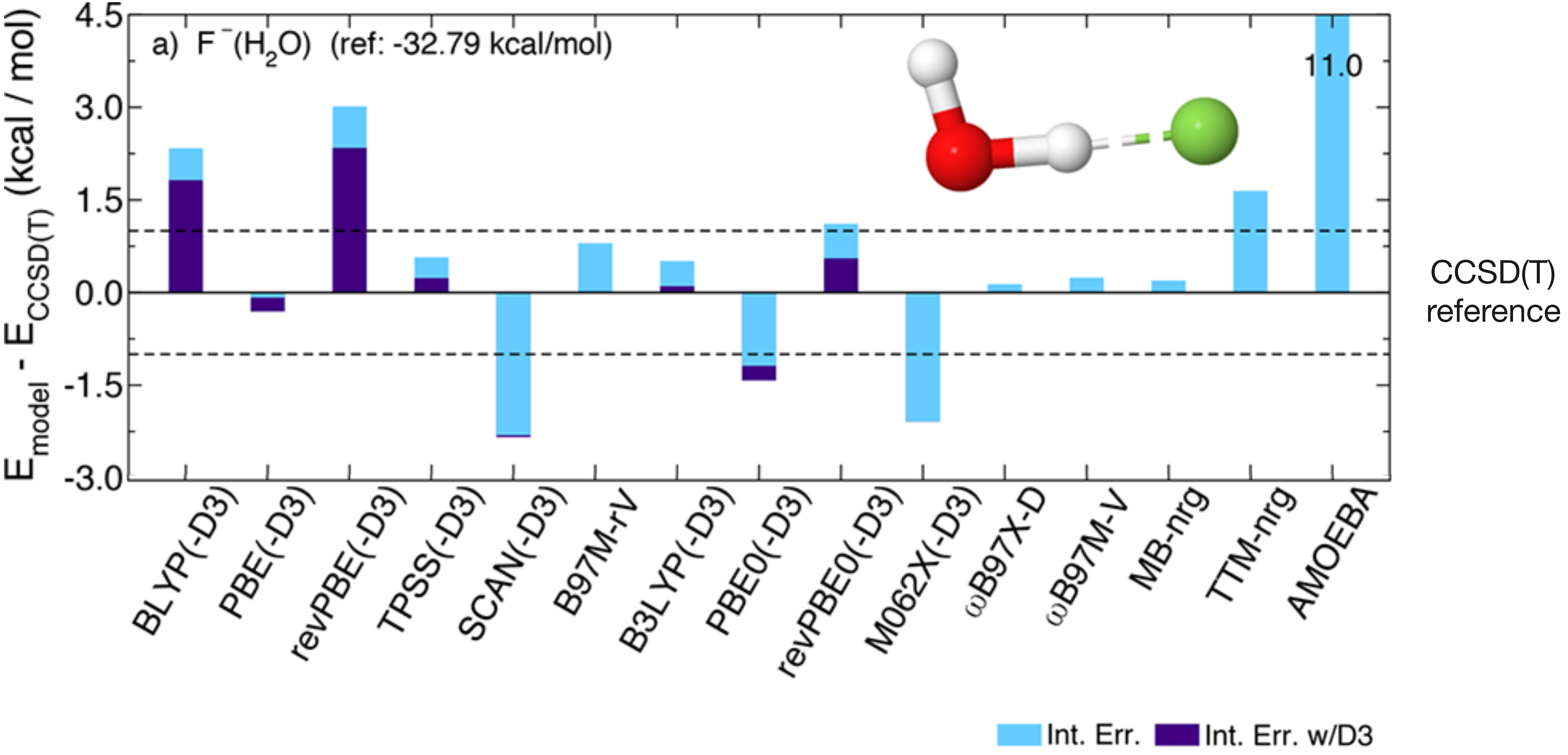
From Water to Complex Solutions

From Water to Complex Solutions

http://paesanigroup.ucsd.edu

Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

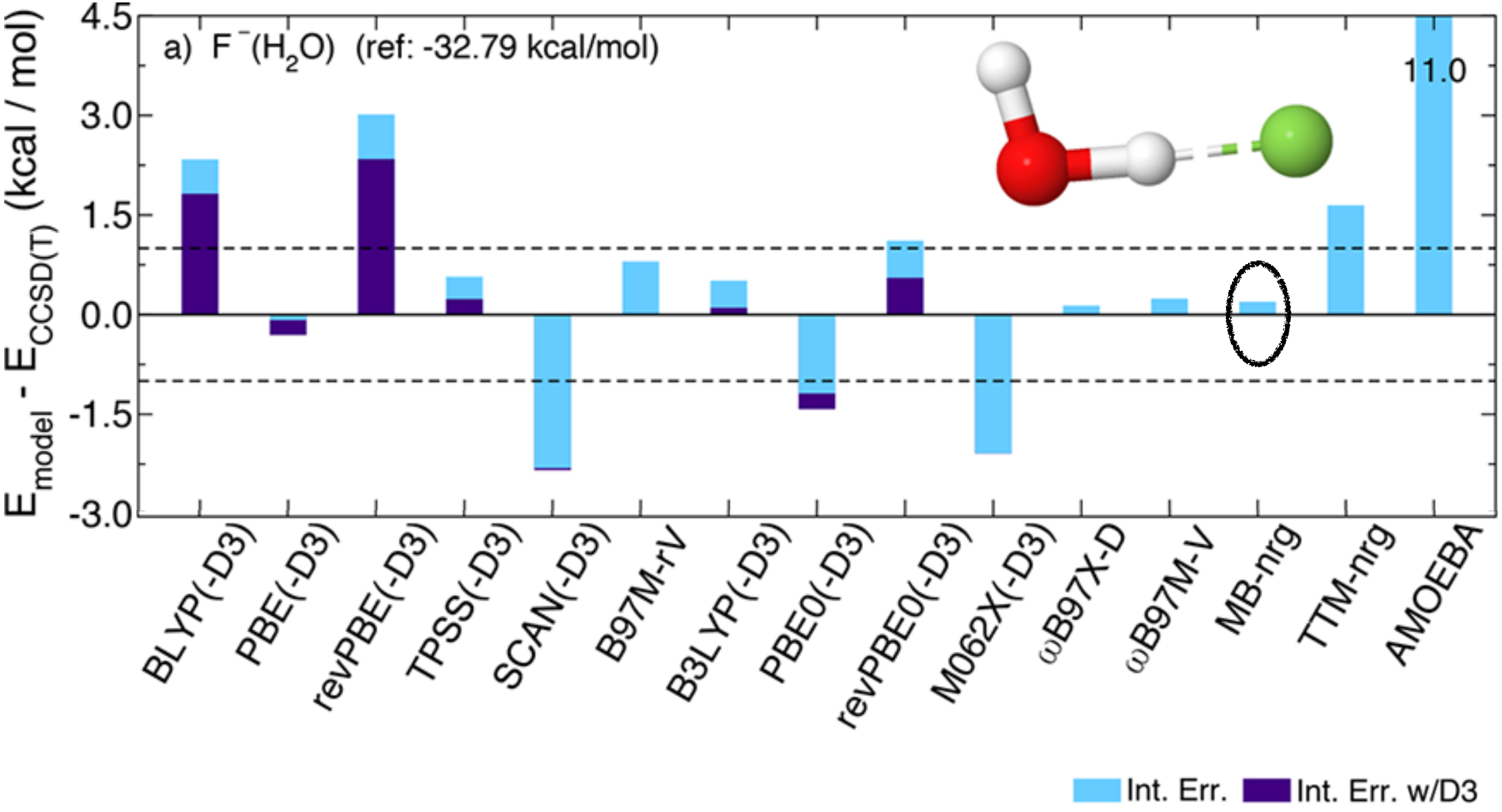
$F^-(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS



Bizzarro, Egan & FP, J. Chem. Theory Comput. *15*, *2983 (2019)*

Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

$F^-(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS



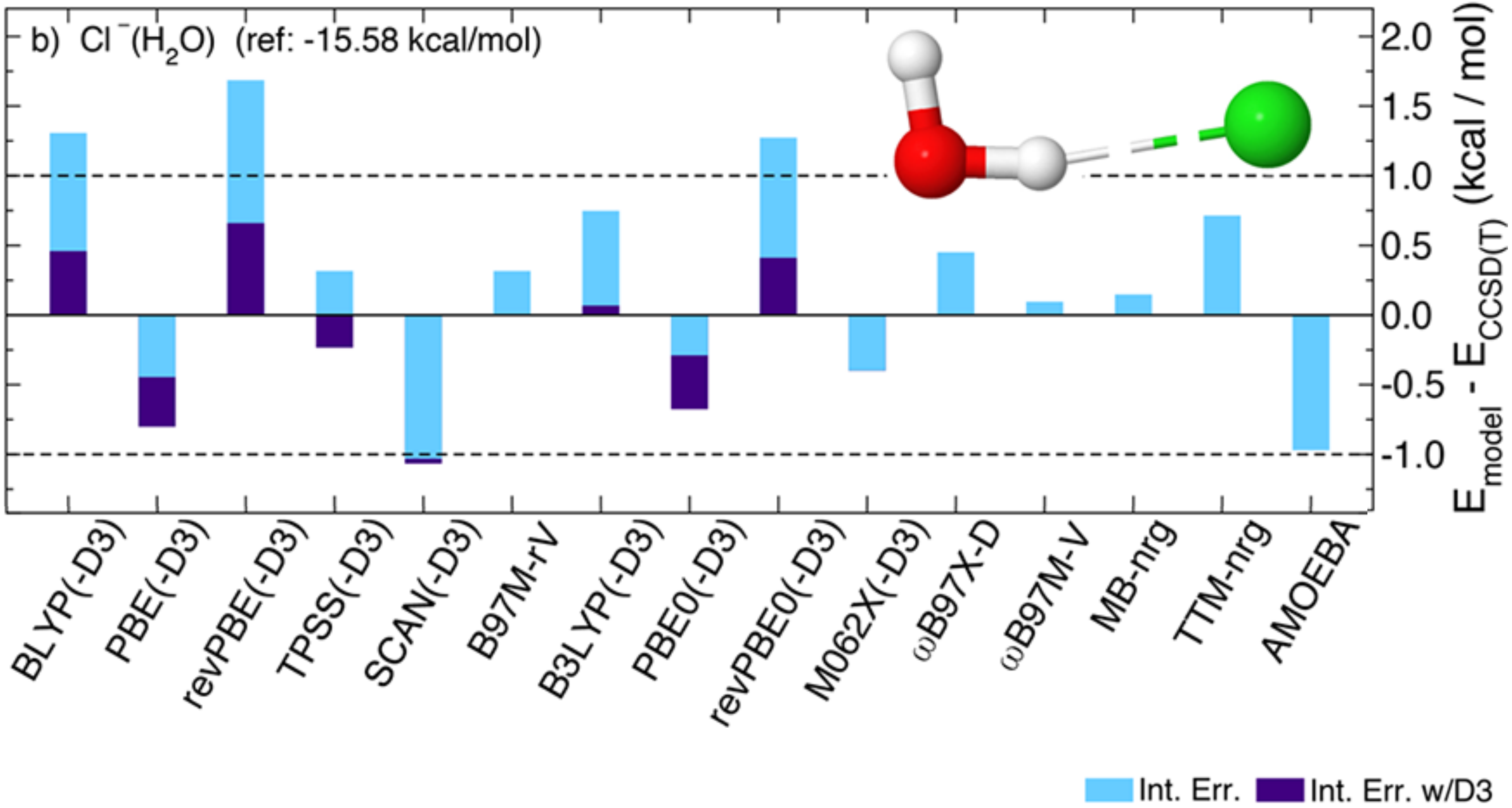
CCSD(T)
reference

Bizzarro, Egan & FP, J. Chem. Theory Comput. 15, 2983 (2019)

Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

$Cl^-(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS

CCSD(T) reference

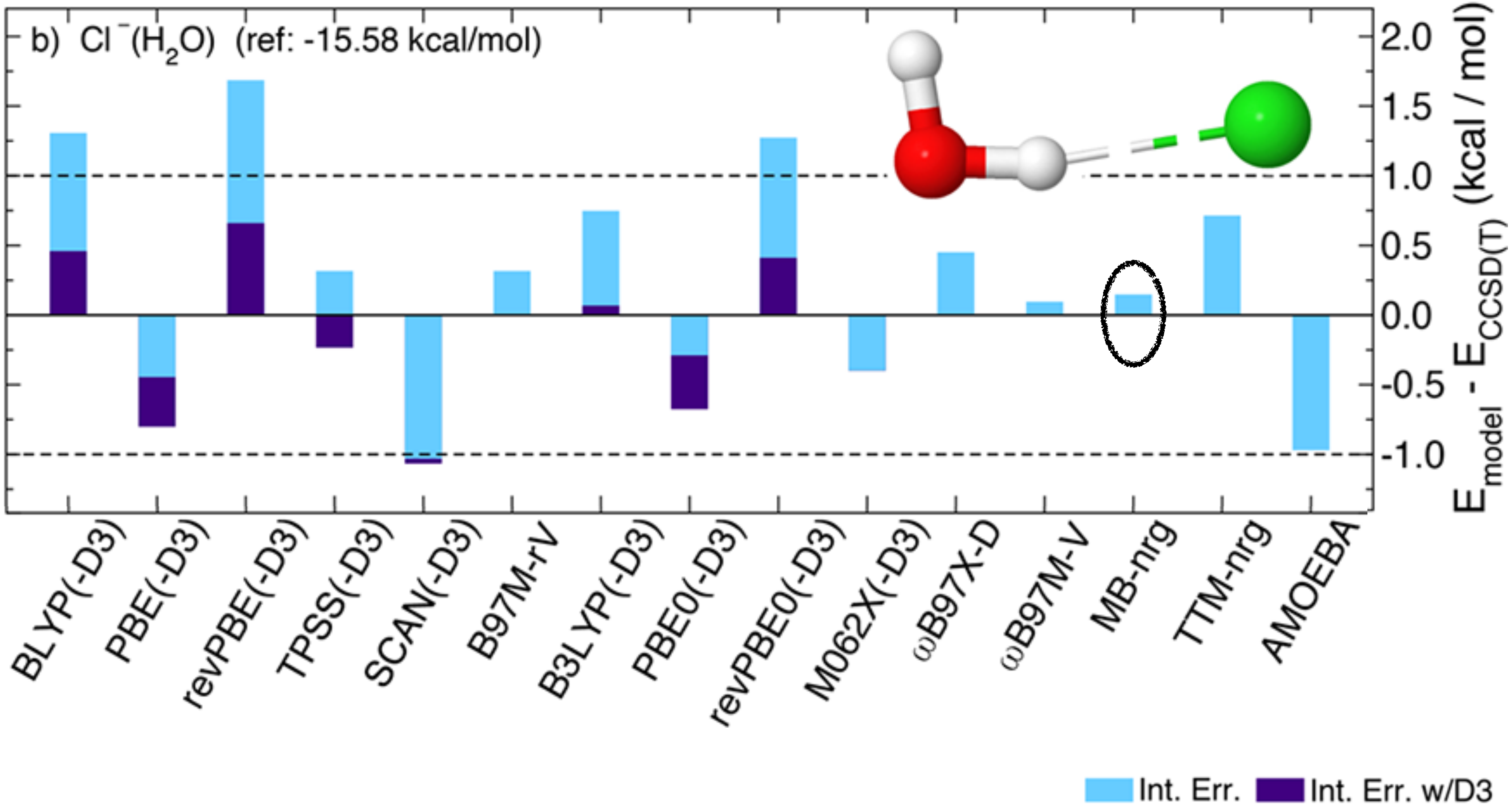


Bizzarro, Egan & FP, J. Chem. Theory Comput. 15, 2983 (2019)

Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

$Cl^-(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS

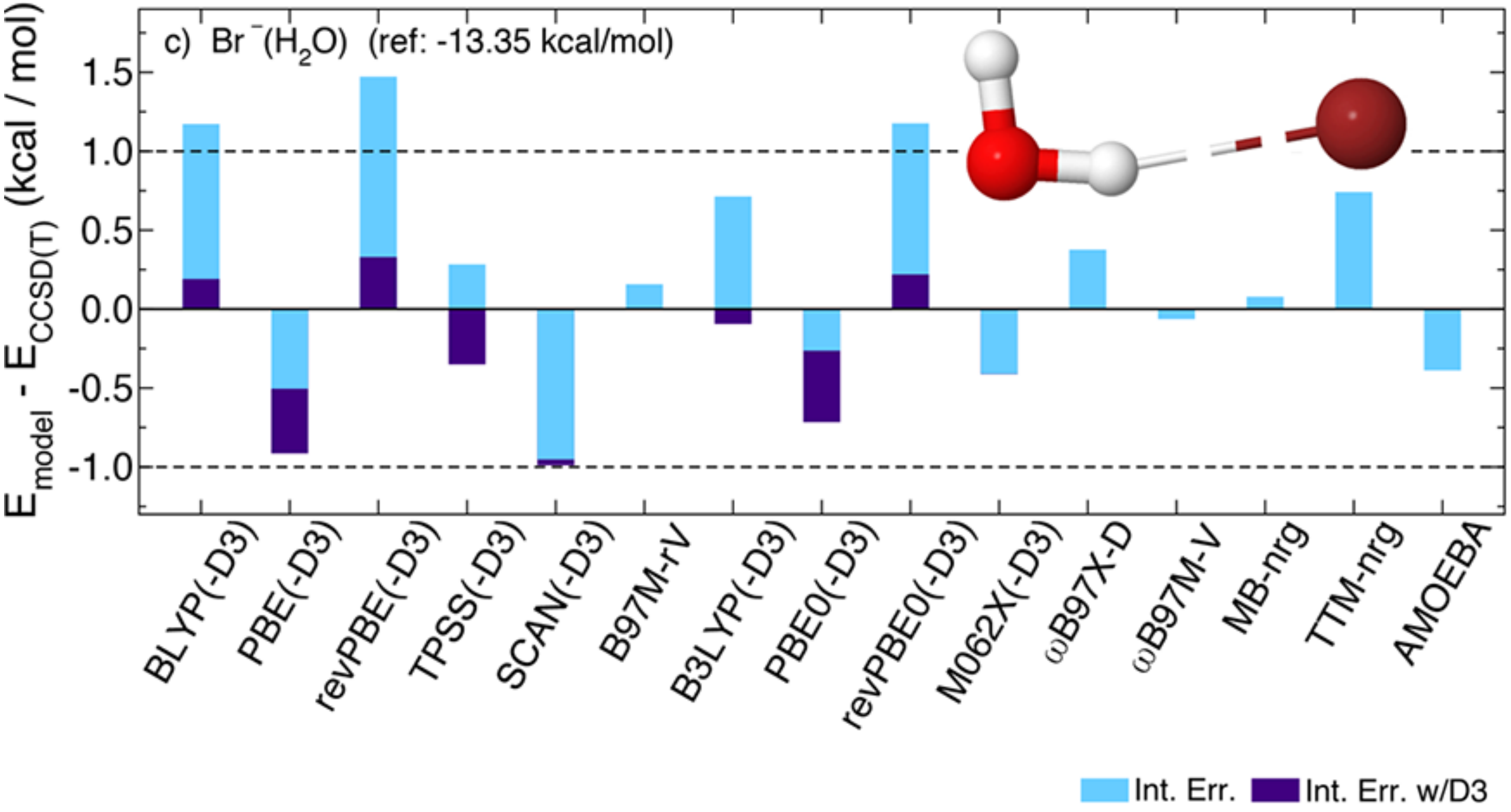
CCSD(T) reference



Bizzarro, Egan & FP, J. Chem. Theory Comput. 15, 2983 (2019)

Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

$Br^-(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS

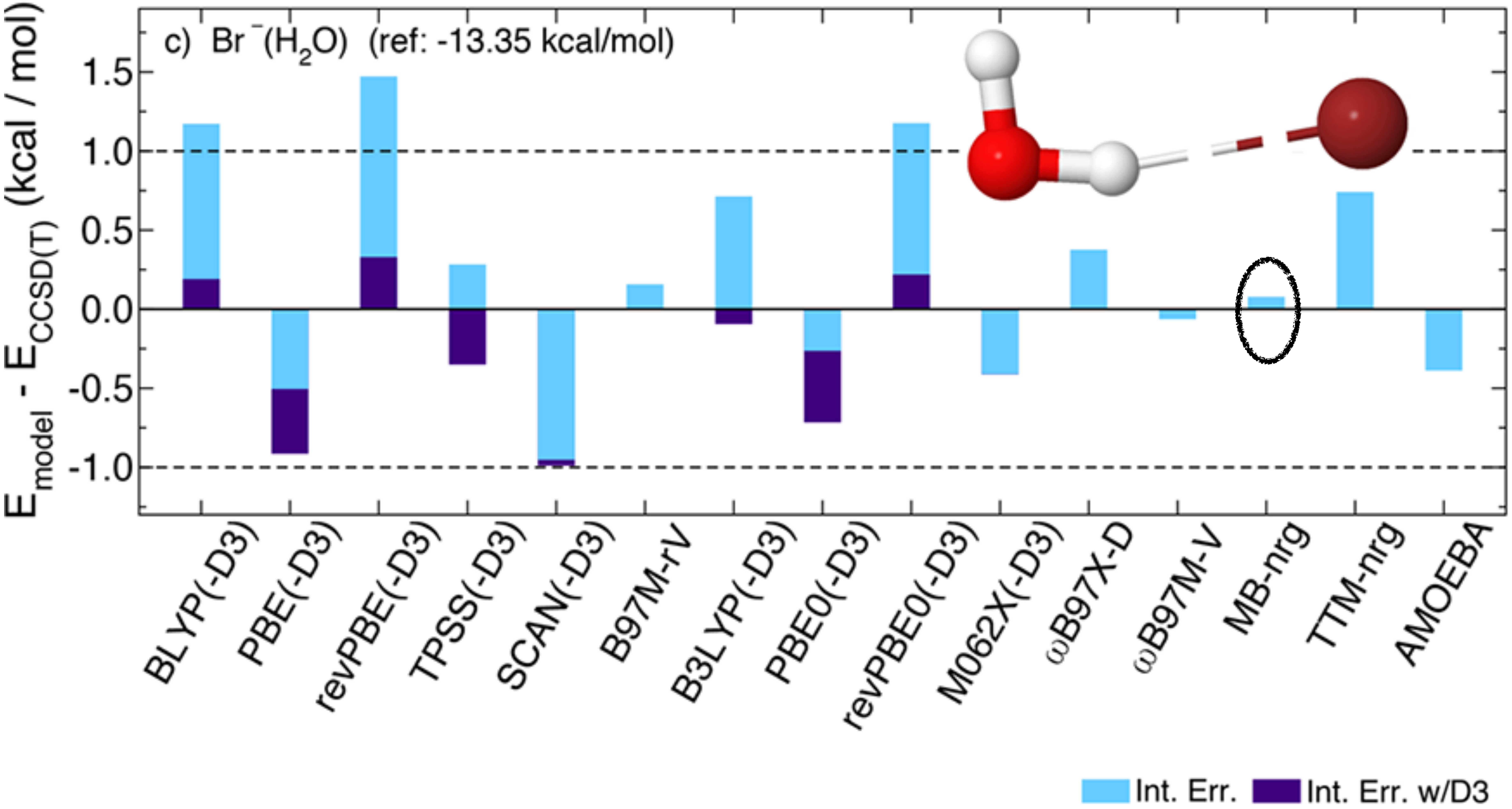


CCSD(T)
reference

Bizzarro, Egan & FP, J. Chem. Theory Comput. *15*, *2983 (2019)*

Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

$Br^-(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS



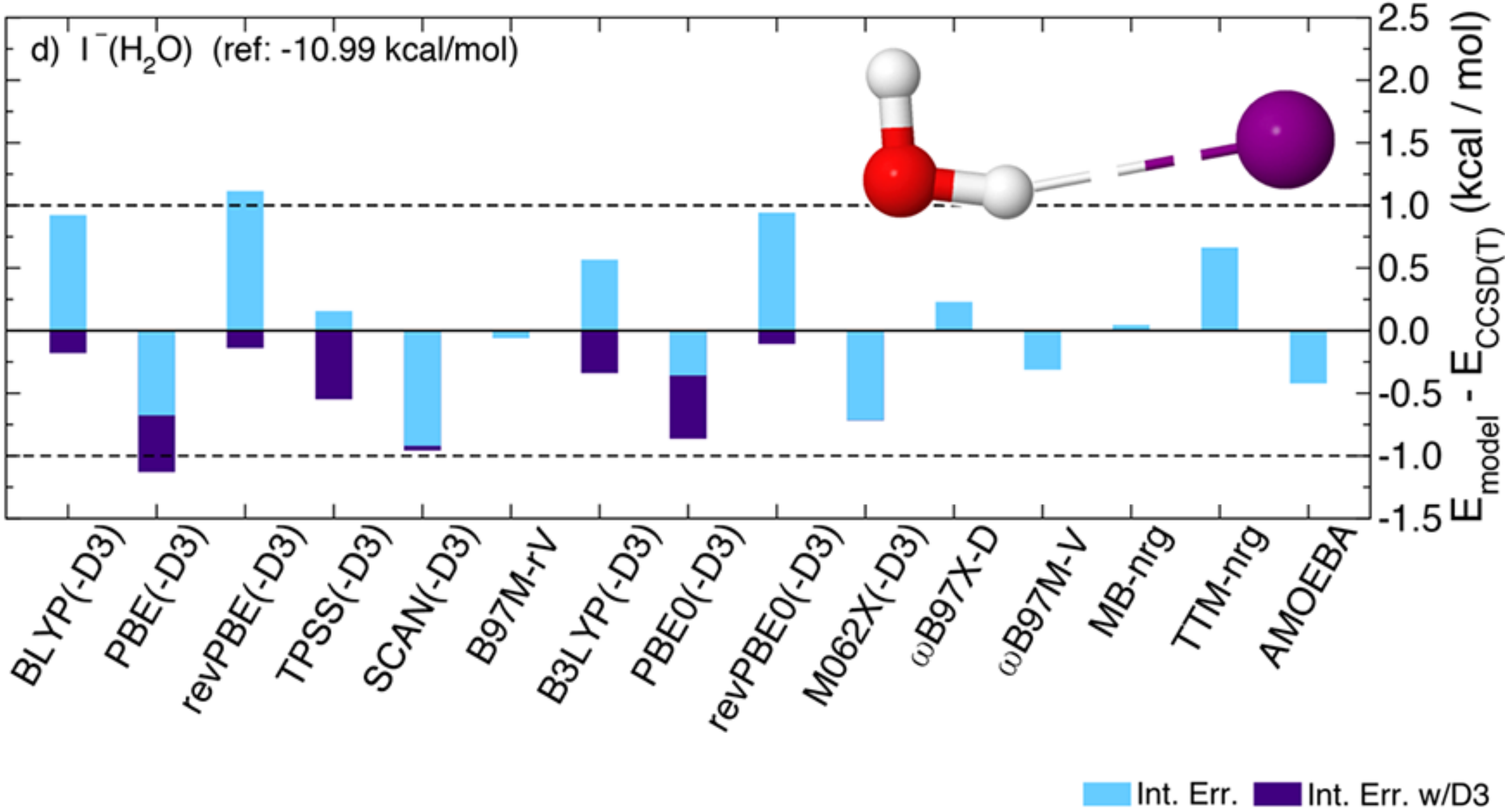
CCSD(T)
reference

Bizzarro, Egan & FP, J. Chem. Theory Comput. 15, 2983 (2019)

Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

$I^-(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS

CCSD(T) reference

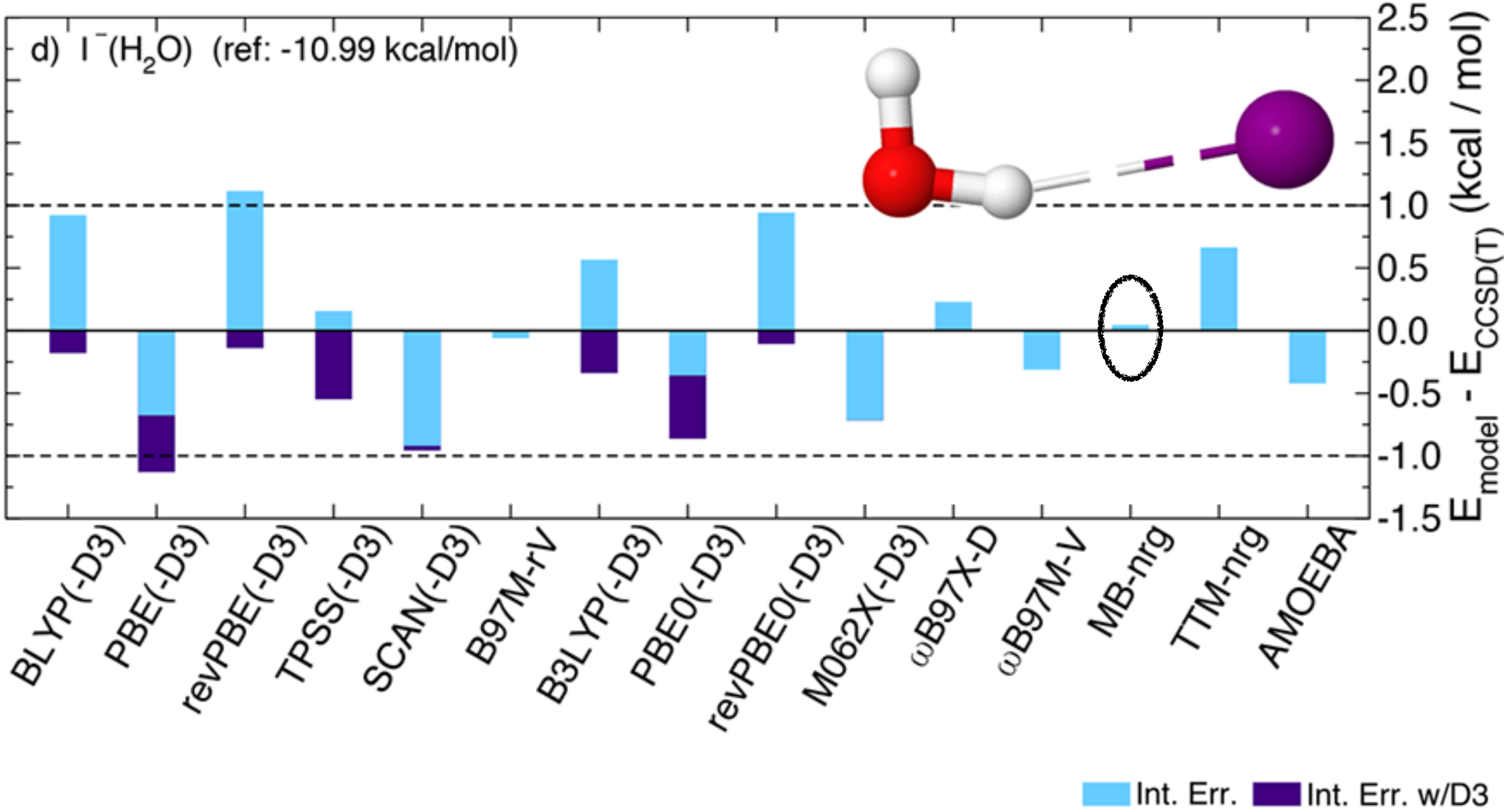


Bizzarro, Egan & FP, J. Chem. Theory Comput. 15, *2983 (2019)*

Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

$I^-(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS

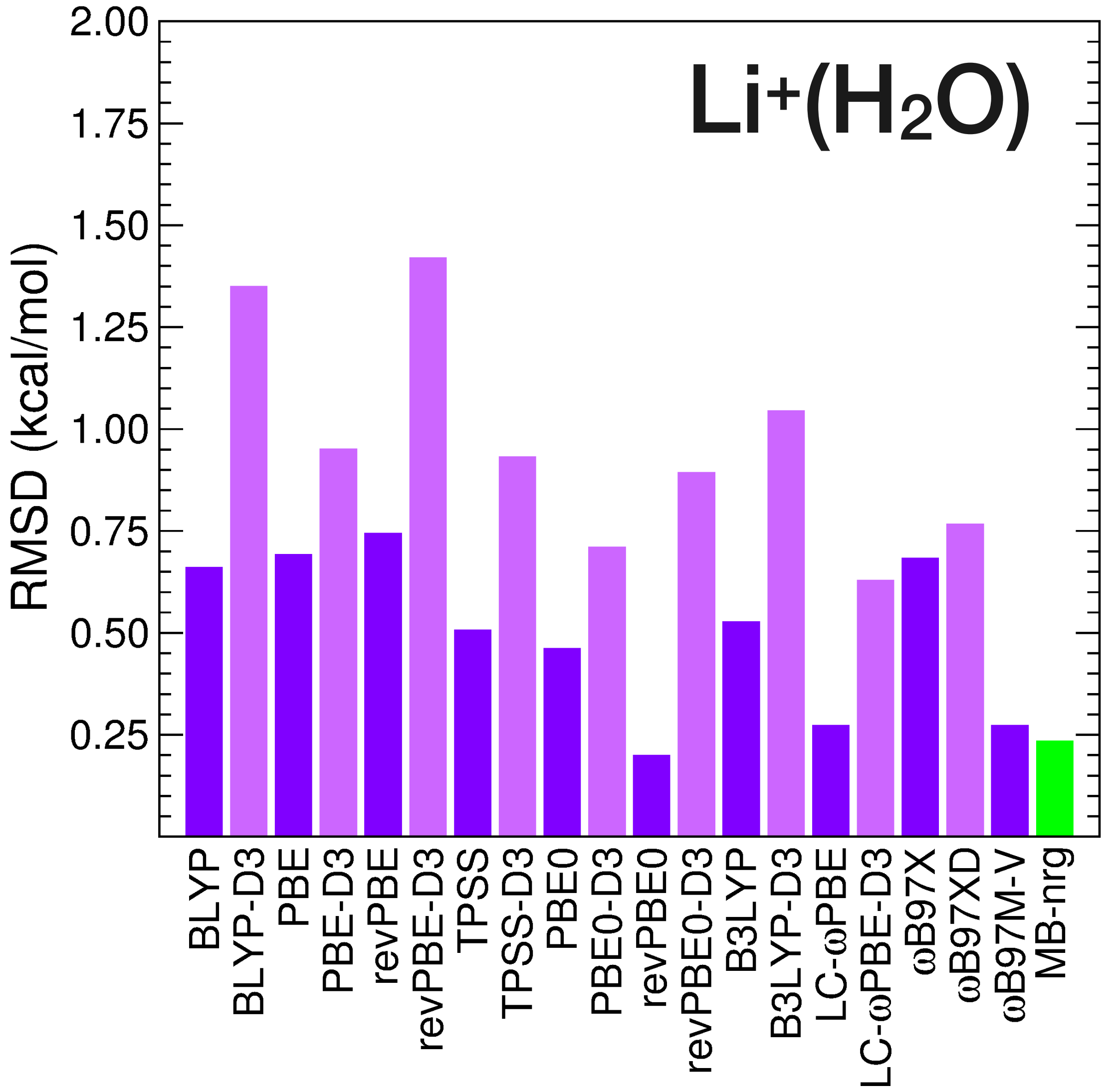
CCSD(T) reference

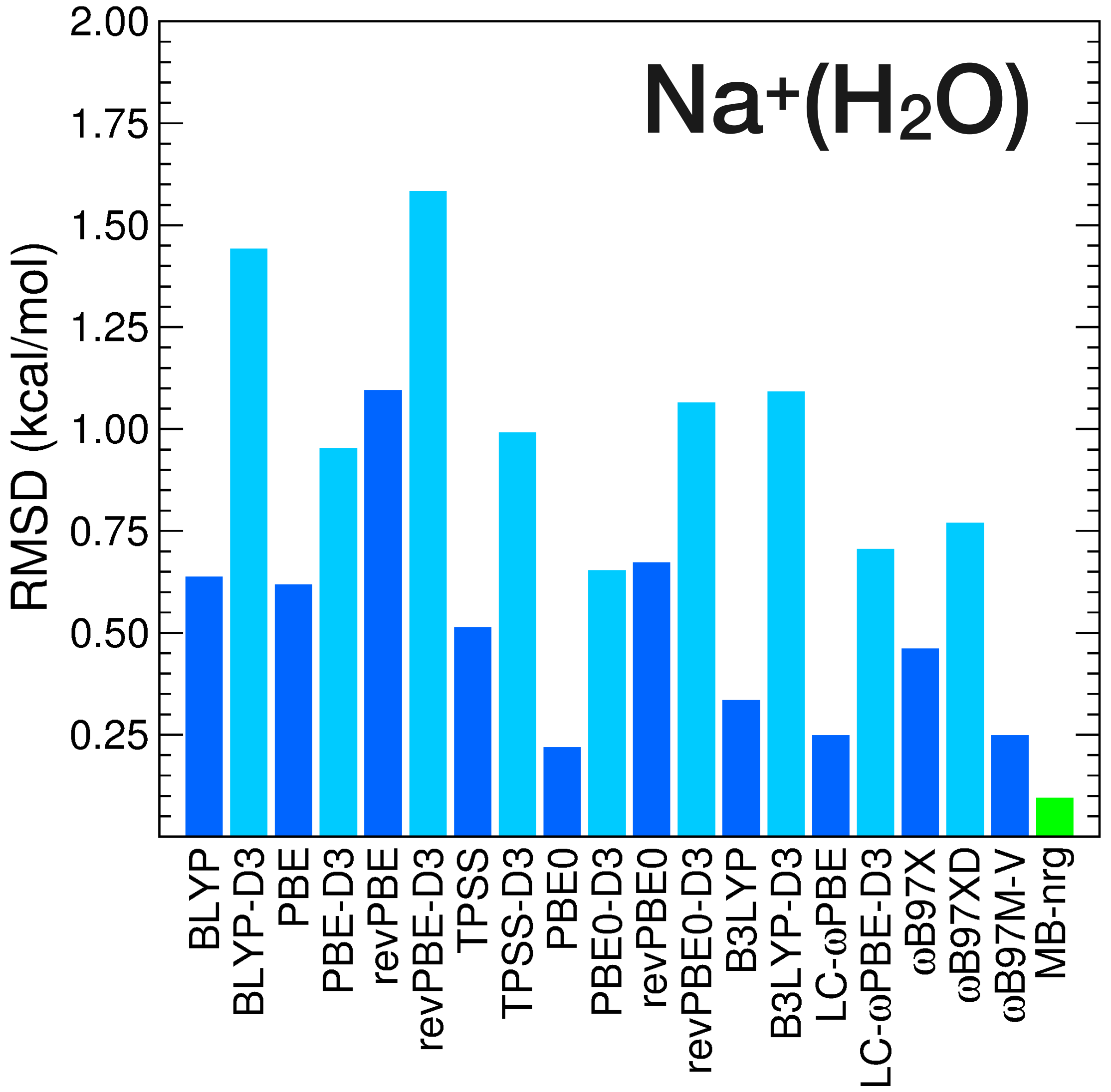


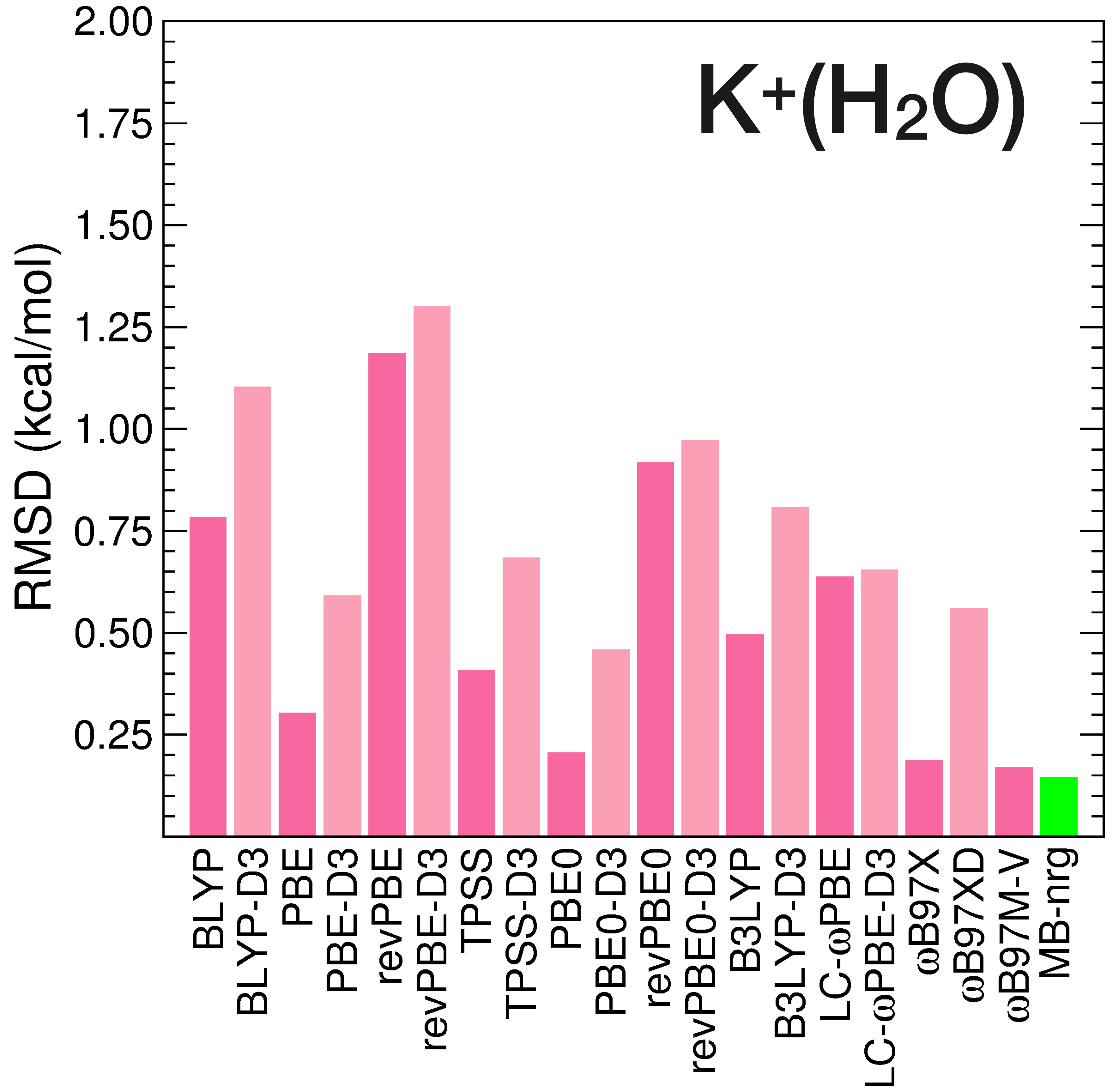
Bizzarro, Egan & FP, J. Chem. Theory Comput. 15, 2983 (2019)

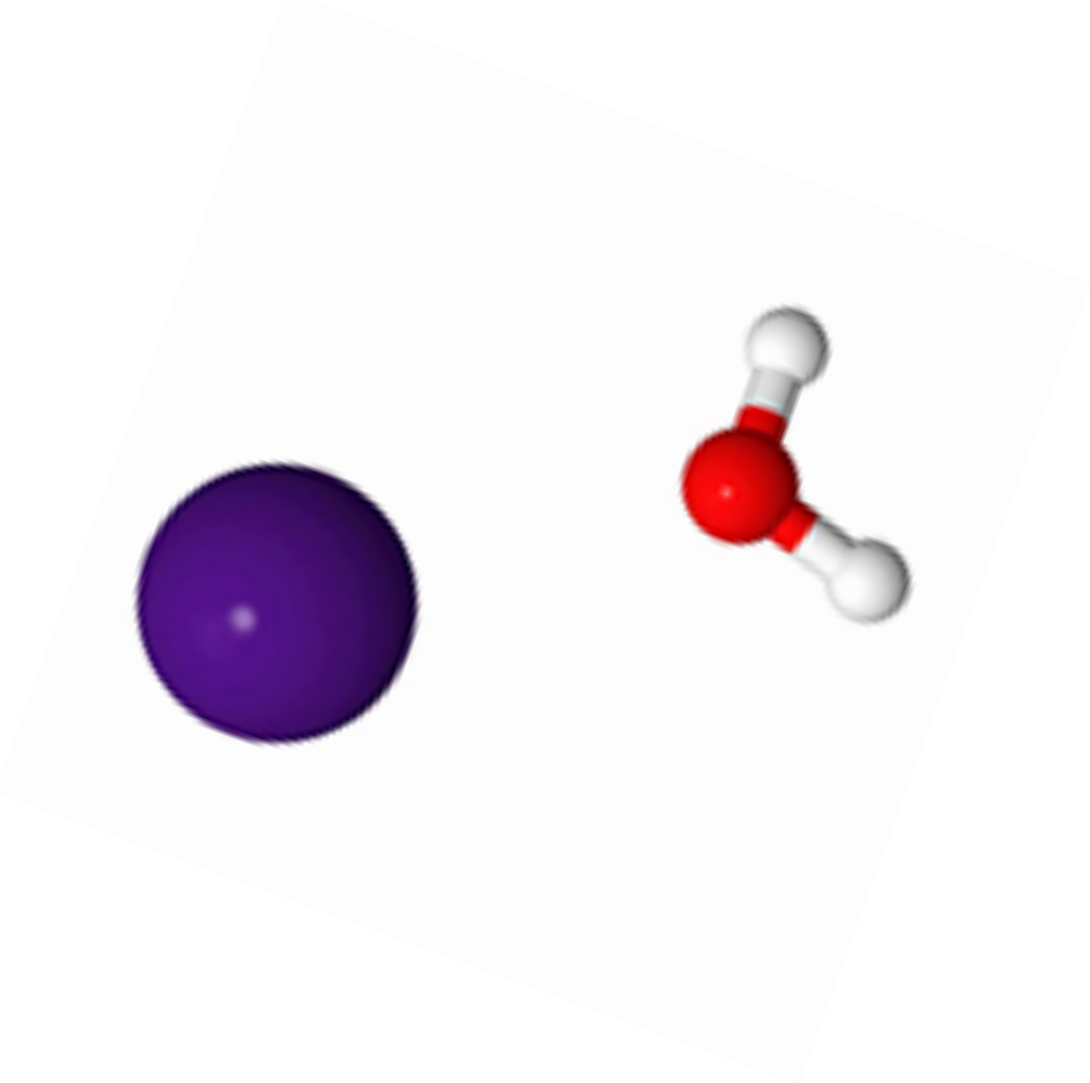
Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

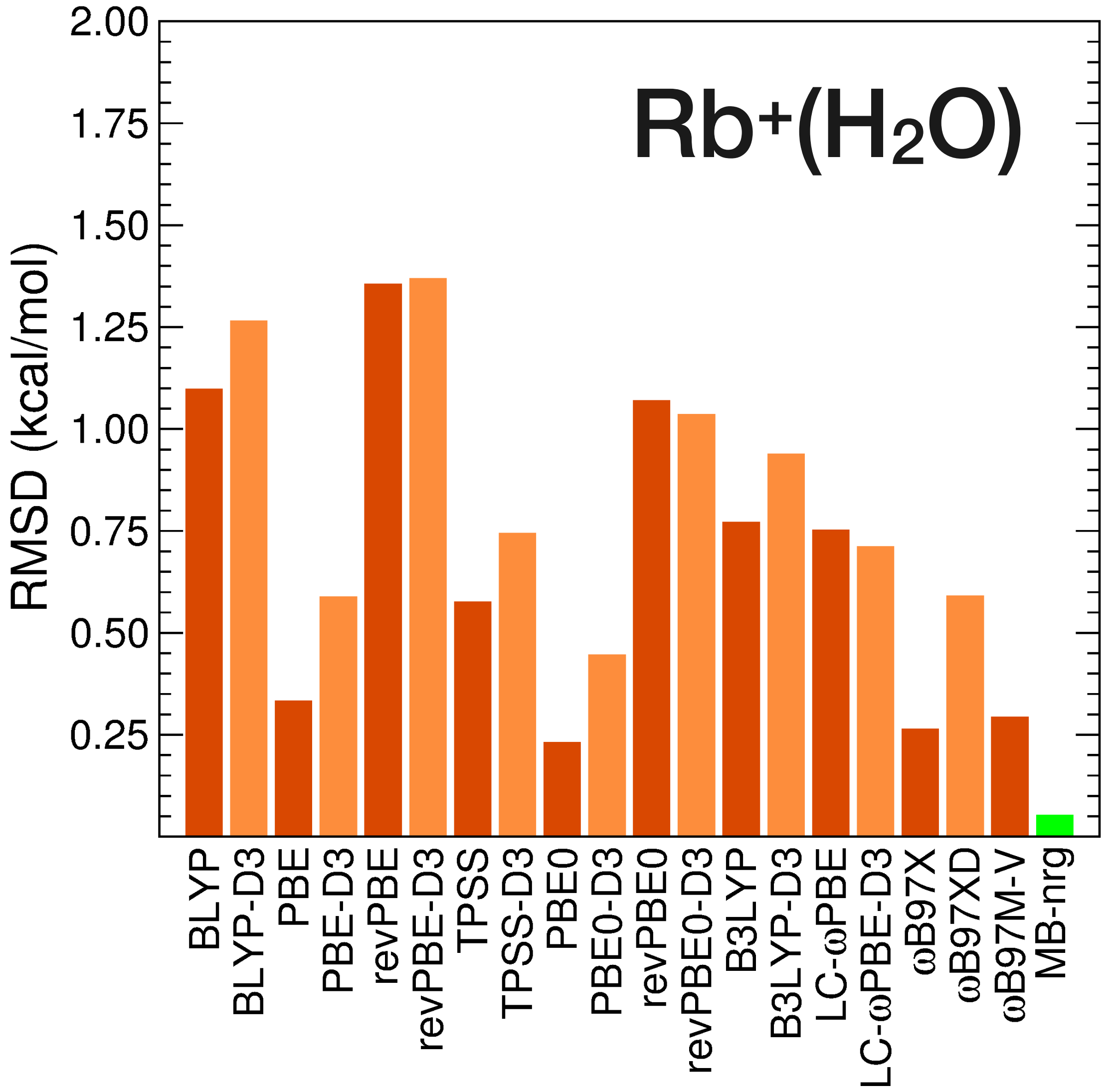
$M^+(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS

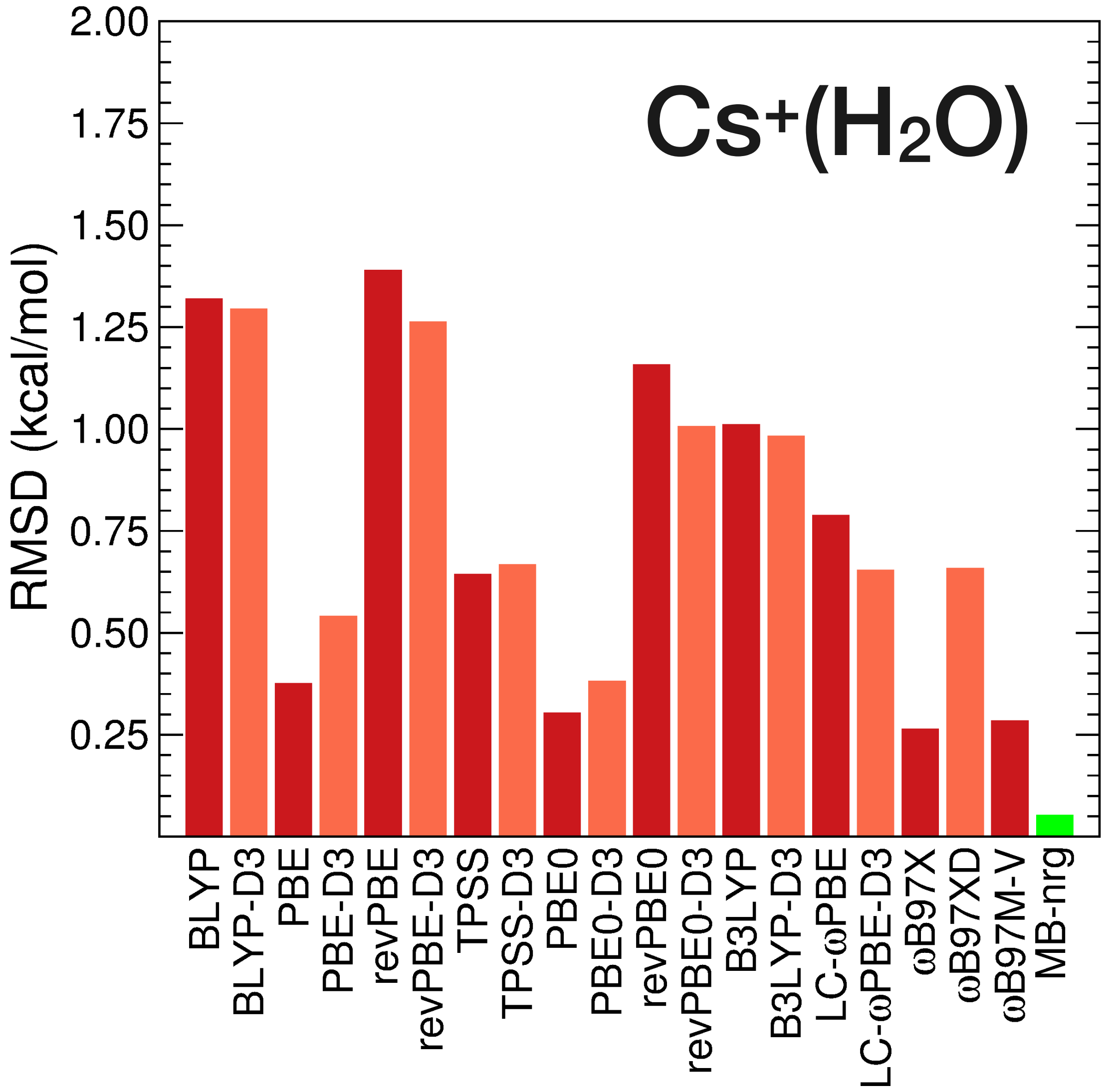








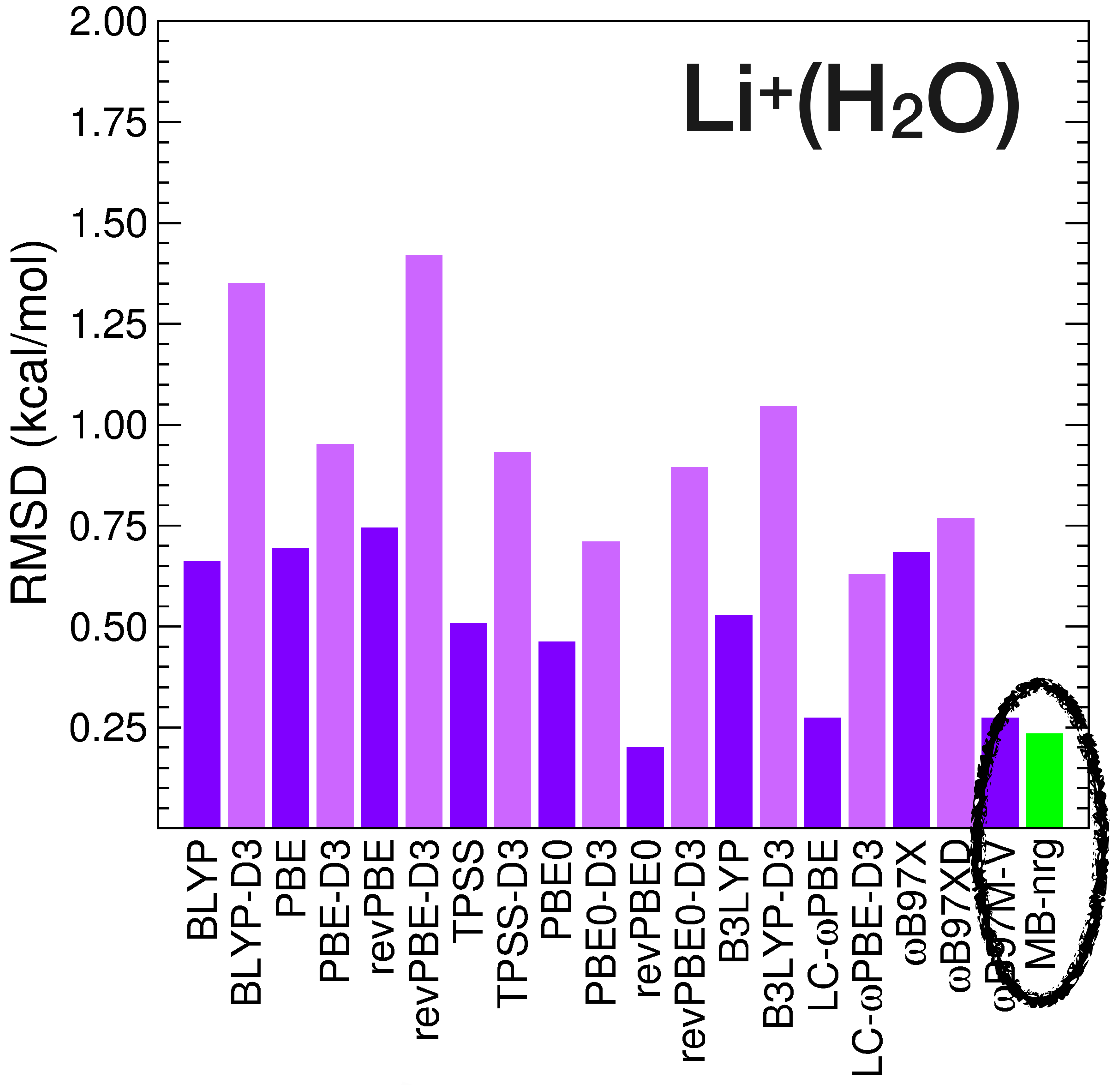


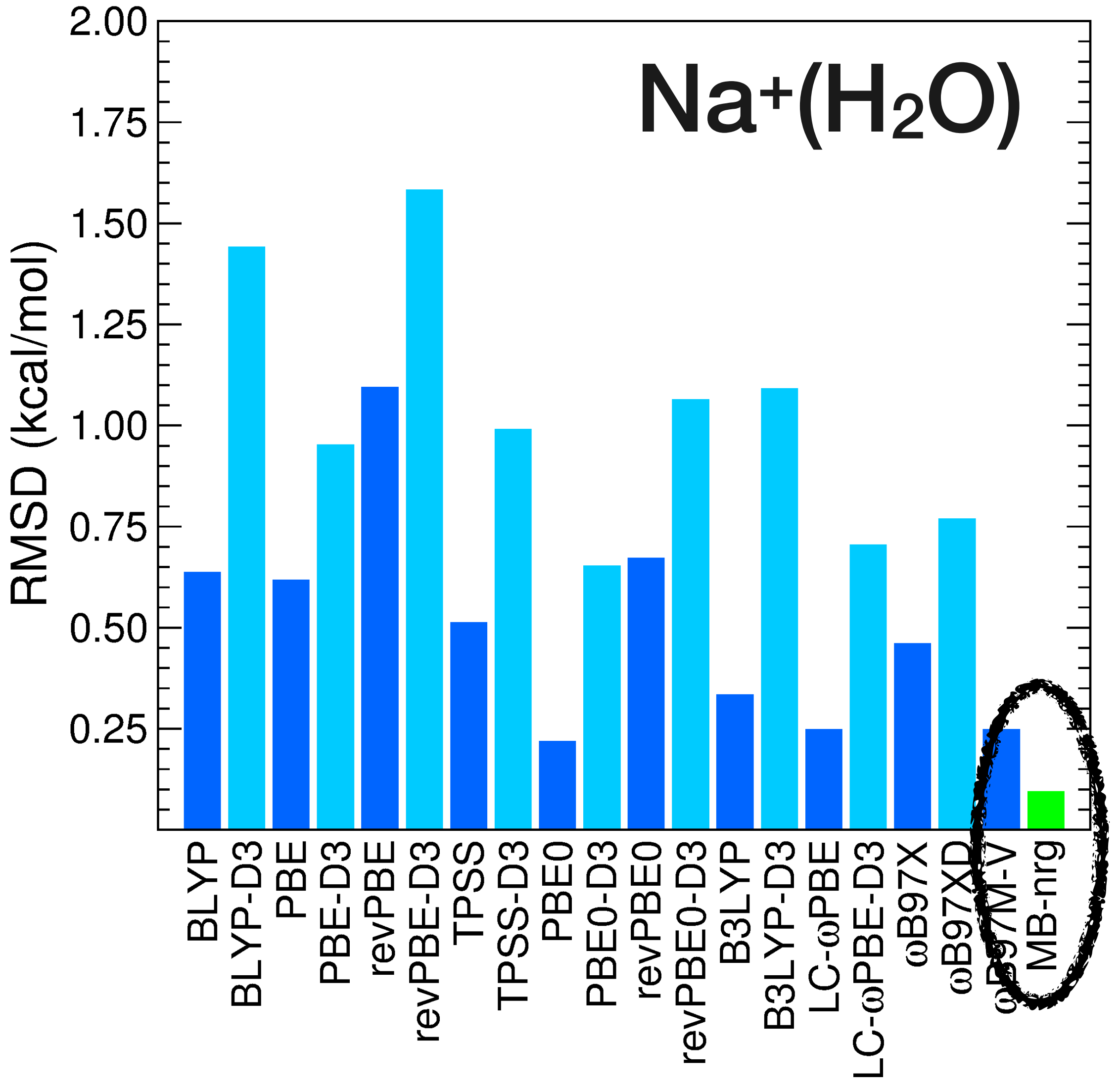


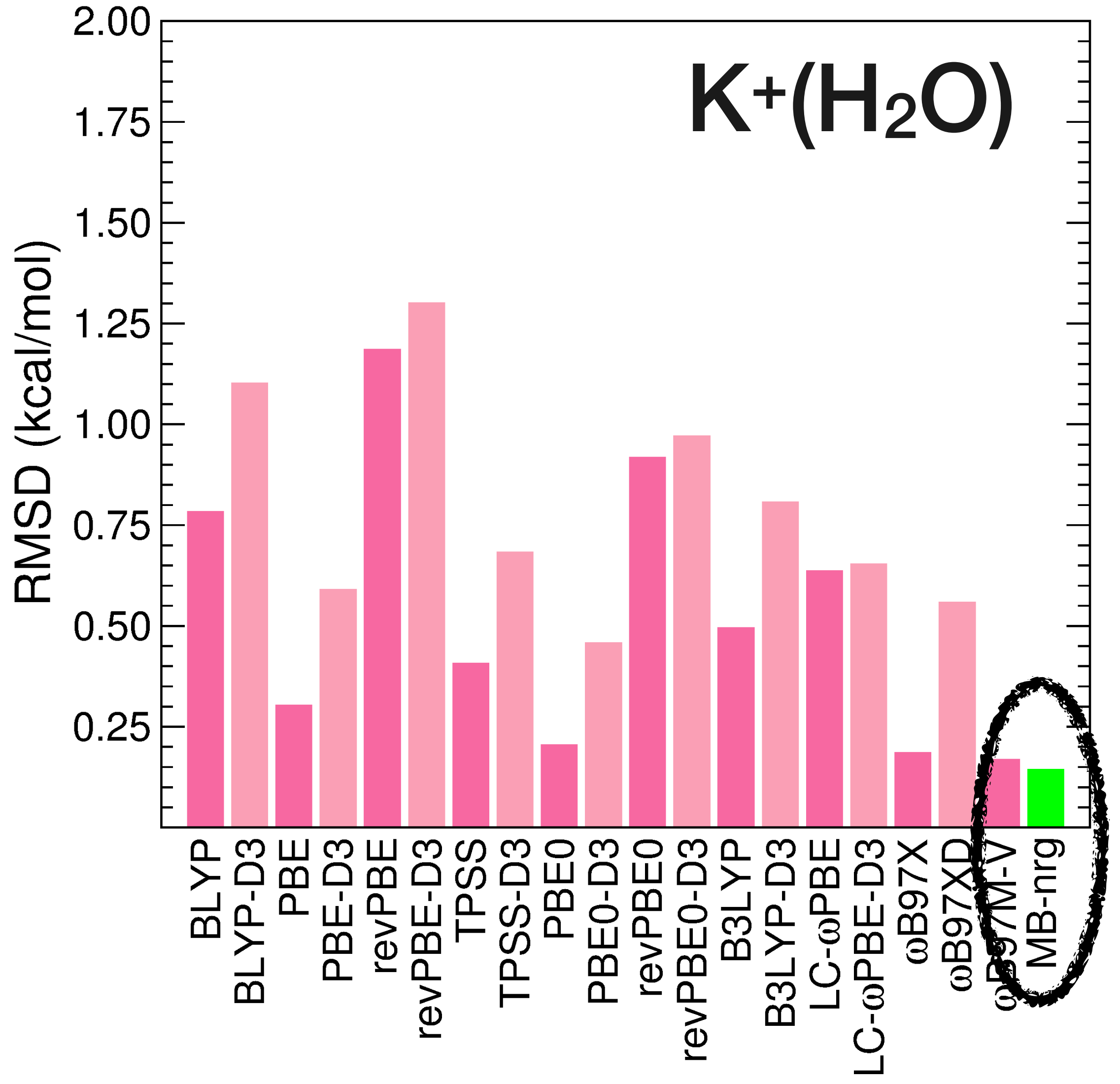
Riera, Mardirossian, Bajaj, Götz & FP, J. Chem. Phys. *147*, *161715 (2017)*

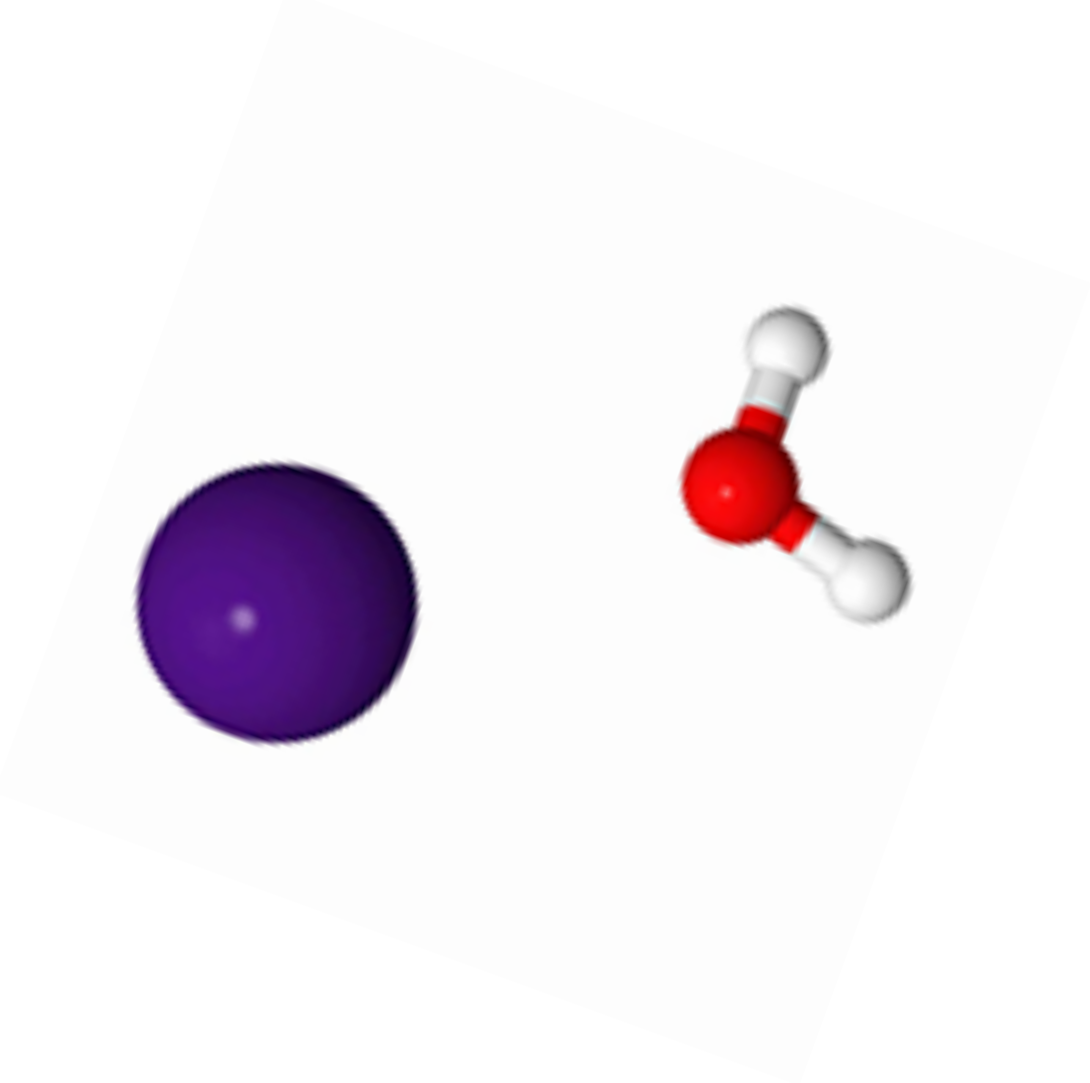
Putting Things in Perspective: MB-nrg vs. Force Fields vs. DFT

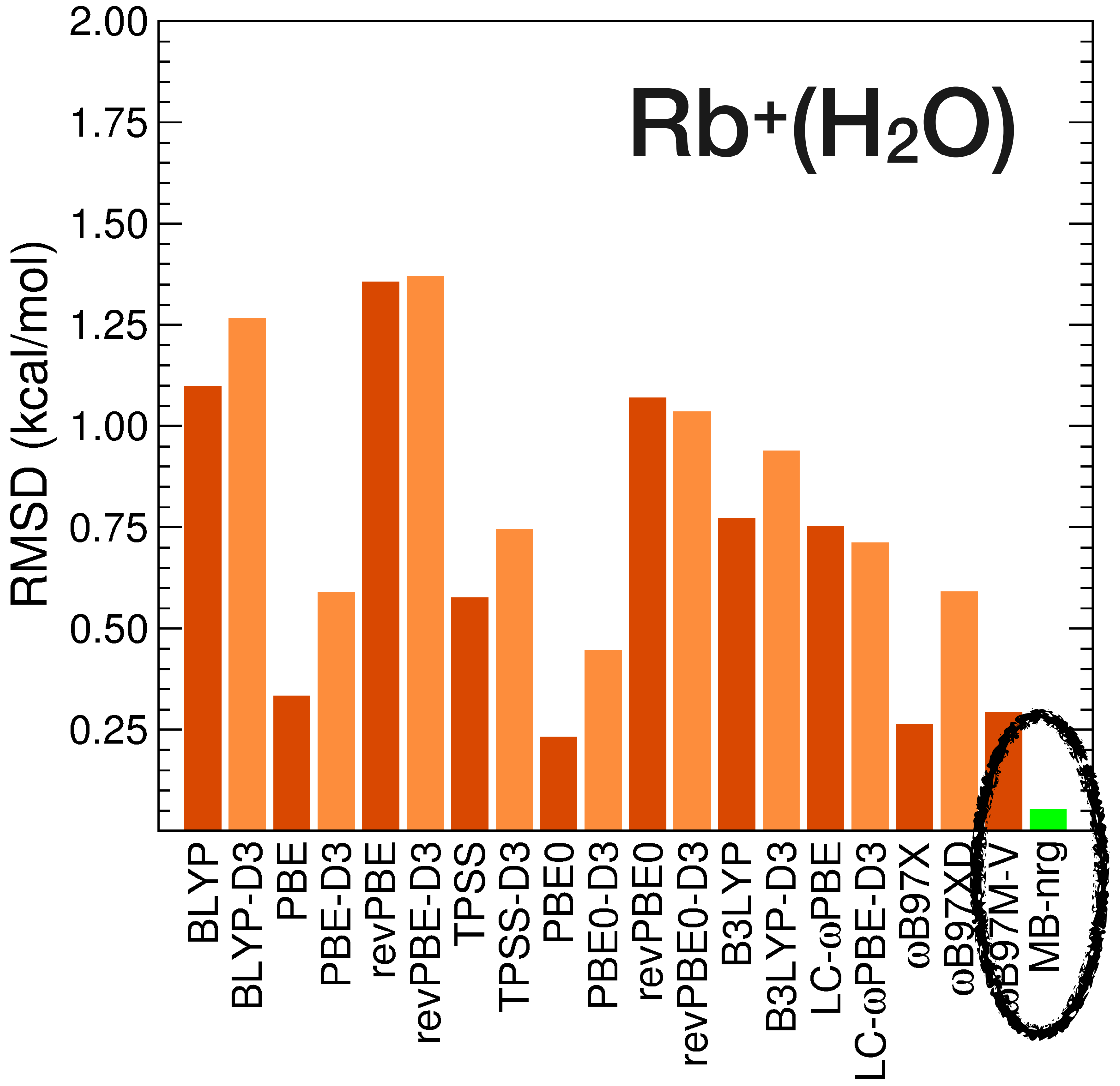
$M^+(H_2O)$: Errors in interaction energies relative to CCSD(T)/CBS

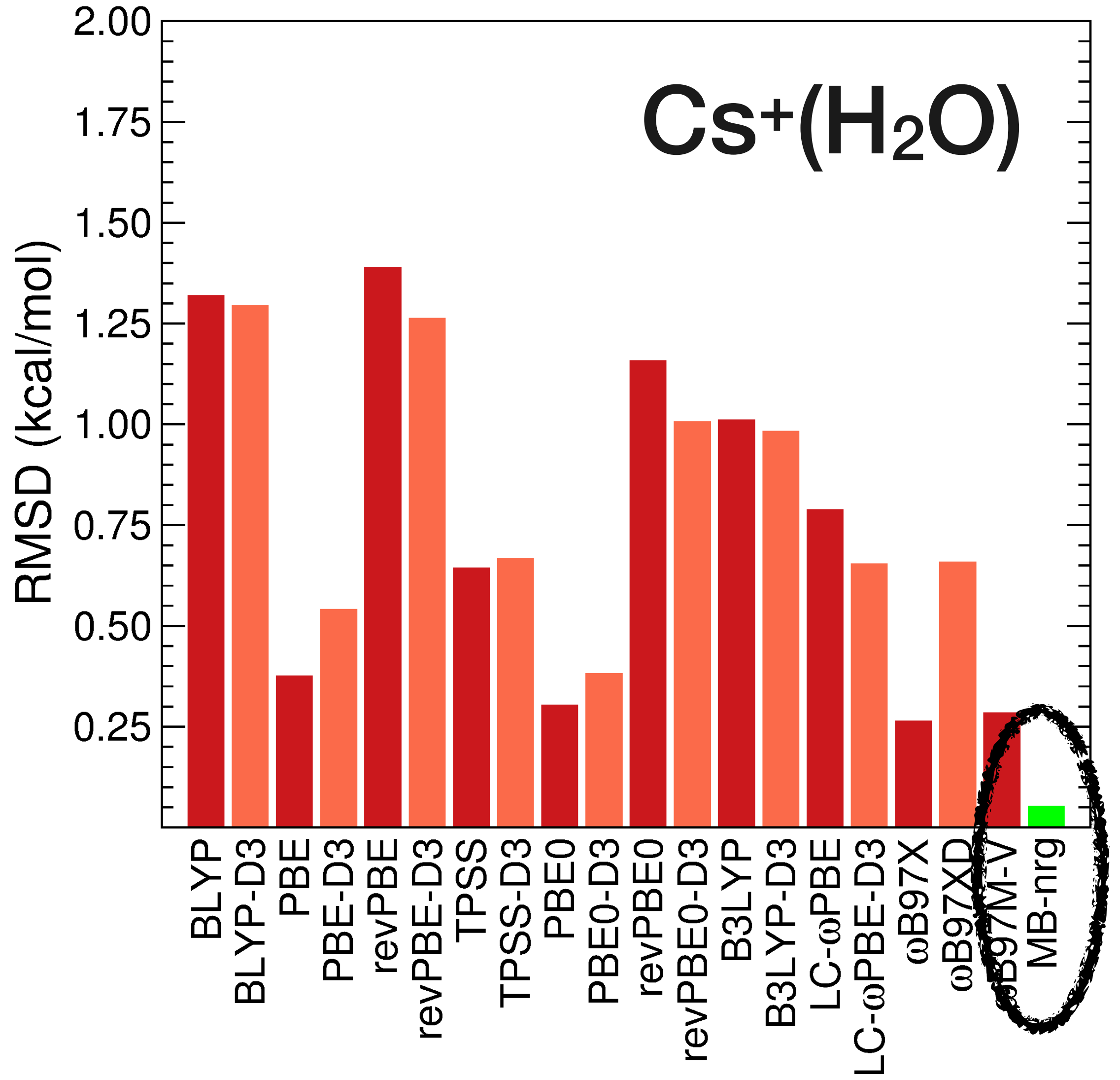








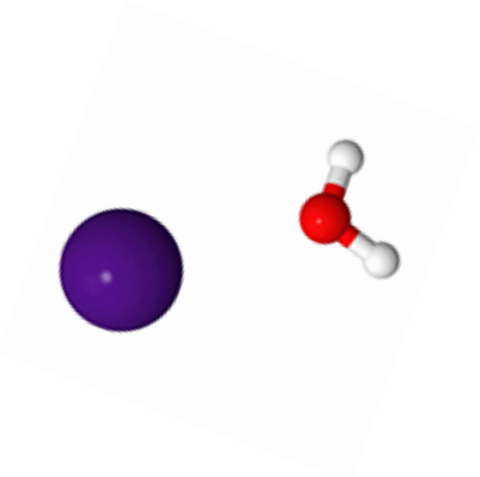




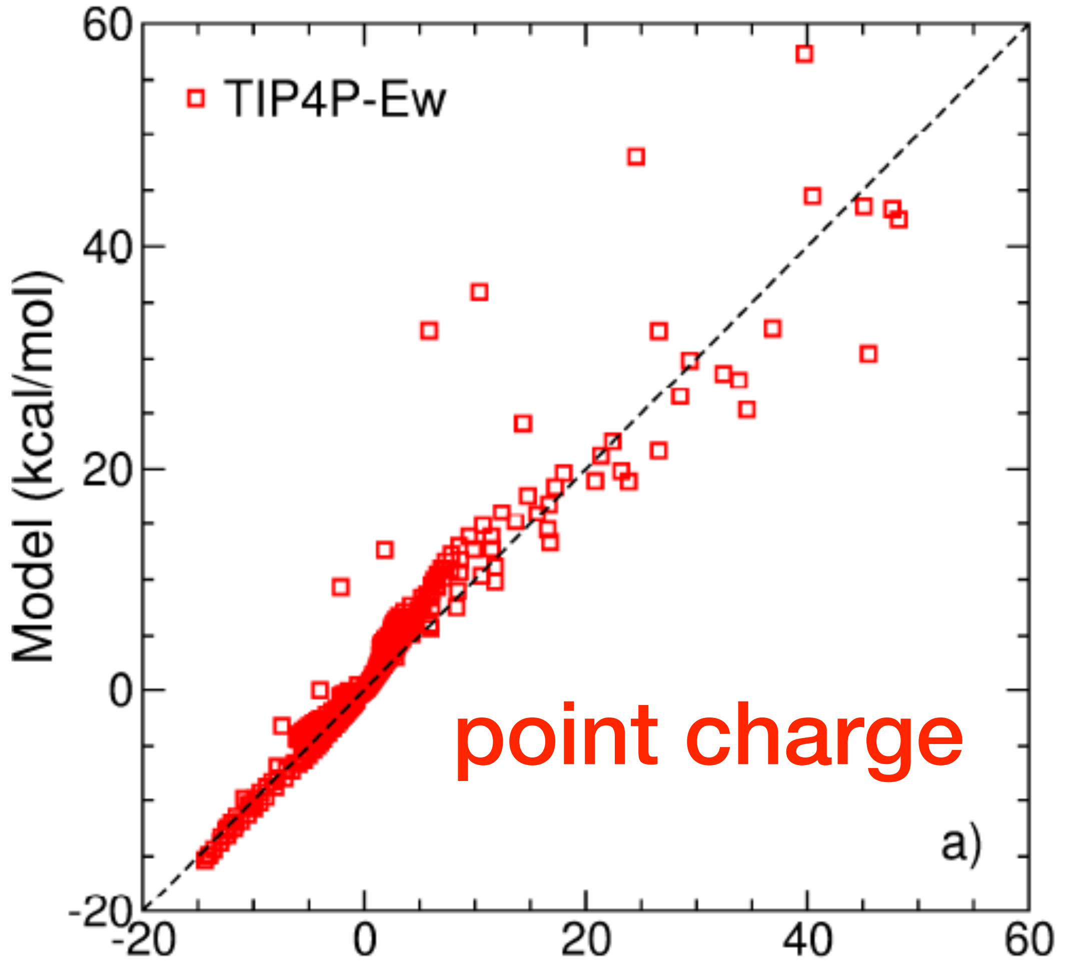
Riera, Mardirossian, Bajaj, Götz & FP, J. Chem. Phys. 147, 161715 (2017)

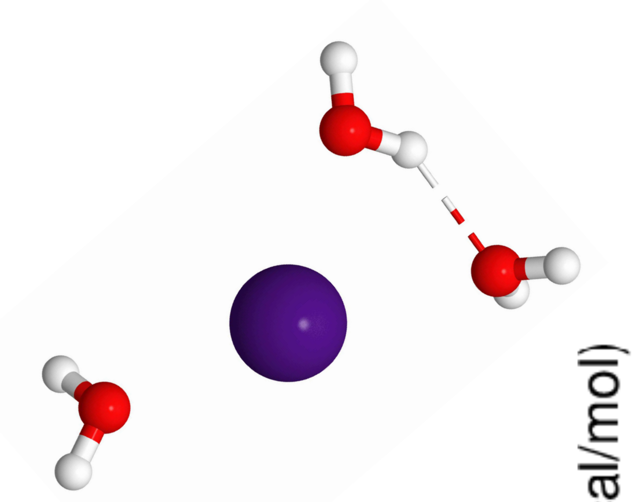
Many-Body Effects in Cs+ Hydration

Comparison between different water-Cs+ models

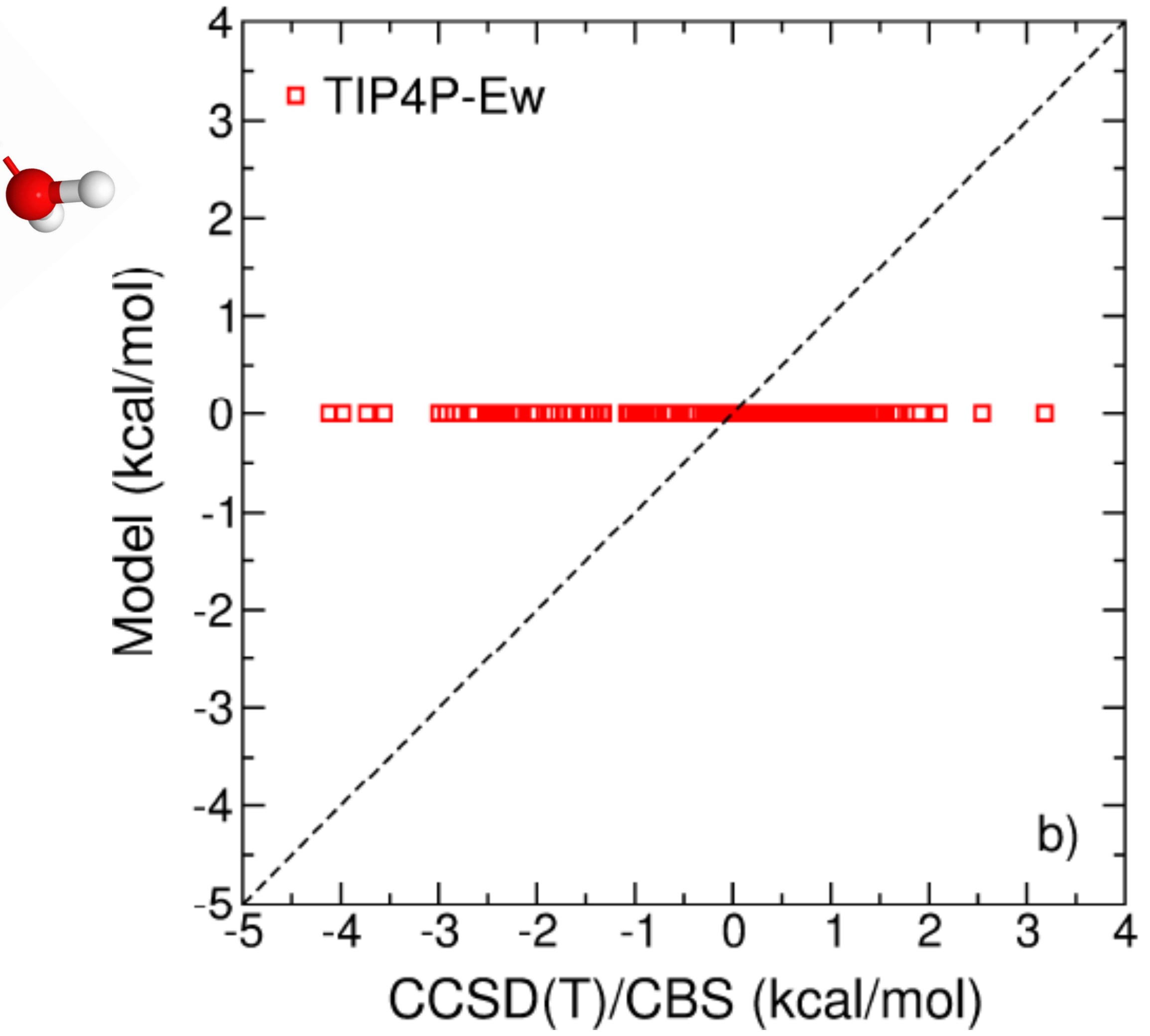


2-body





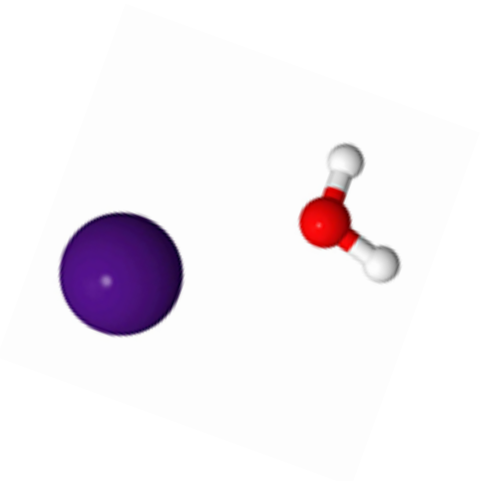
3-body



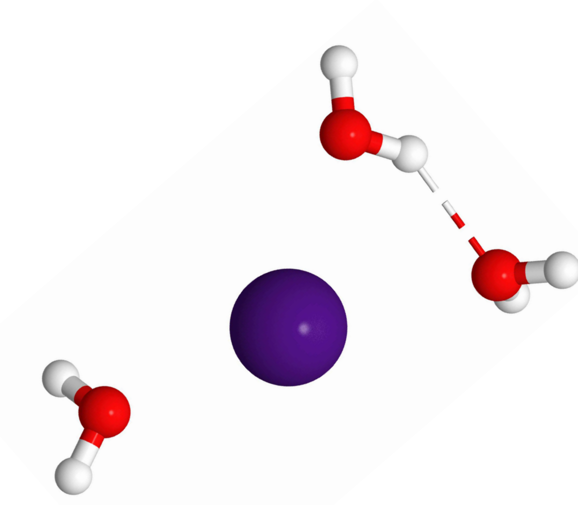
Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Many-Body Effects in Cs+ Hydration

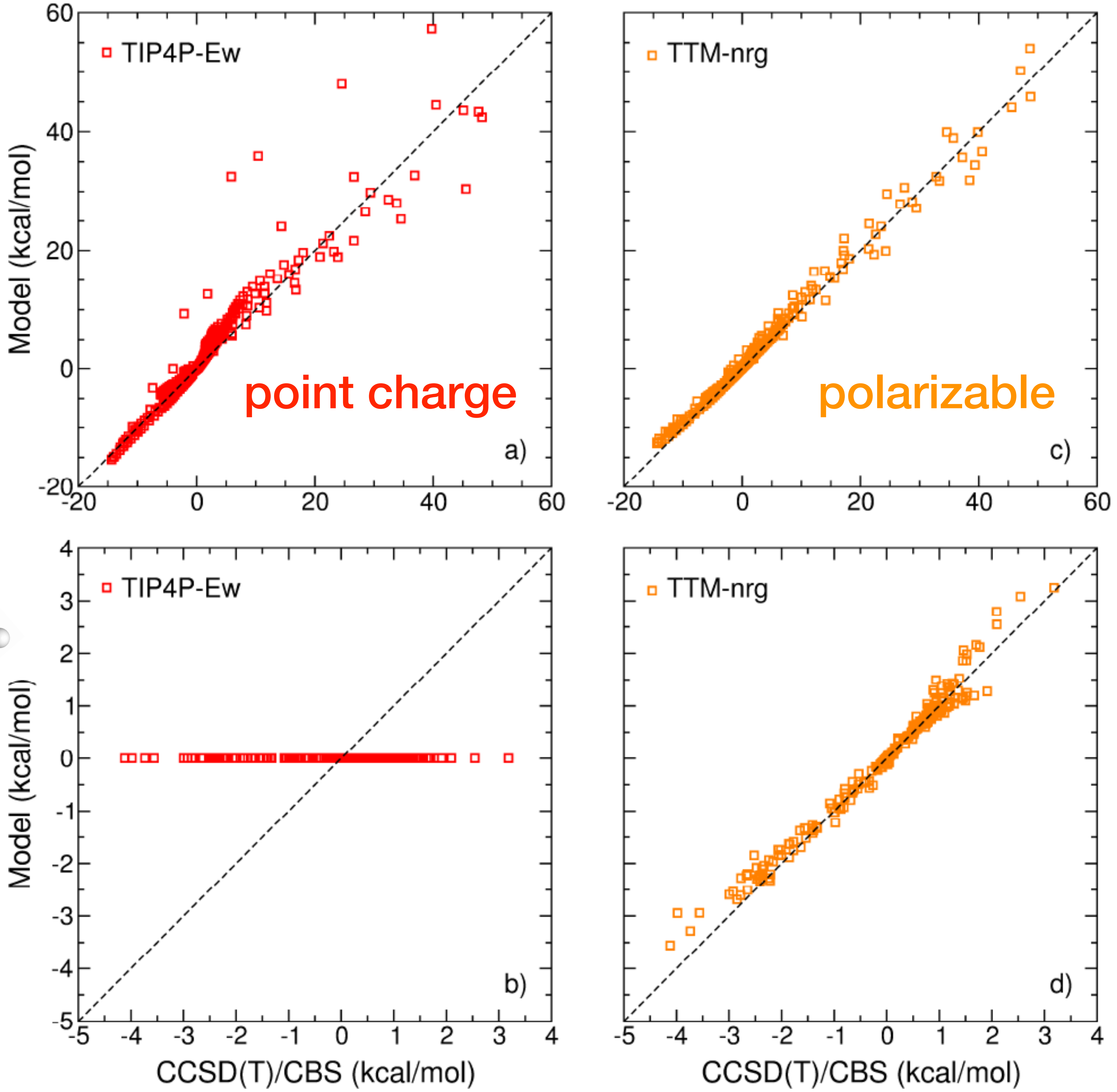
Comparison between different water-Cs+ models



2-body



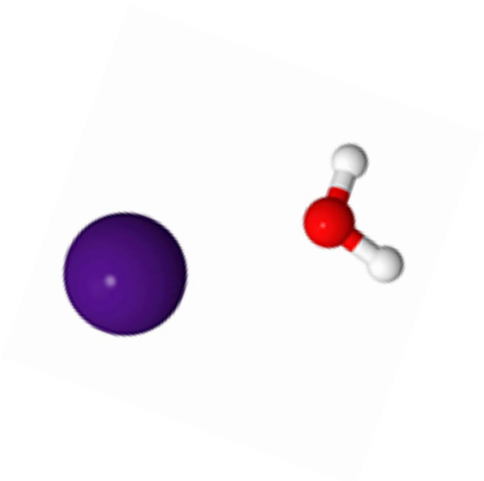
3-body



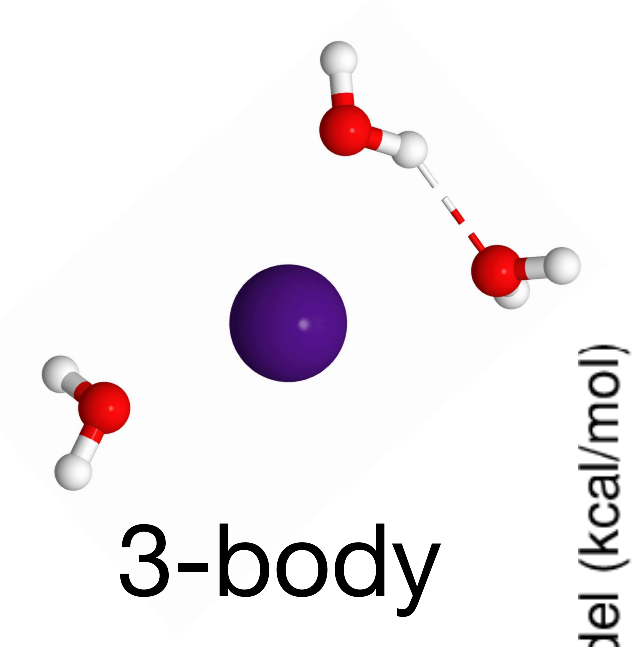
Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Many-Body Effects in Cs+ Hydration

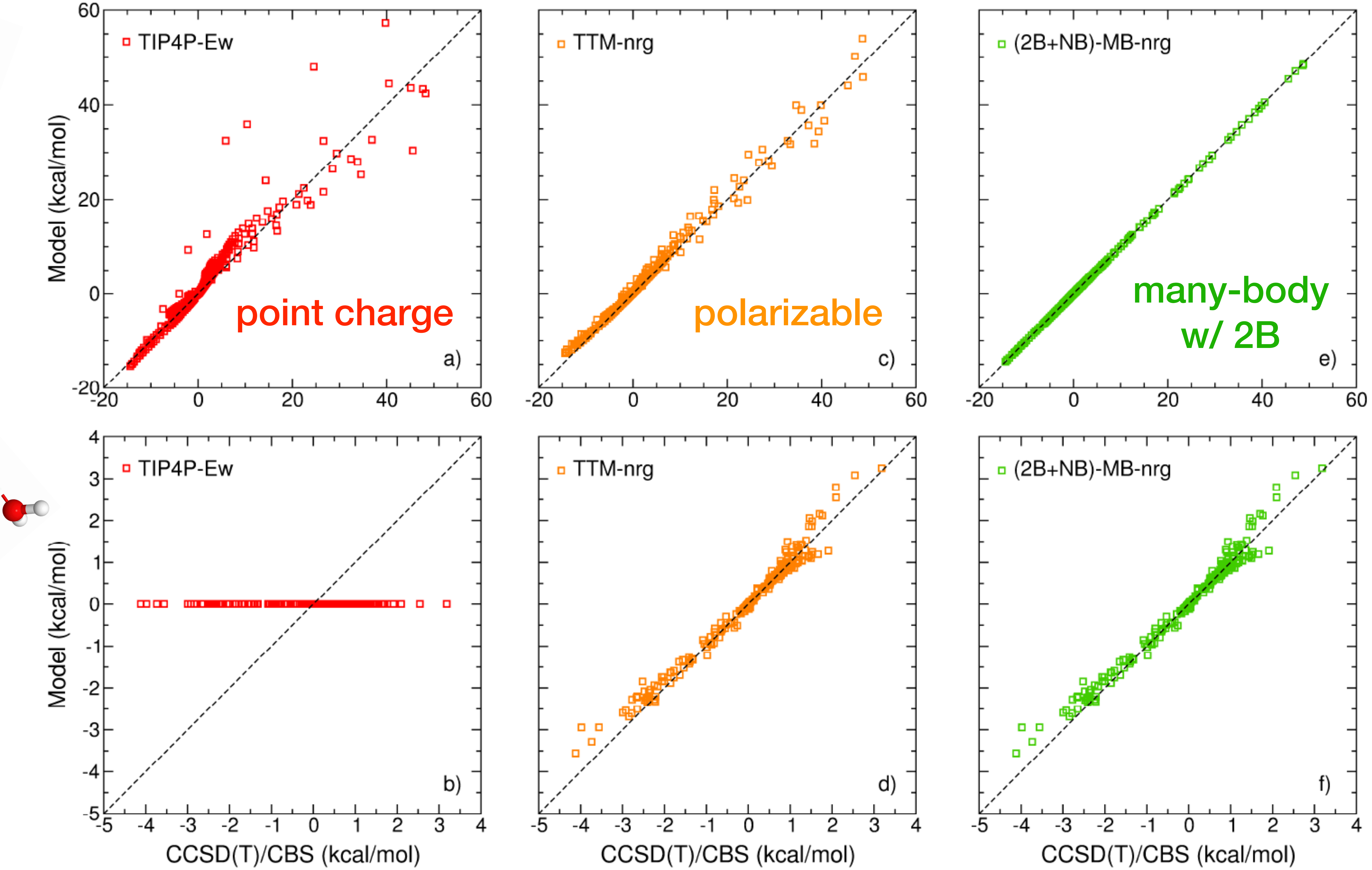
Comparison between different water-Cs+ models



2-body



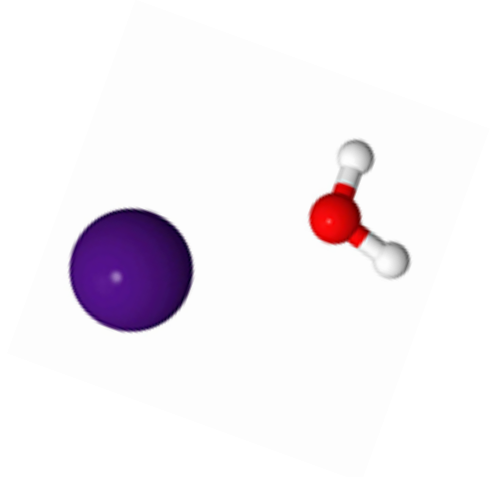
3-body



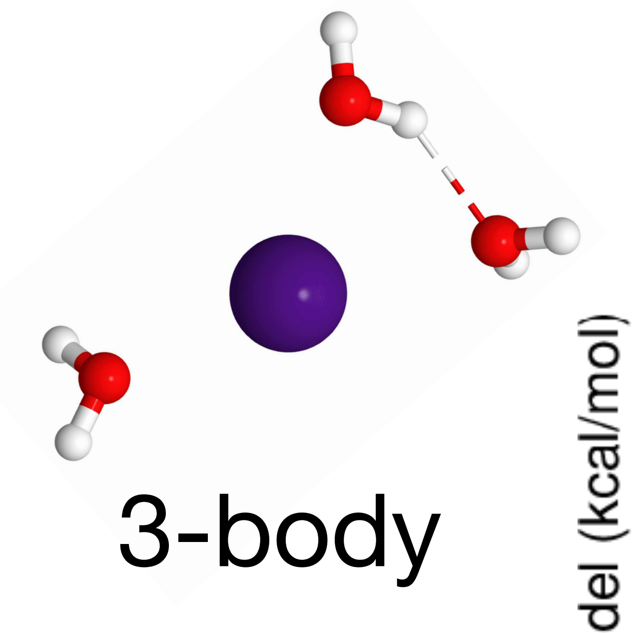
Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Many-Body Effects in Cs⁺ Hydration

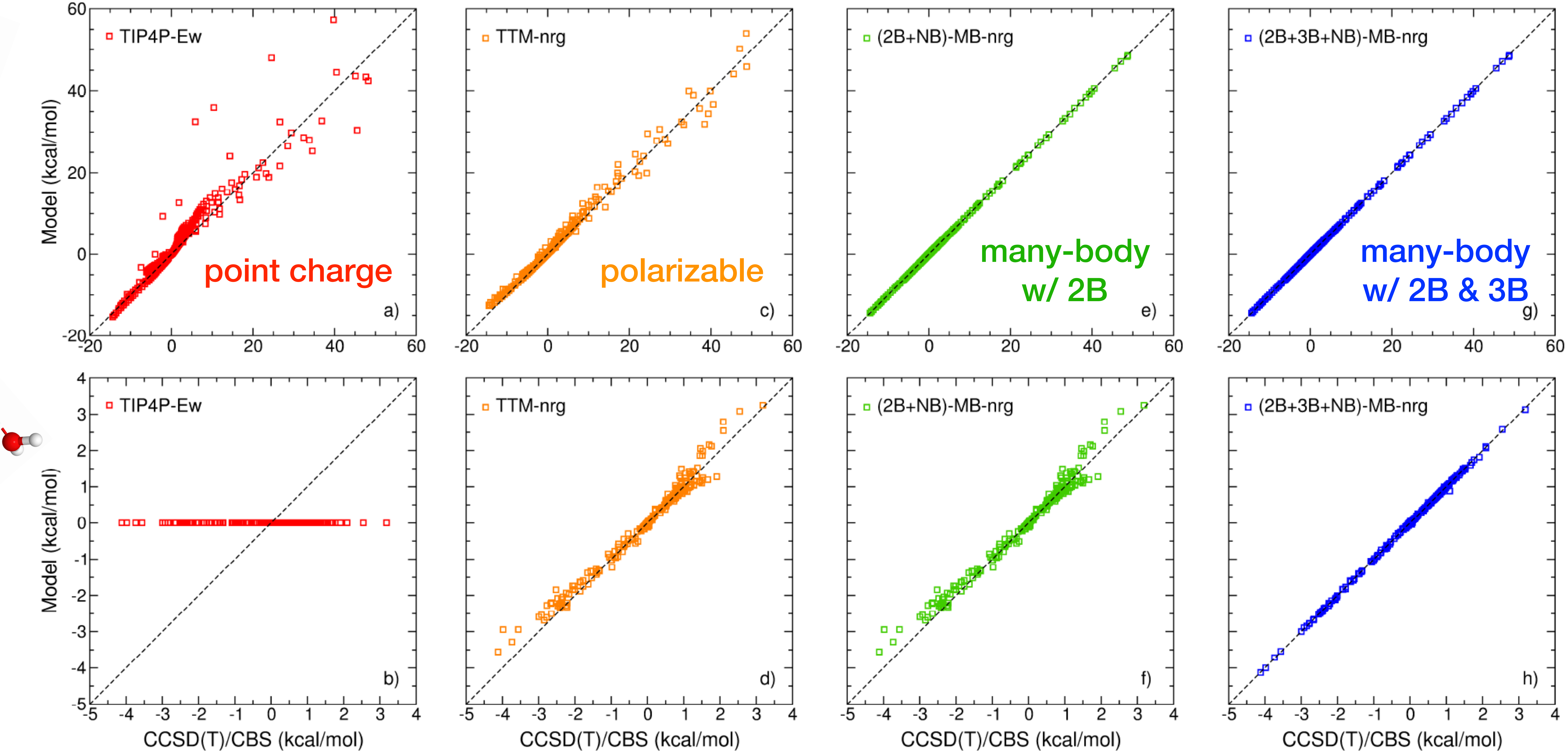
Comparison between different water-Cs⁺ models



2-body



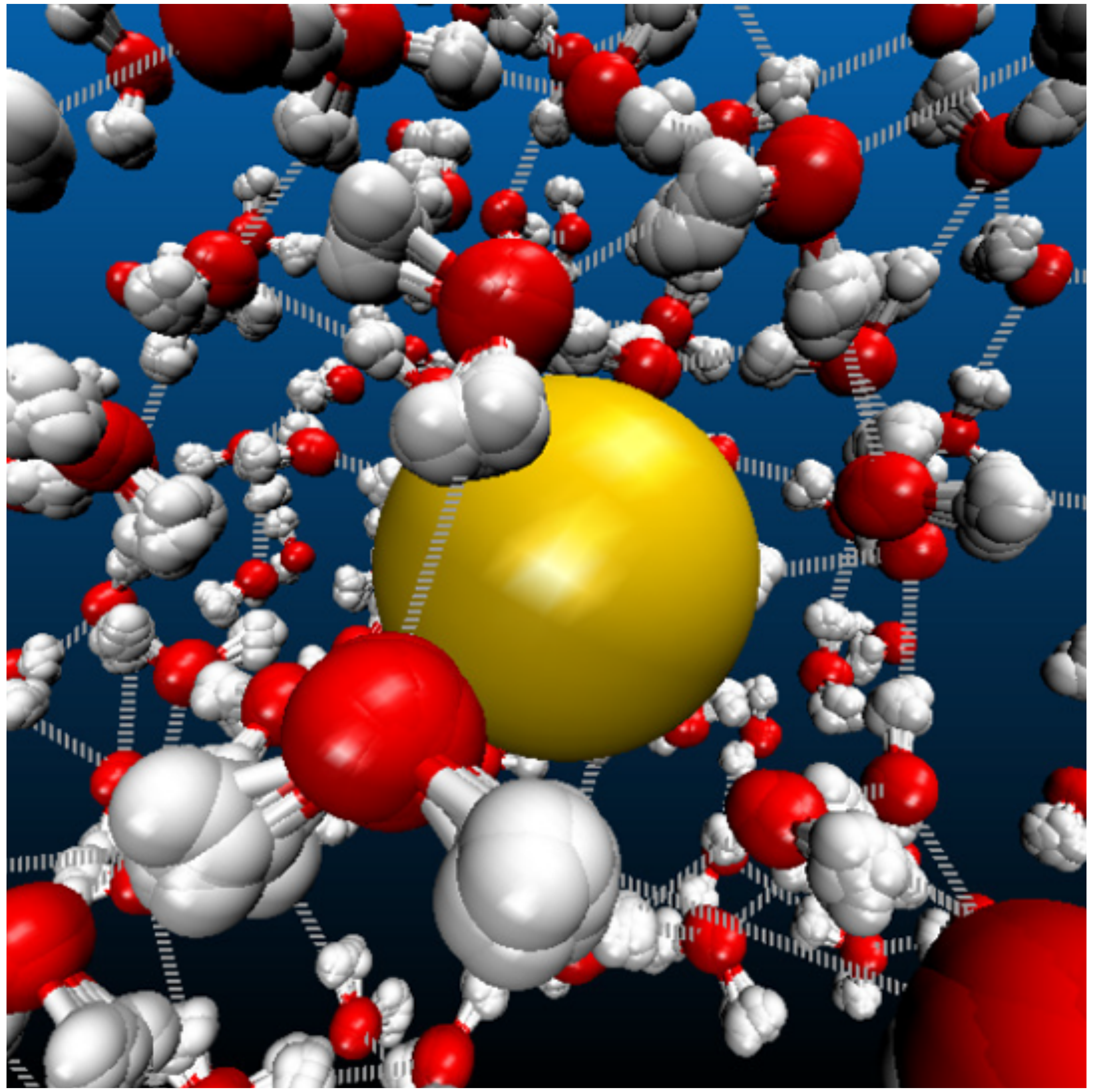
3-body

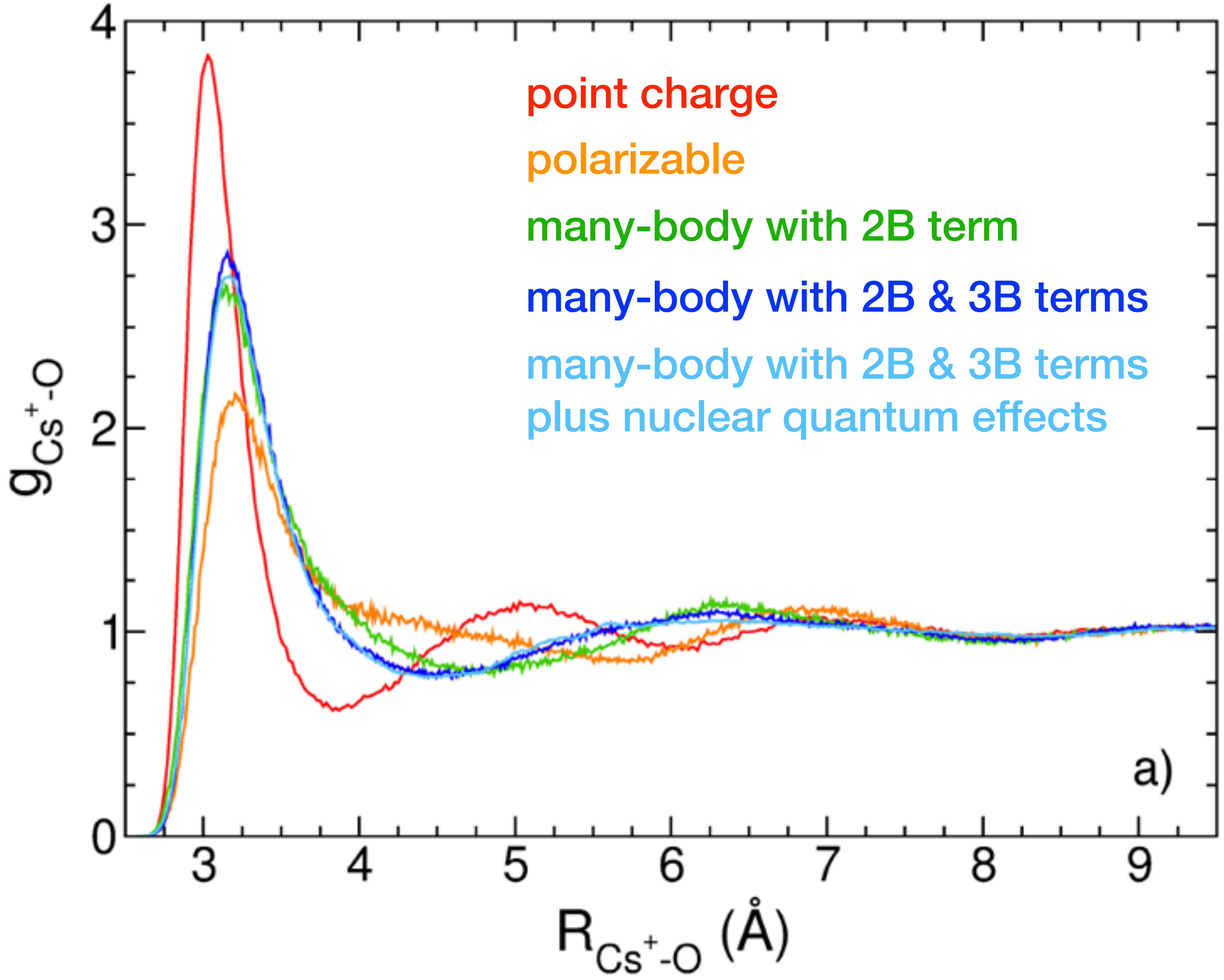


Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Many-Body Effects in Cs^+ Hydration

Comparison between different water-Cs^+ models





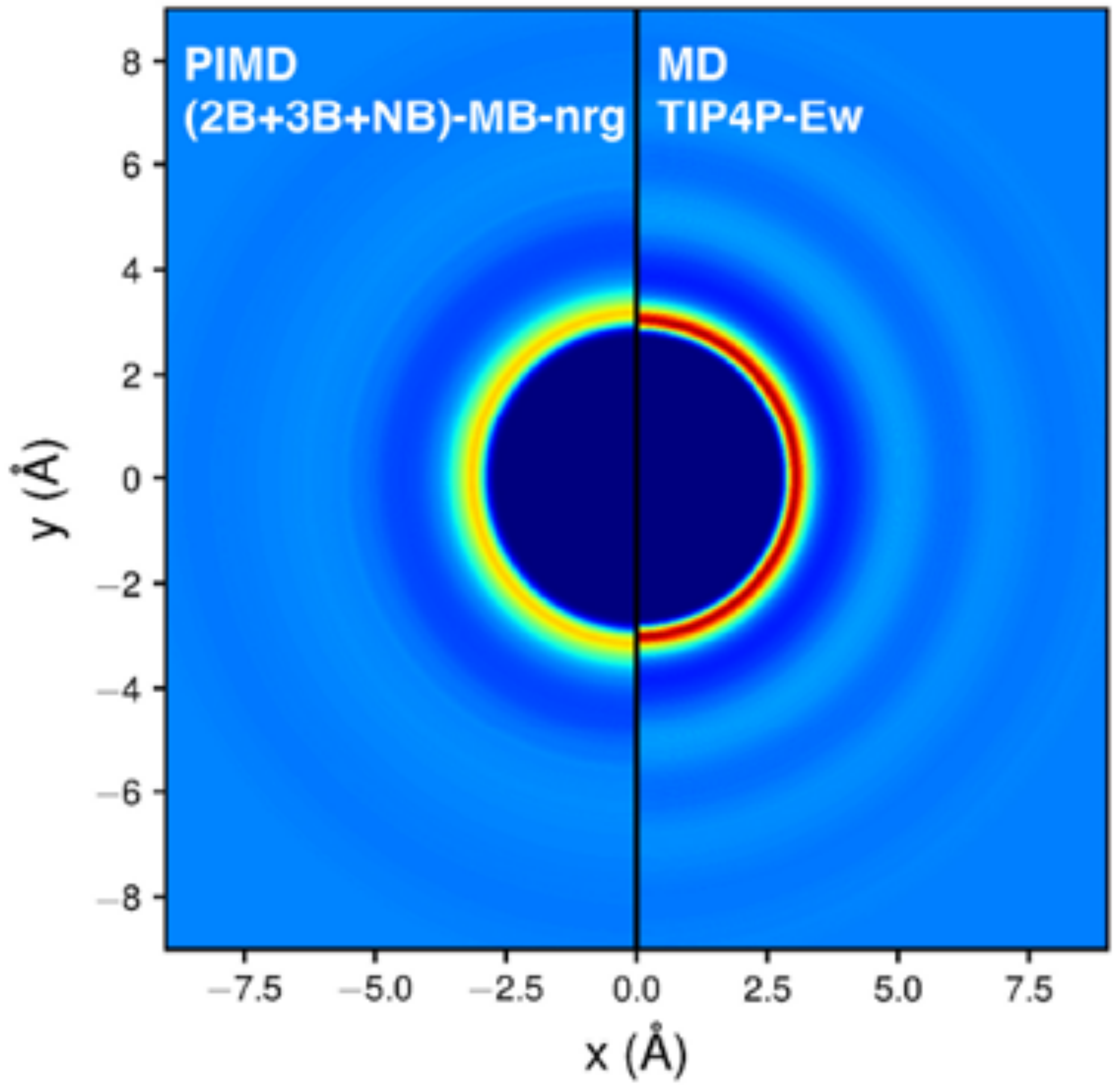
Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Many-Body Effects in Cs$^+$ Hydration

Comparison between different water-Cs$^+$ models

many-body
w/ 2B & 3B
+ NQE

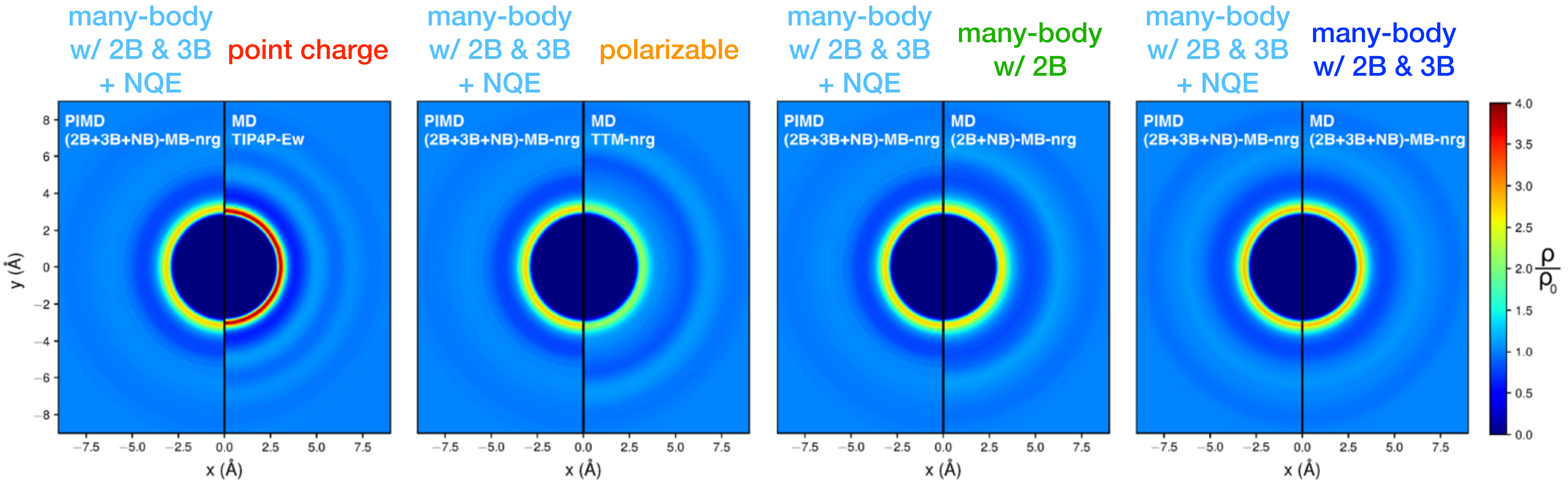
point charge



Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Many-Body Effects in Cs$^+$ Hydration

Comparison between different water-Cs$^+$ models



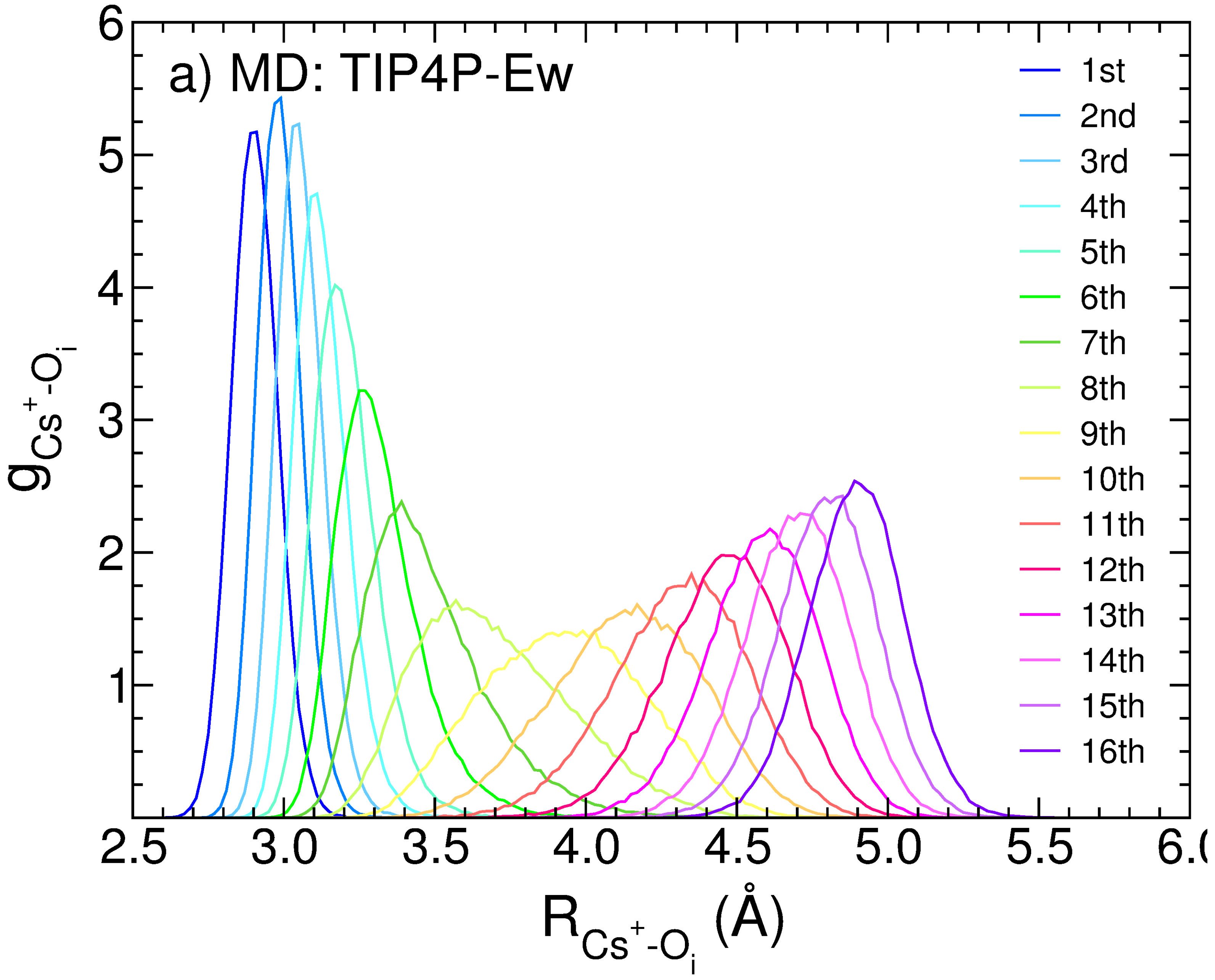
Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Many-Body Effects in Cs+ Hydration

Cs+ hydration: Experiment vs. Simulation

point charge model

incremental RDFs

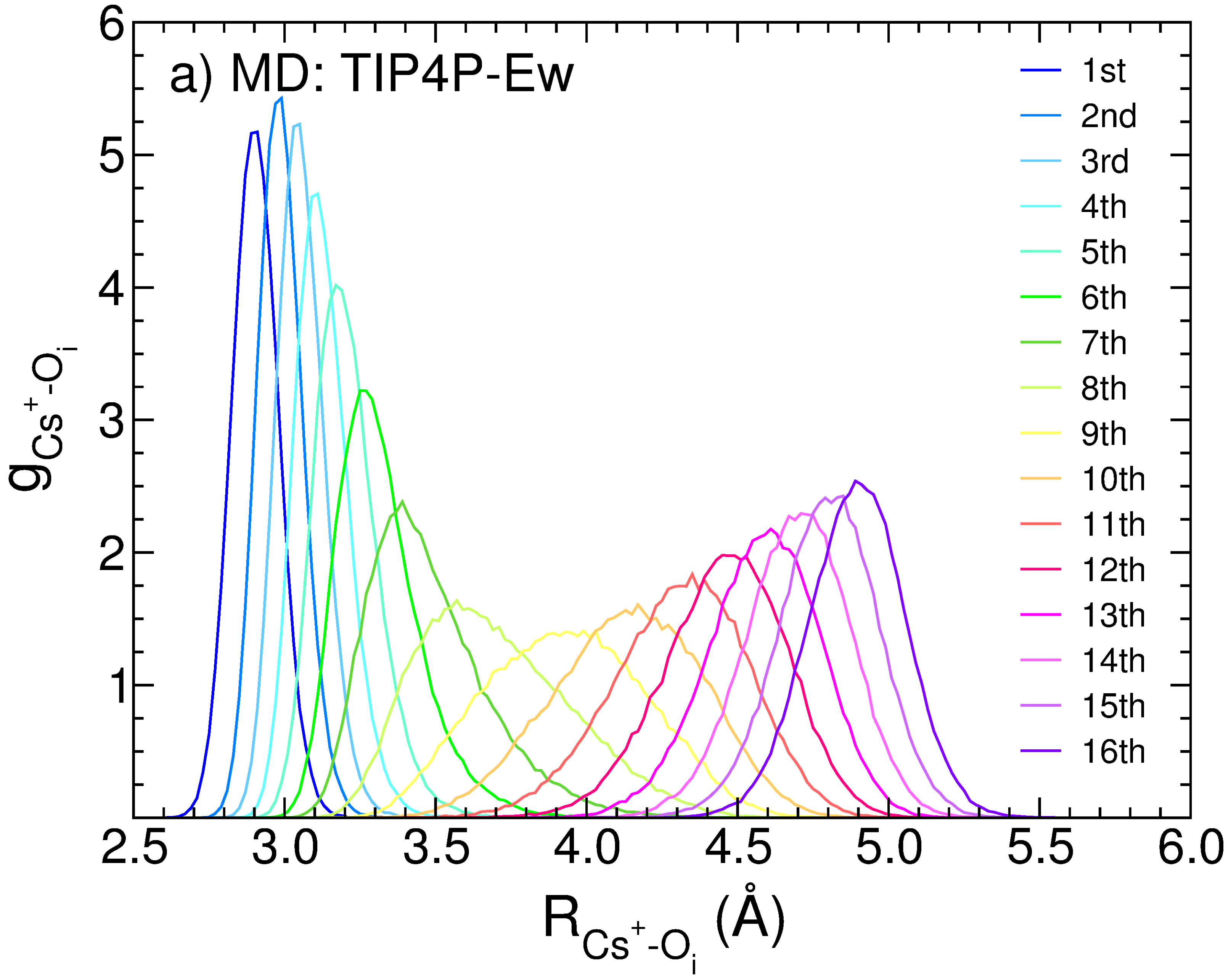


Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

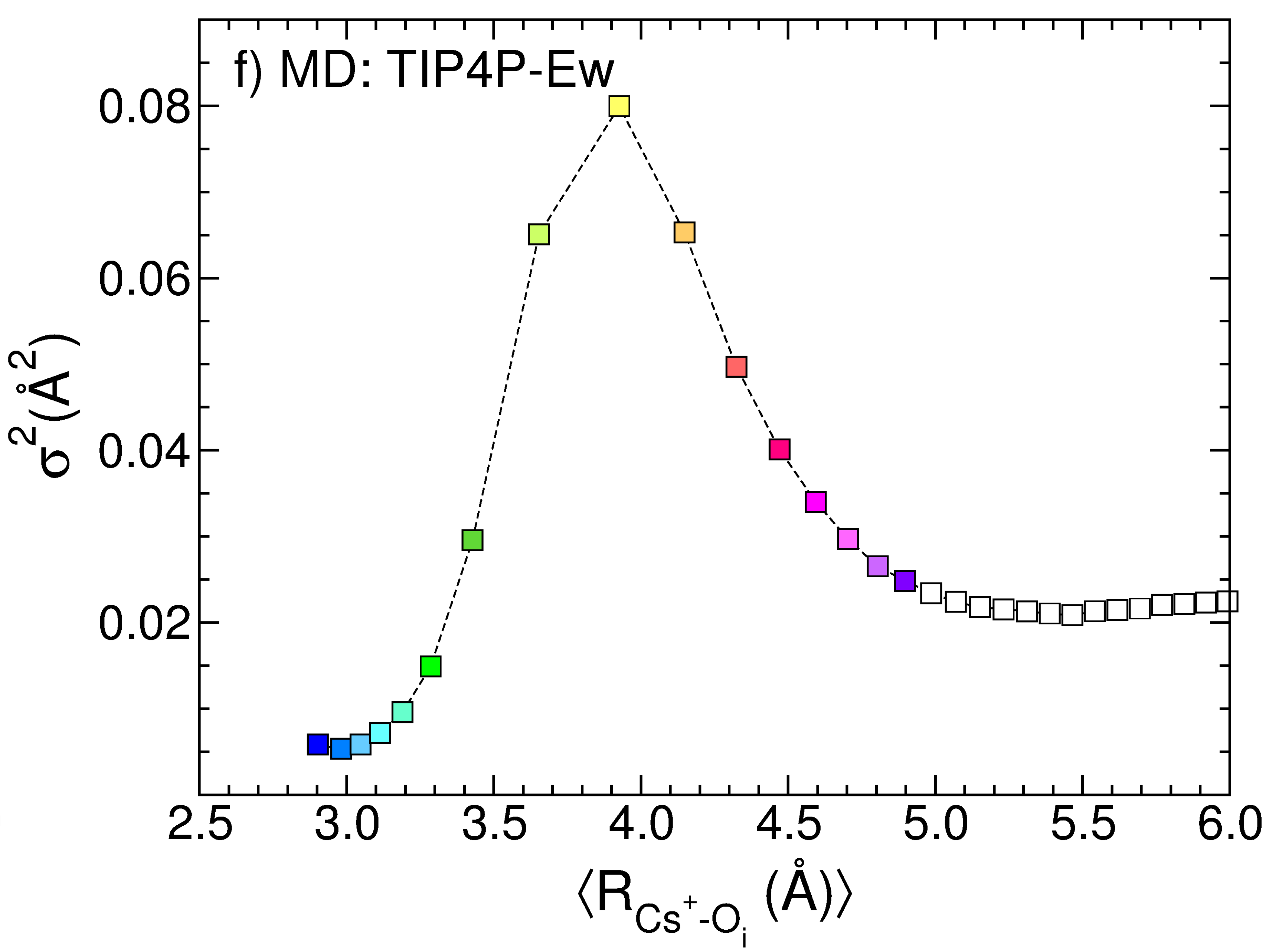
Many-Body Effects in Cs^+ Hydration

Cs^+ hydration: Experiment vs. Simulation

point charge model



incremental RDFs



average distance and variance

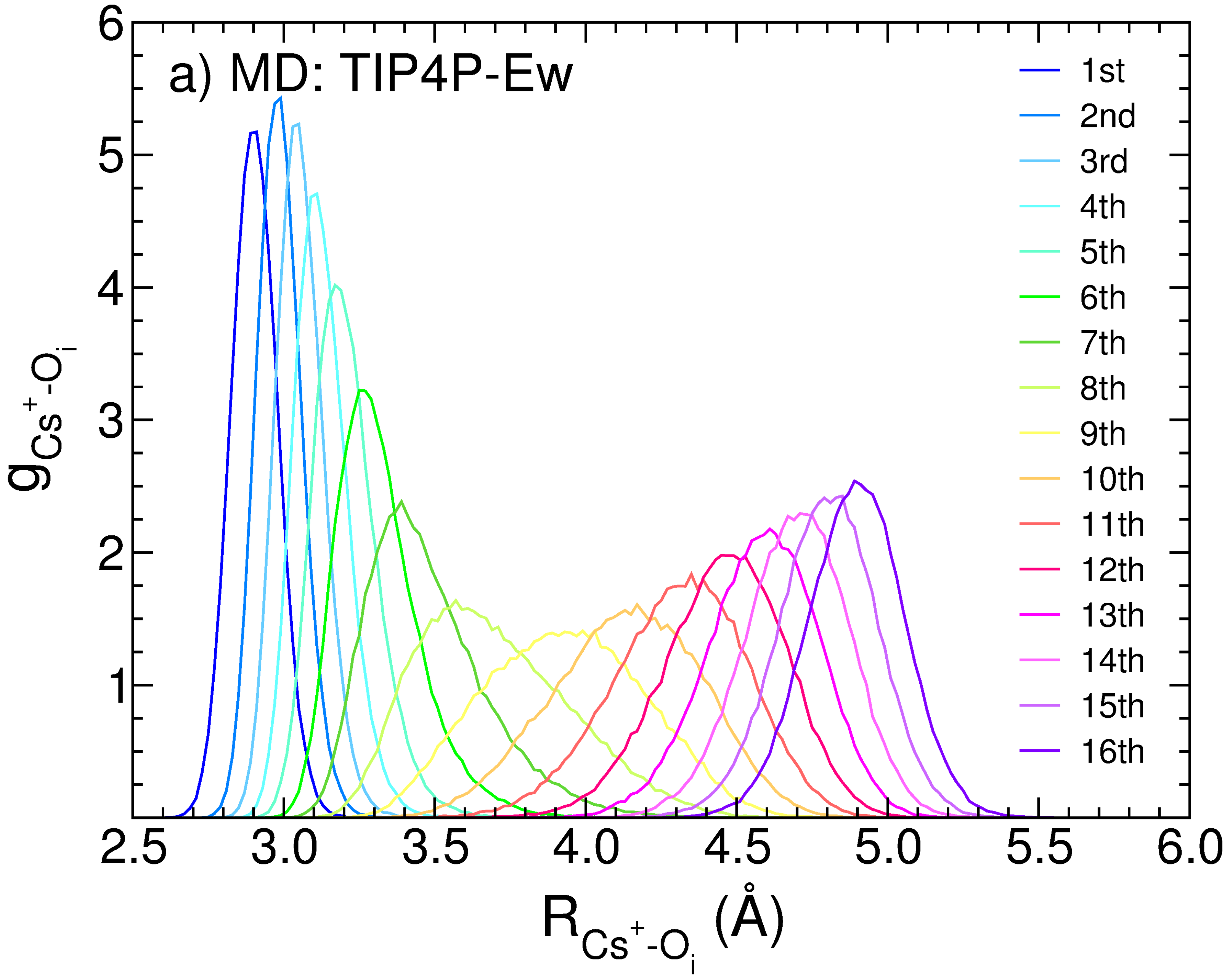
Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. *10*, *406 (2019)*

Many-Body Effects in Cs+ Hydration

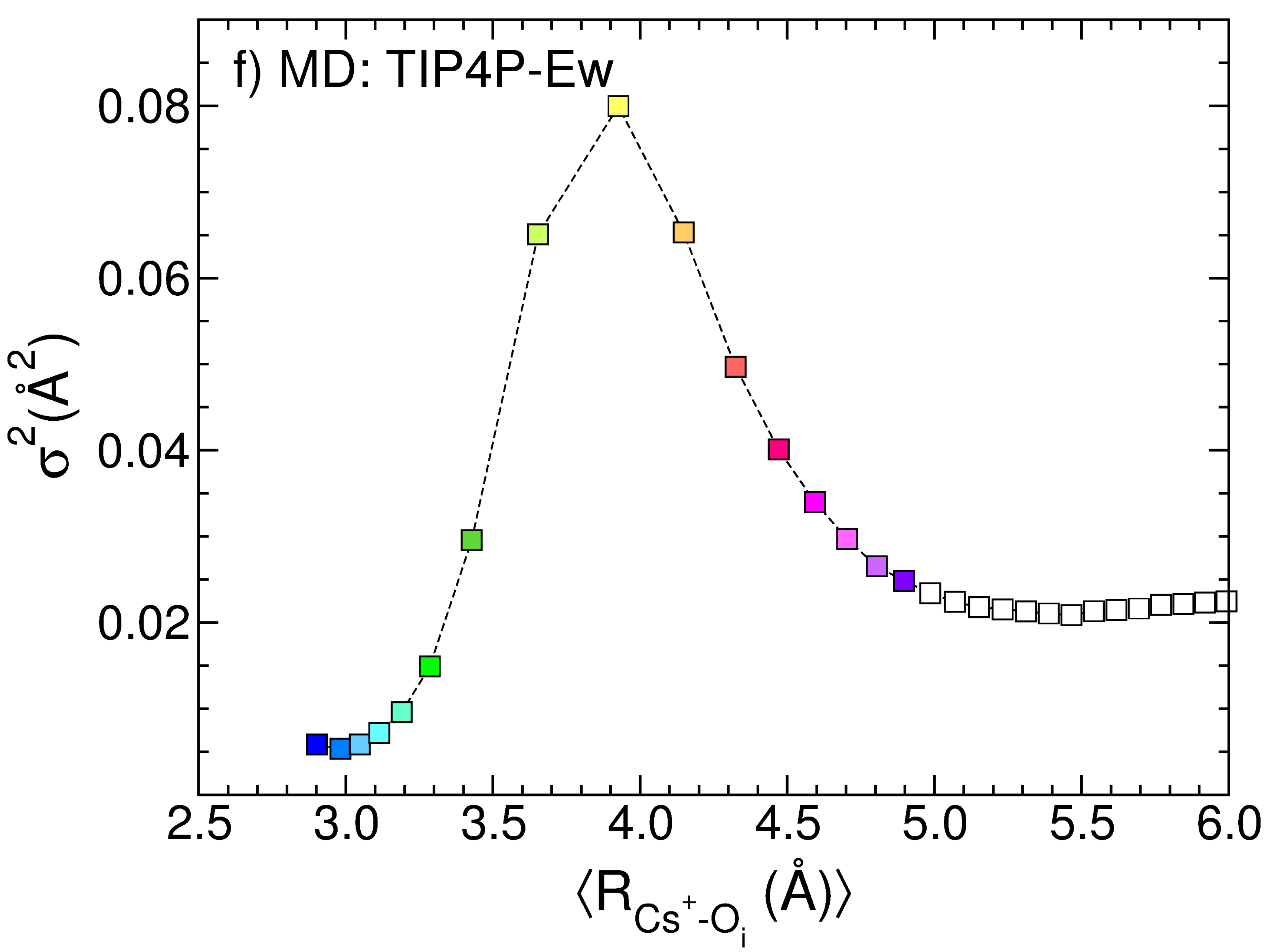
Cs+ hydration: Experiment vs. Simulation

point charge model

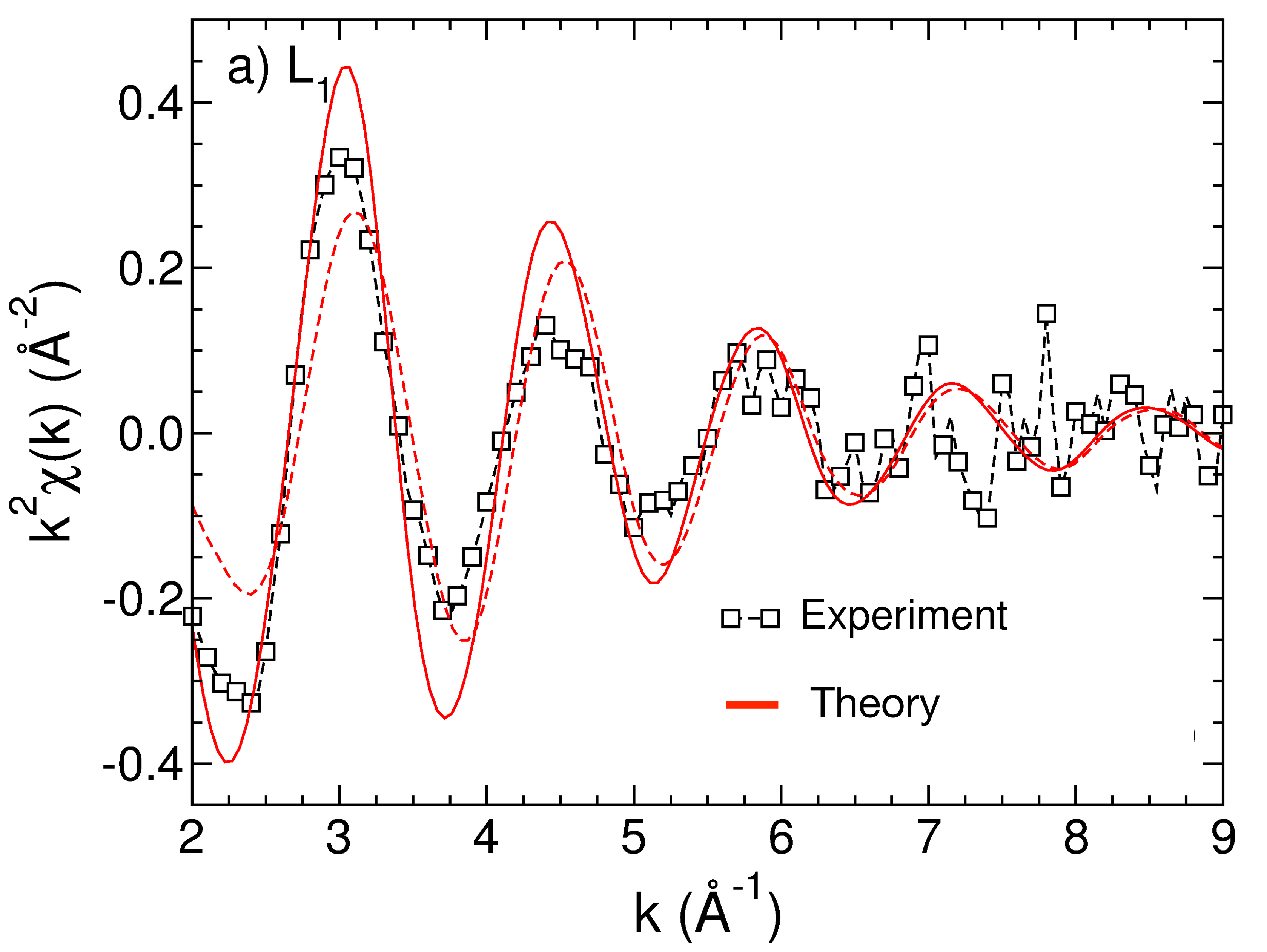
incremental RDFs



average distance and variance



EXAFS spectrum



Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Many-Body Effects in Cs+ Hydration

Cs+ hydration: Experiment vs. Simulation

polarizable model

incremental RDFs

average distance and variance

EXAFS spectrum

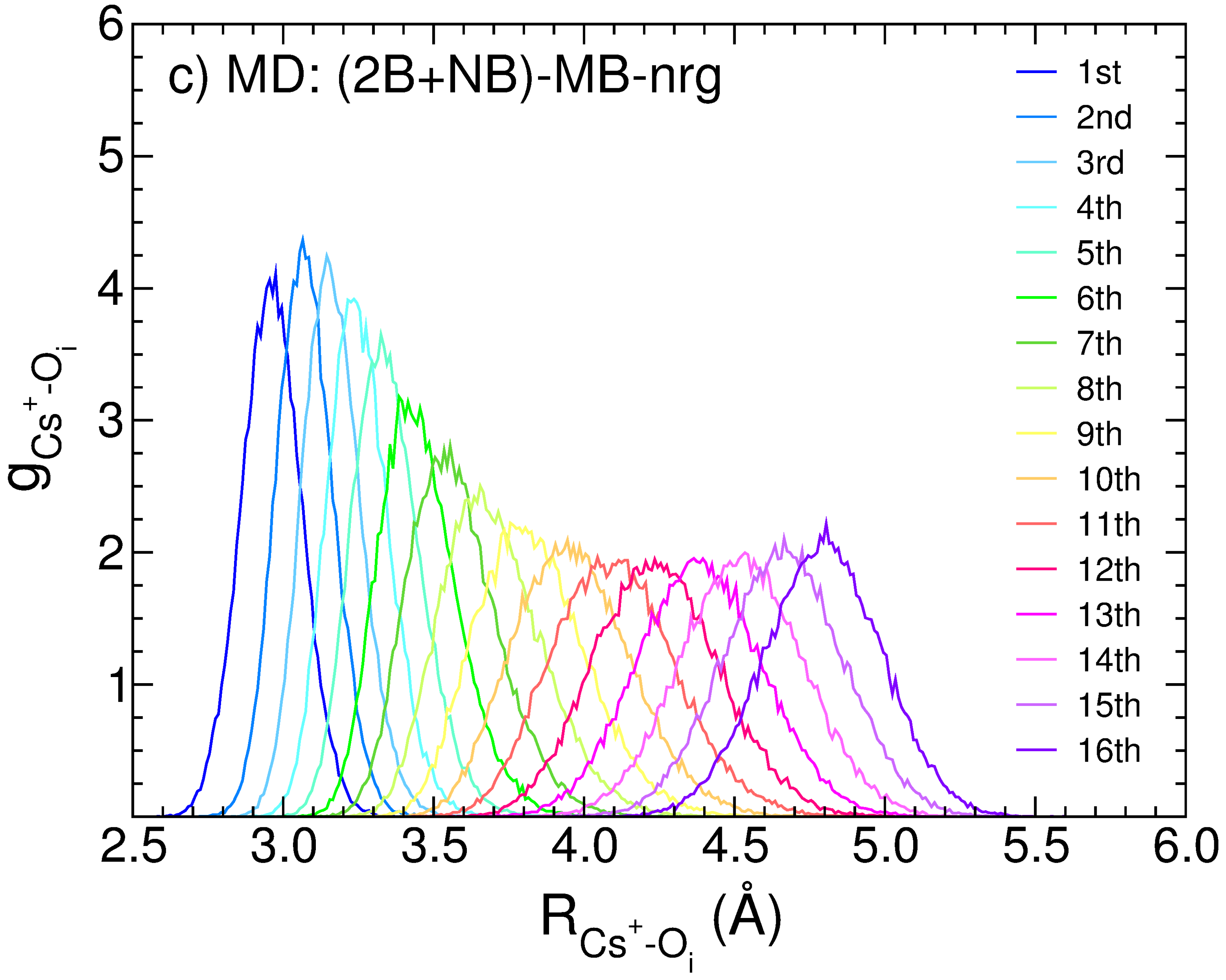
Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. <u>*10*</u>*, 406 (2019)*

Many-Body Effects in Cs^+ Hydration

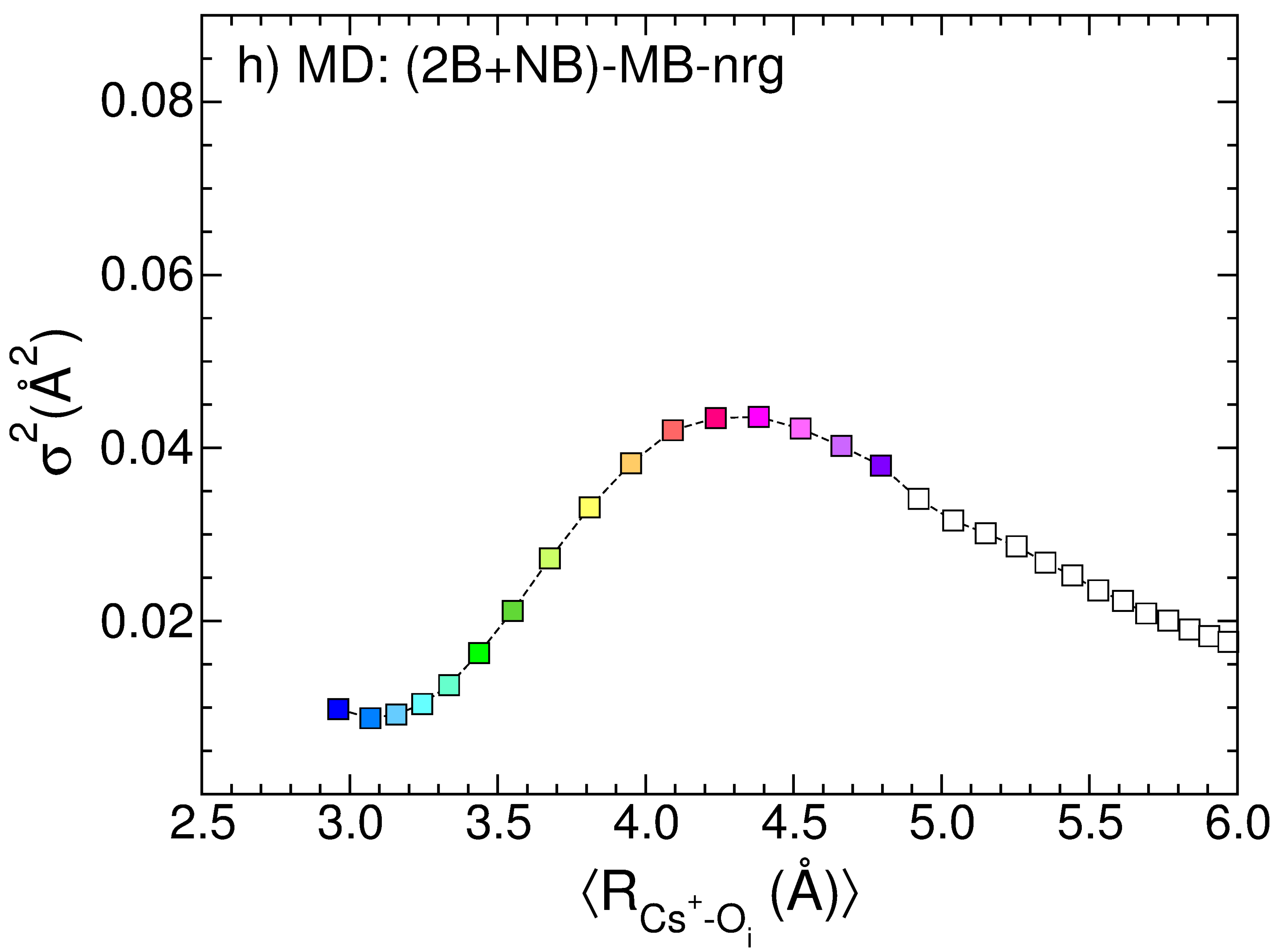
Cs^+ hydration: Experiment vs. Simulation

many-body model with explicit 2B term

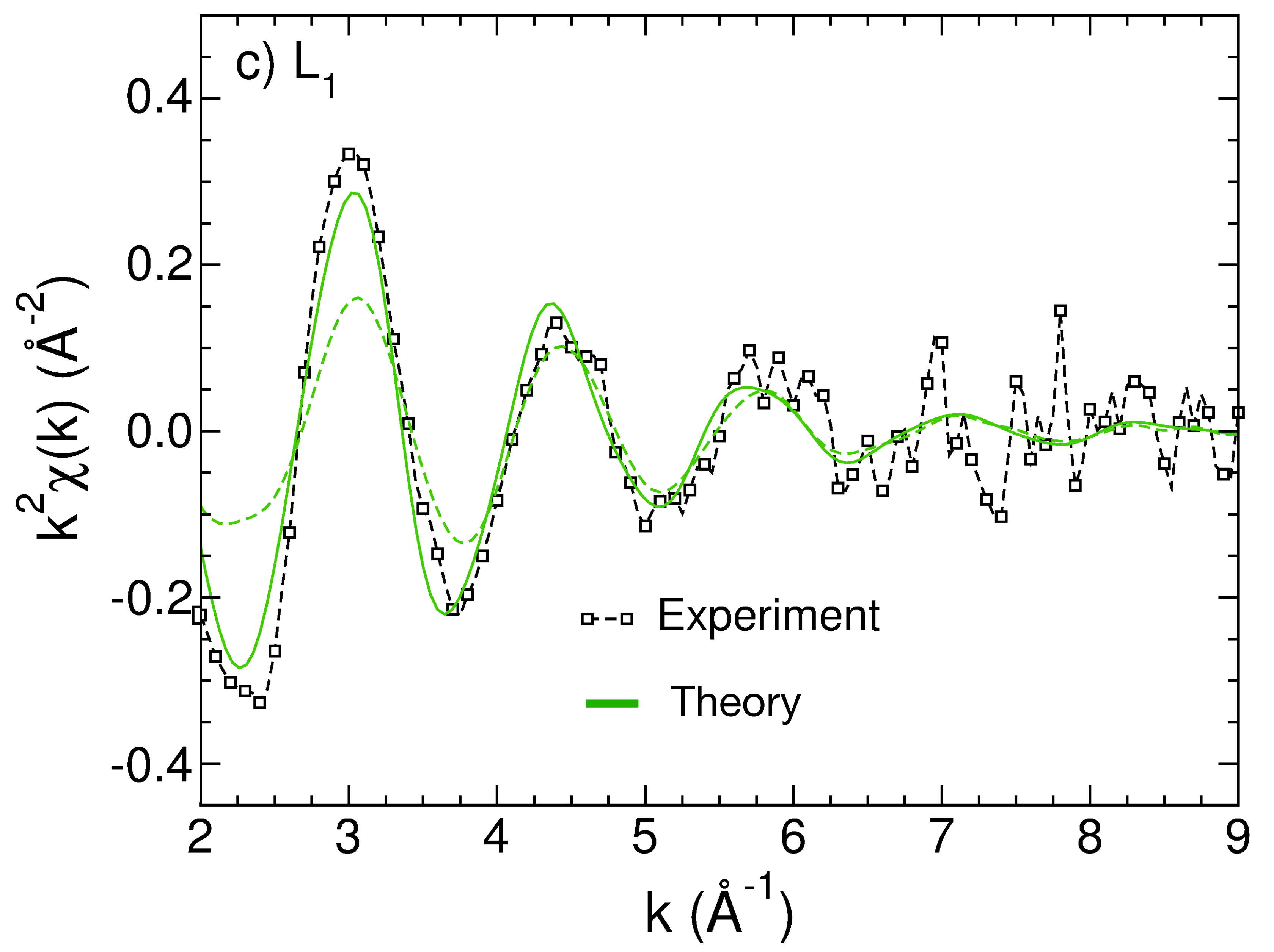
incremental RDFs



average distance and variance



EXAFS spectrum

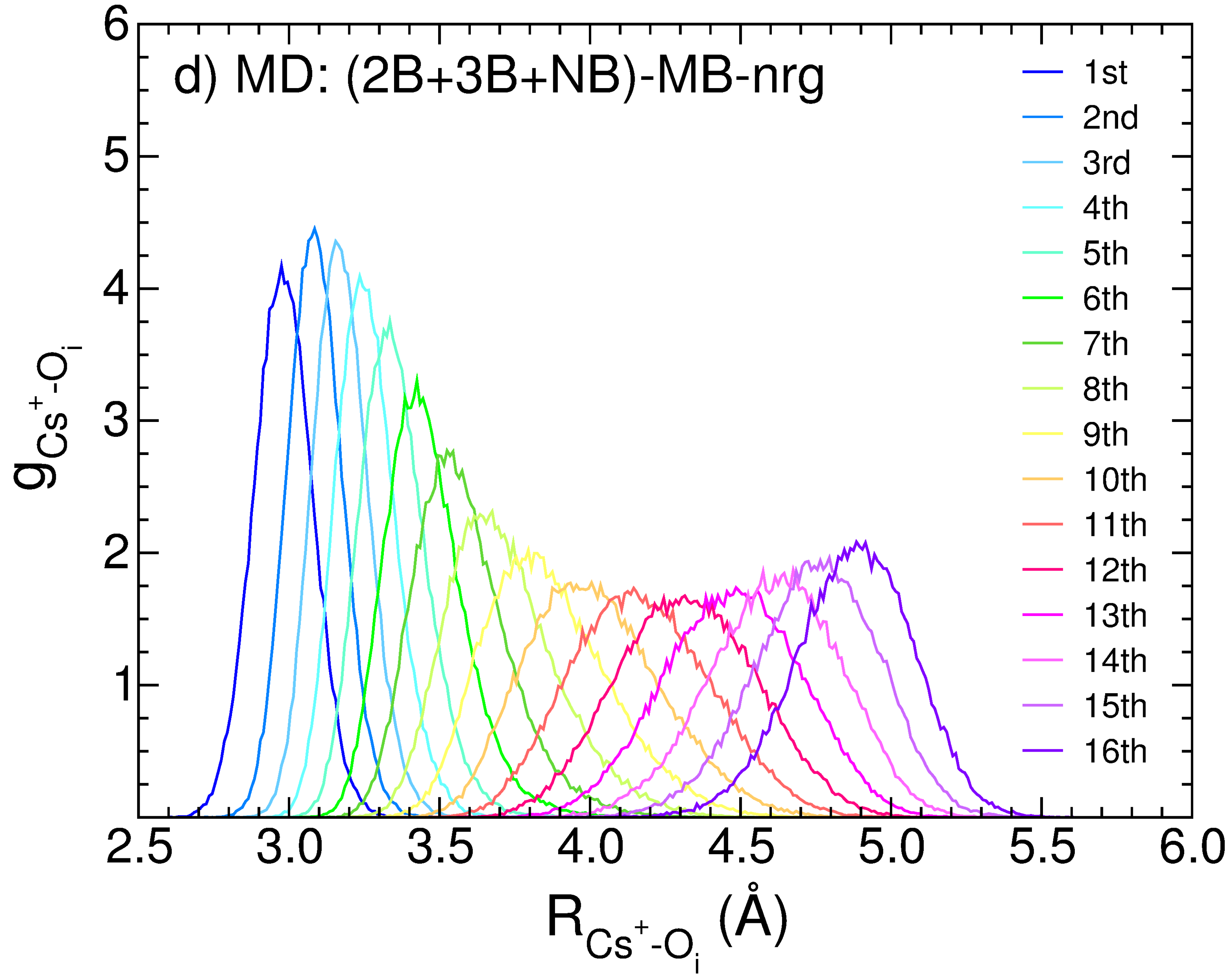


Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

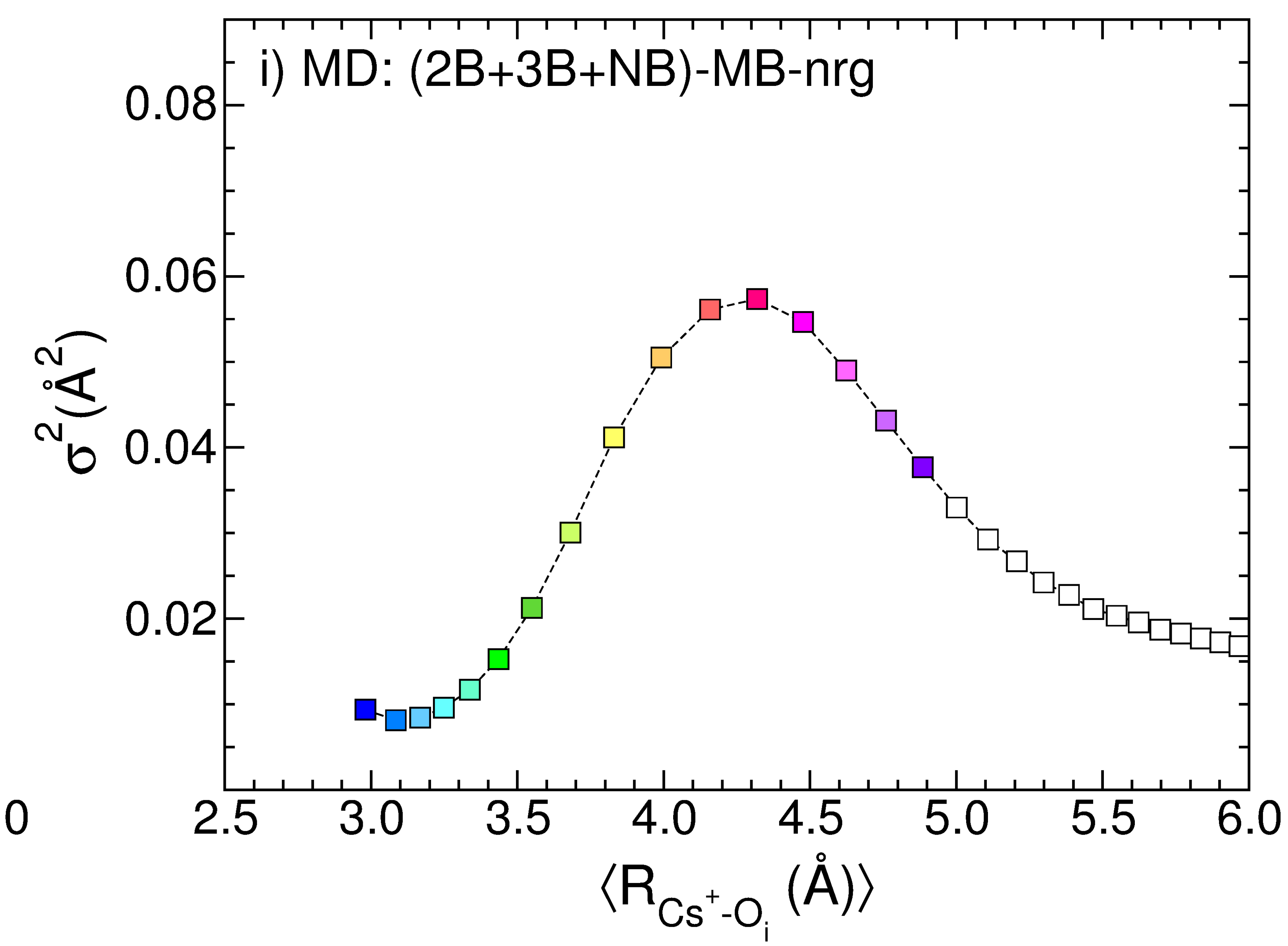
Many-Body Effects in Cs+ Hydration

Cs^+ hydration: Experiment vs. Simulation

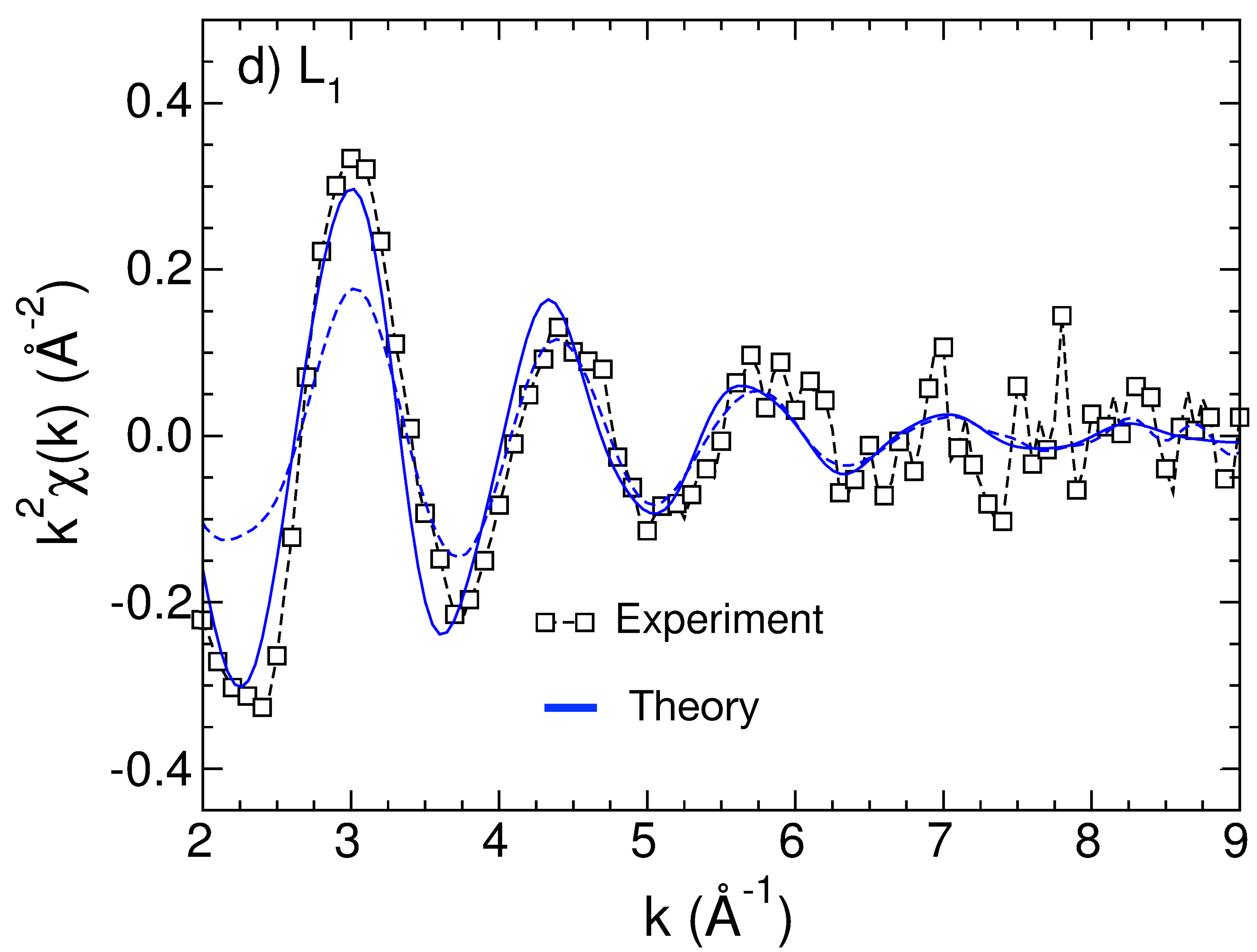
many-body model with explicit 2B & 3B terms



incremental RDFs



average distance and variance



EXAFS spectrum

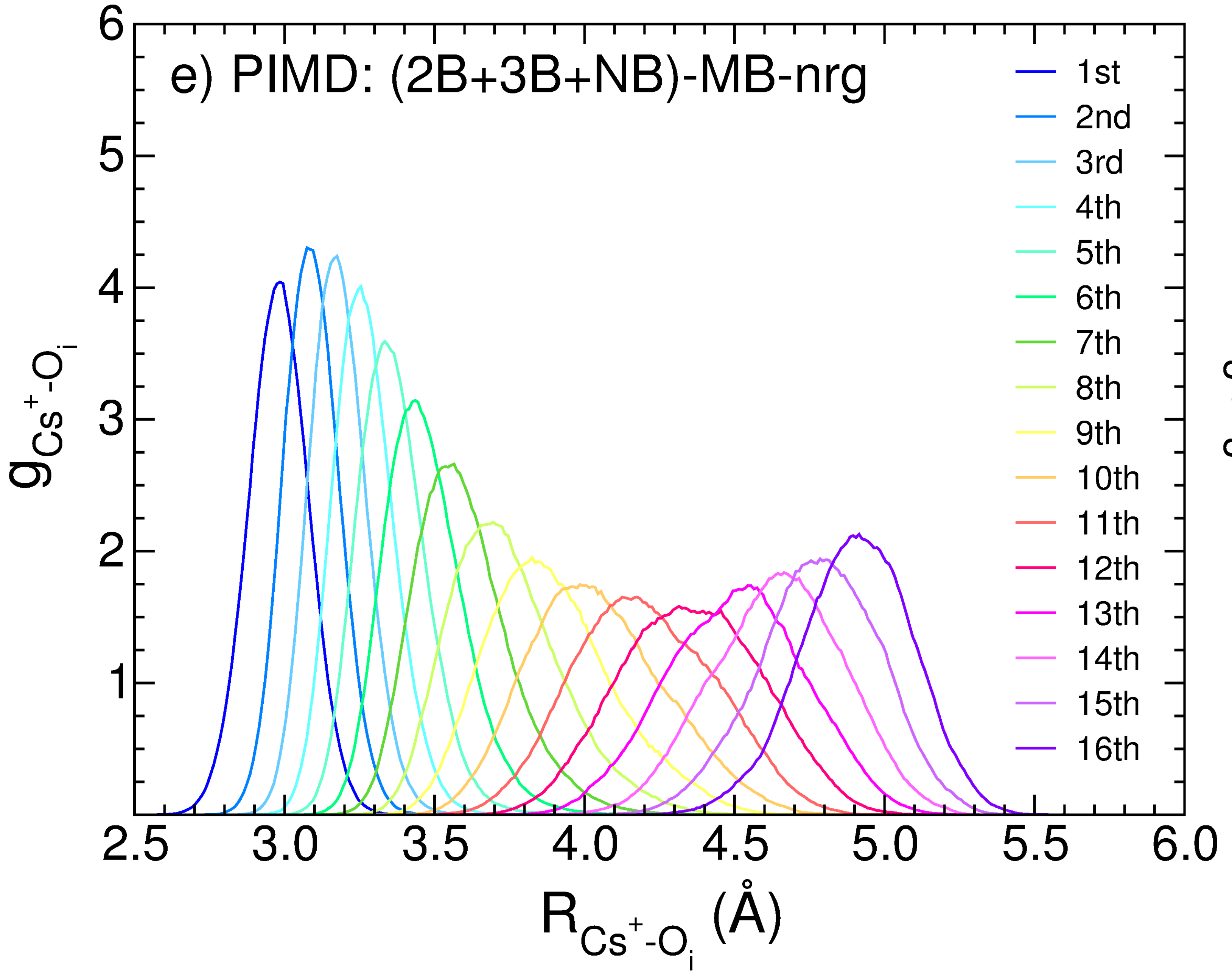
Zhuang, Riera, Schenter, Fulton & FP, *J. Phys. Chem. Lett.* 10, 406 (2019)

Many-Body Effects in Cs$^+$ Hydration

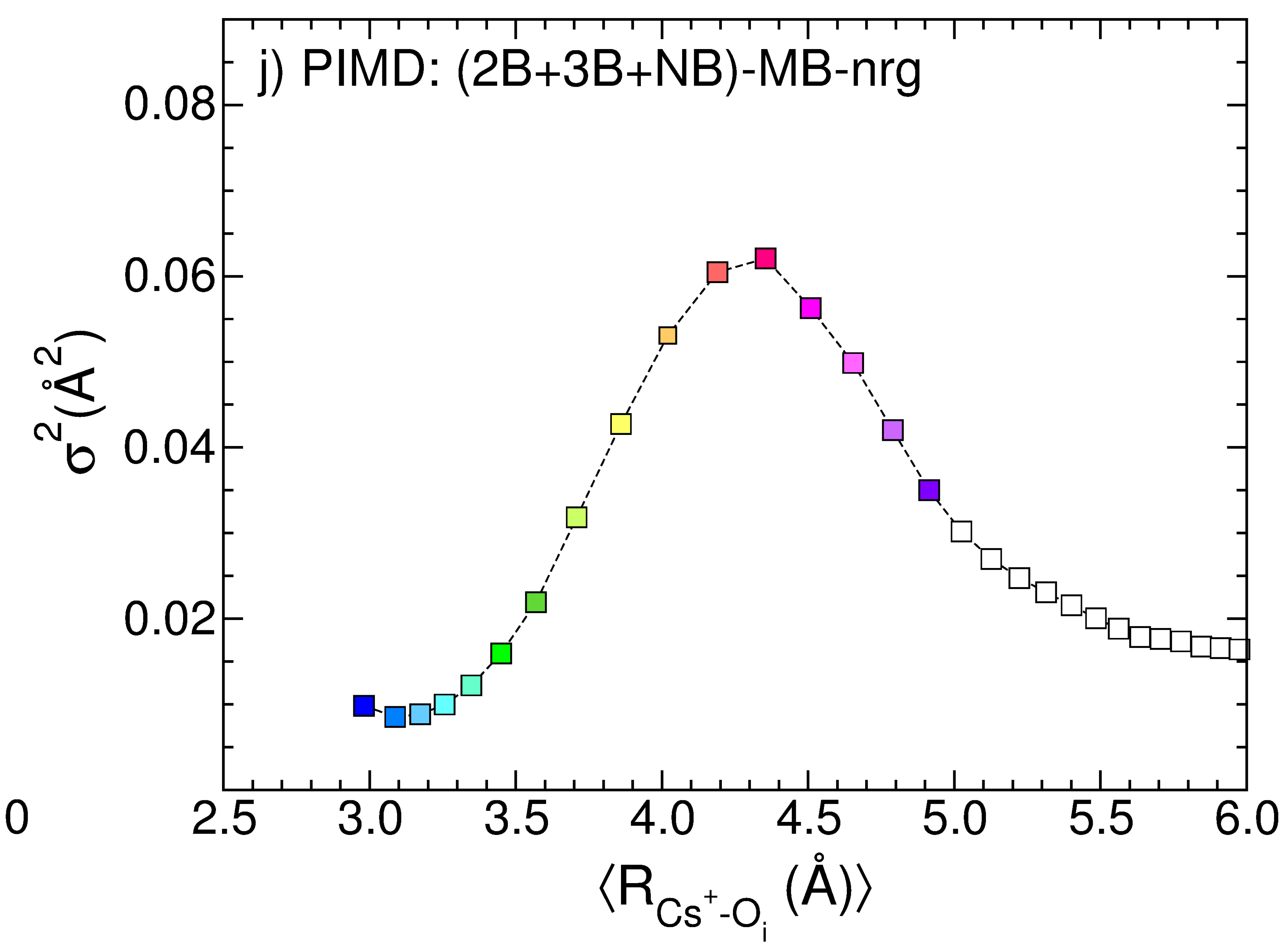
Cs$^+$ hydration: Experiment vs. Simulation

many-body model with explicit 2B & 3B terms and nuclear quantum effects

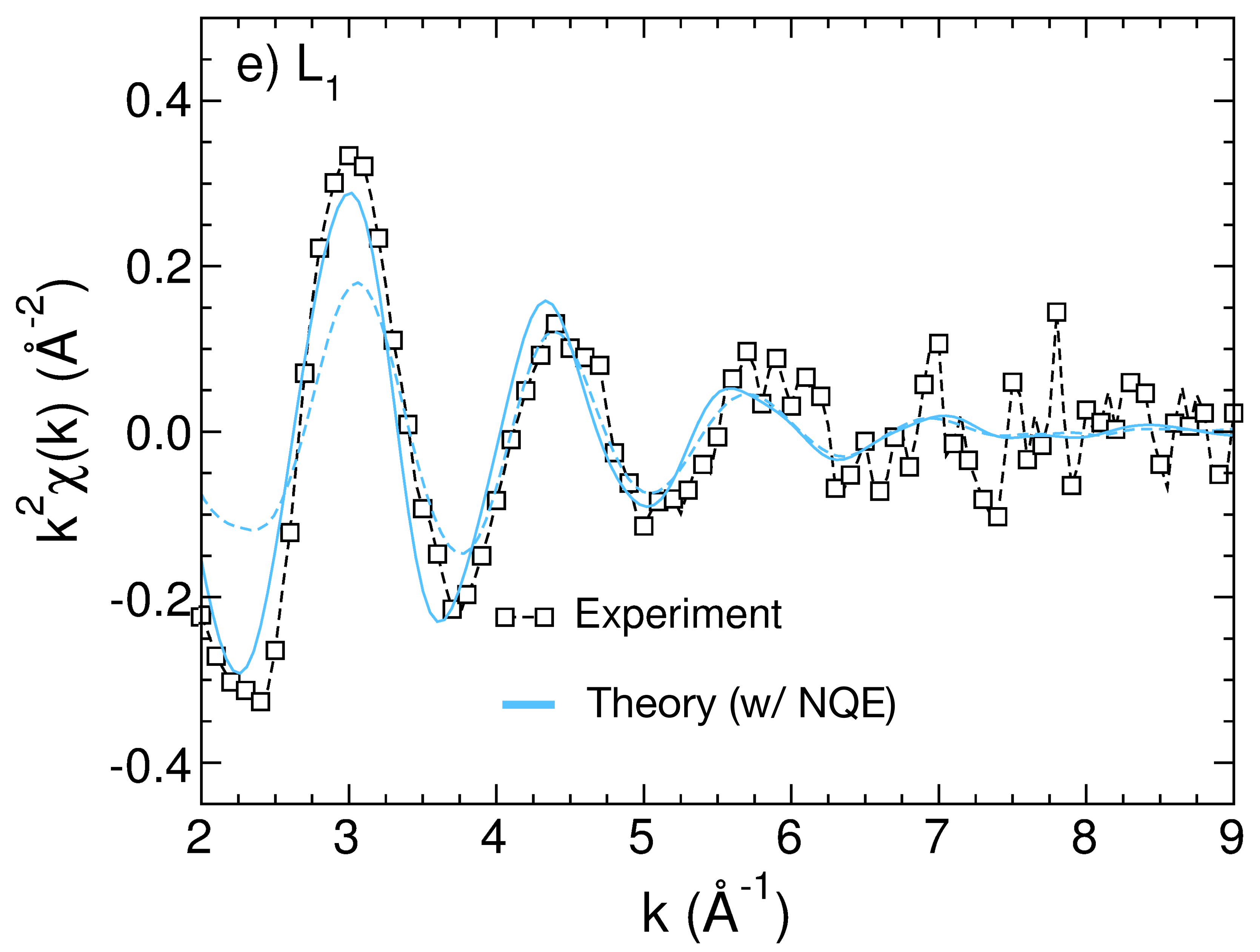
incremental RDFs



average distance and variance



EXAFS spectrum

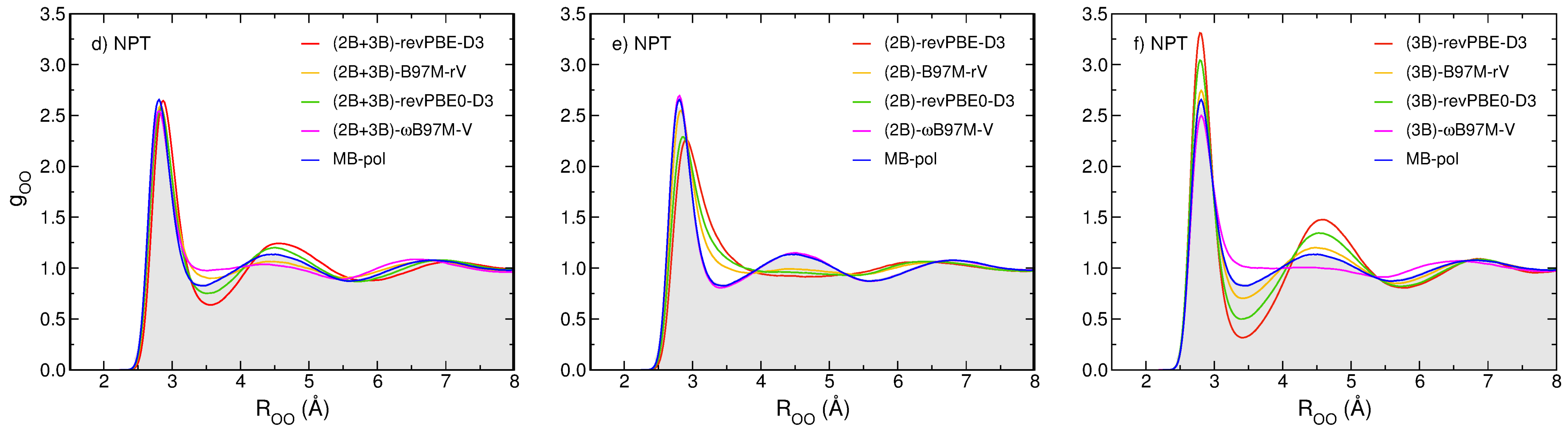


Zhuang, Riera, Schenter, Fulton & FP, J. Phys. Chem. Lett. 10, 406 (2019)

Learning Physics from Many-Body Models

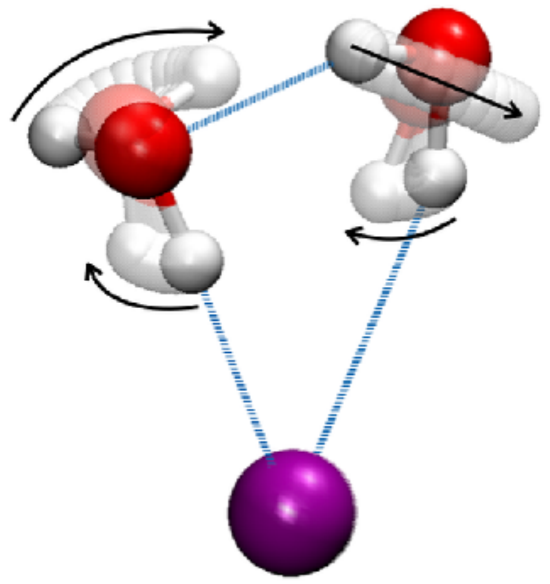
What makes water water?

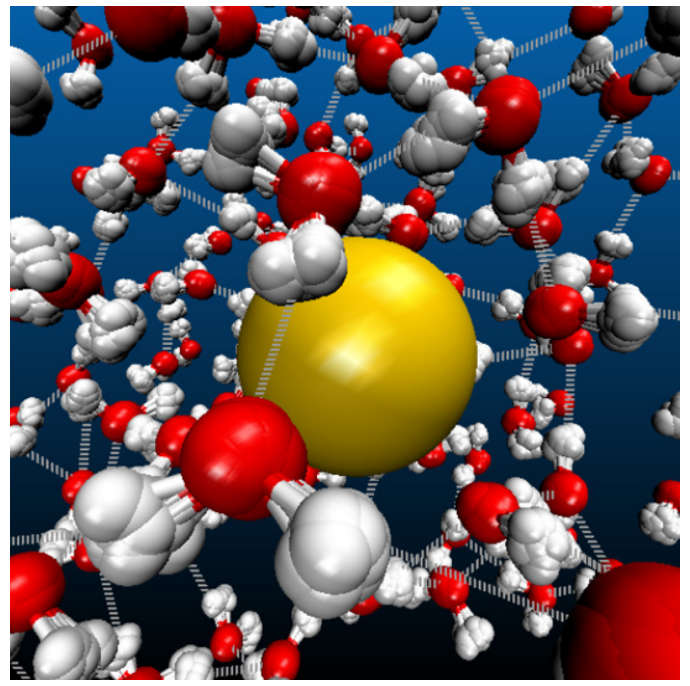
- Low-order many-body interactions determine the local structure of water
- Delocalization errors make current DFT models unreliable for water

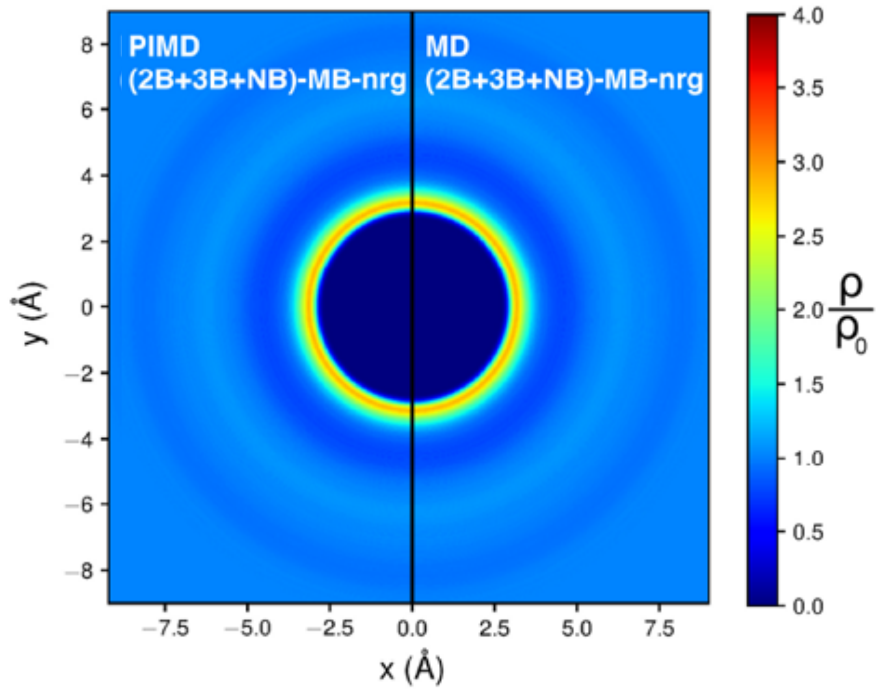


How good are simulations of hydration?

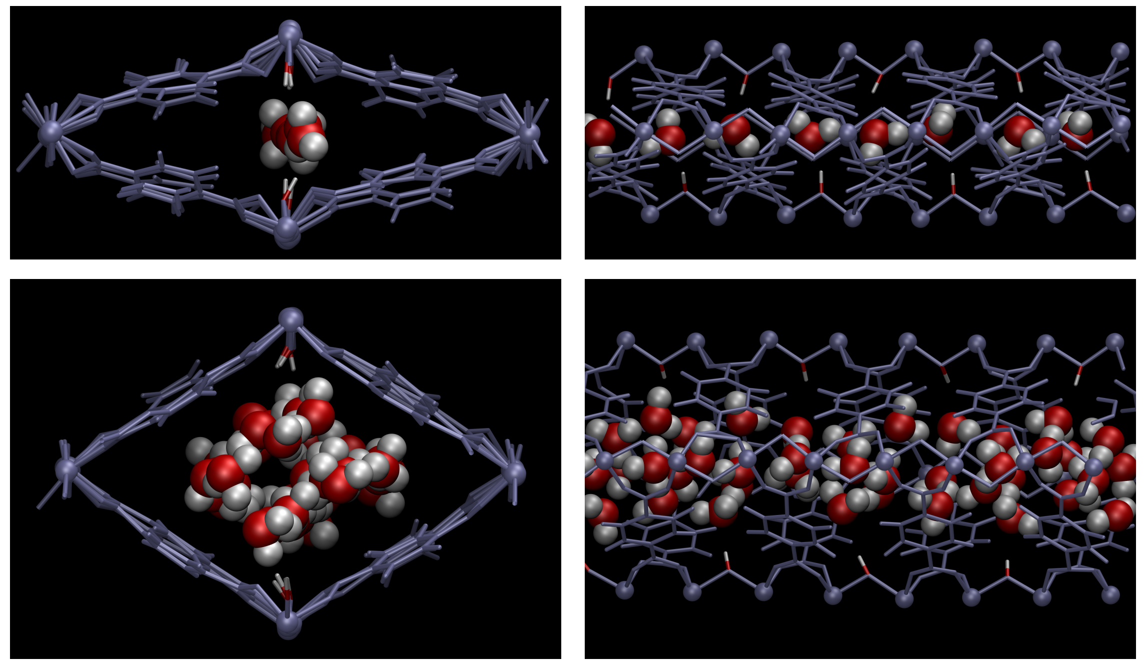
- Many-body interactions at both short and long ranges determine the hydration structure of monovalent ions
- Common force fields are unable to correctly predict the hydration structure of monovalent ions

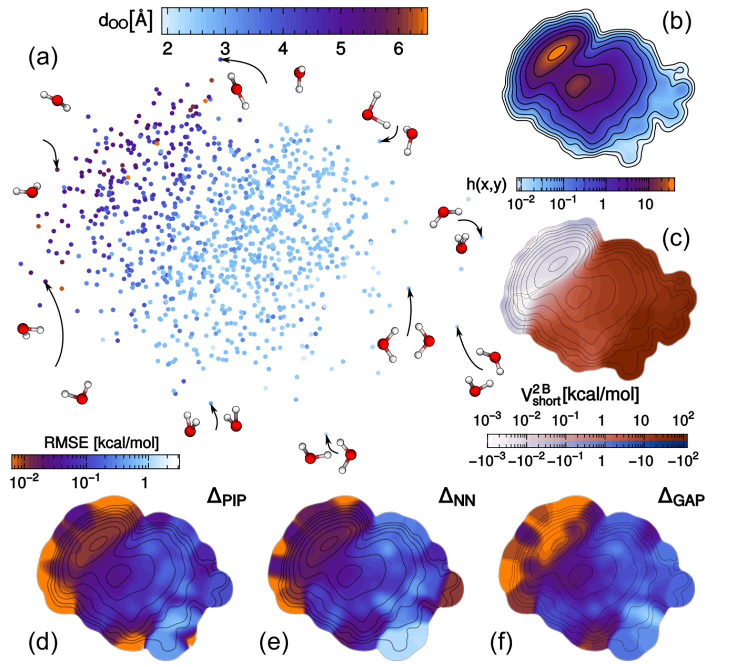






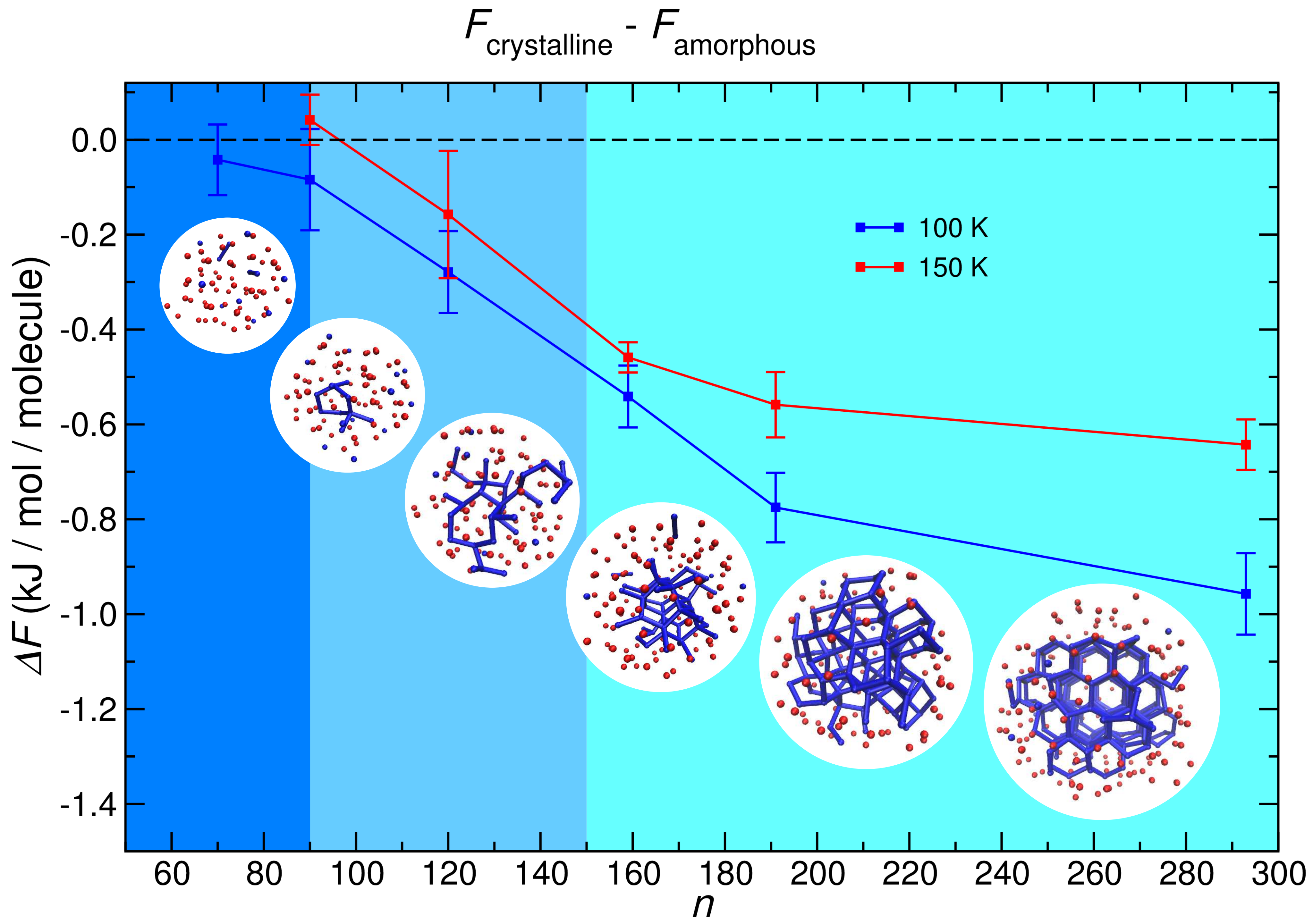
Data-driven many-body models: Quo vadis?





From Water to Ice…

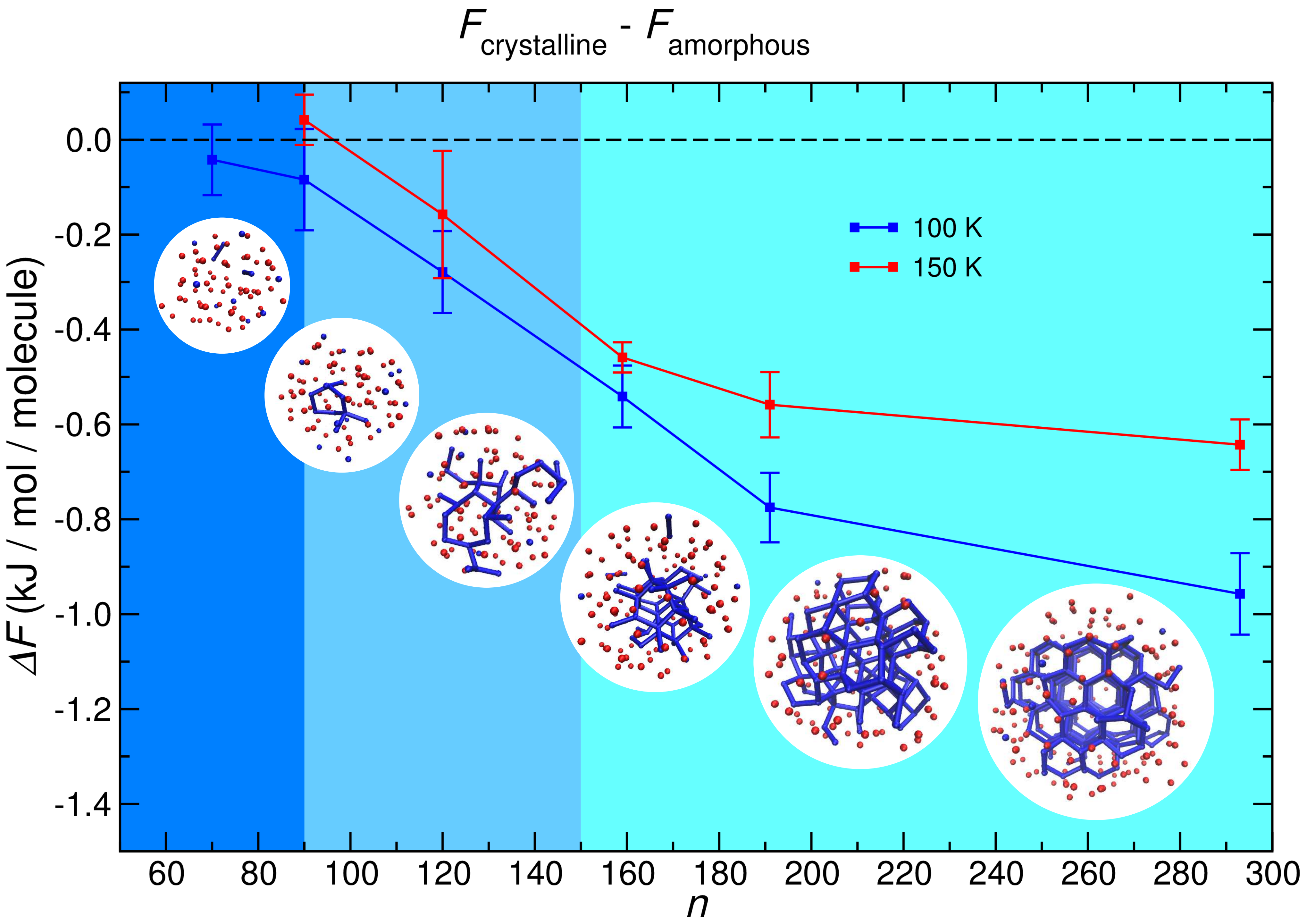
Homogeneous ice nucleation



Moberg, Becker, Dierking, Zurheide, Bandow, Buck, Hudait, Molinero, FP & Zeuch, Proc. Natl. Acad. Sci. U.S.A., in press
https://doi.org/10.26434/chemrxiv.8856053.v1

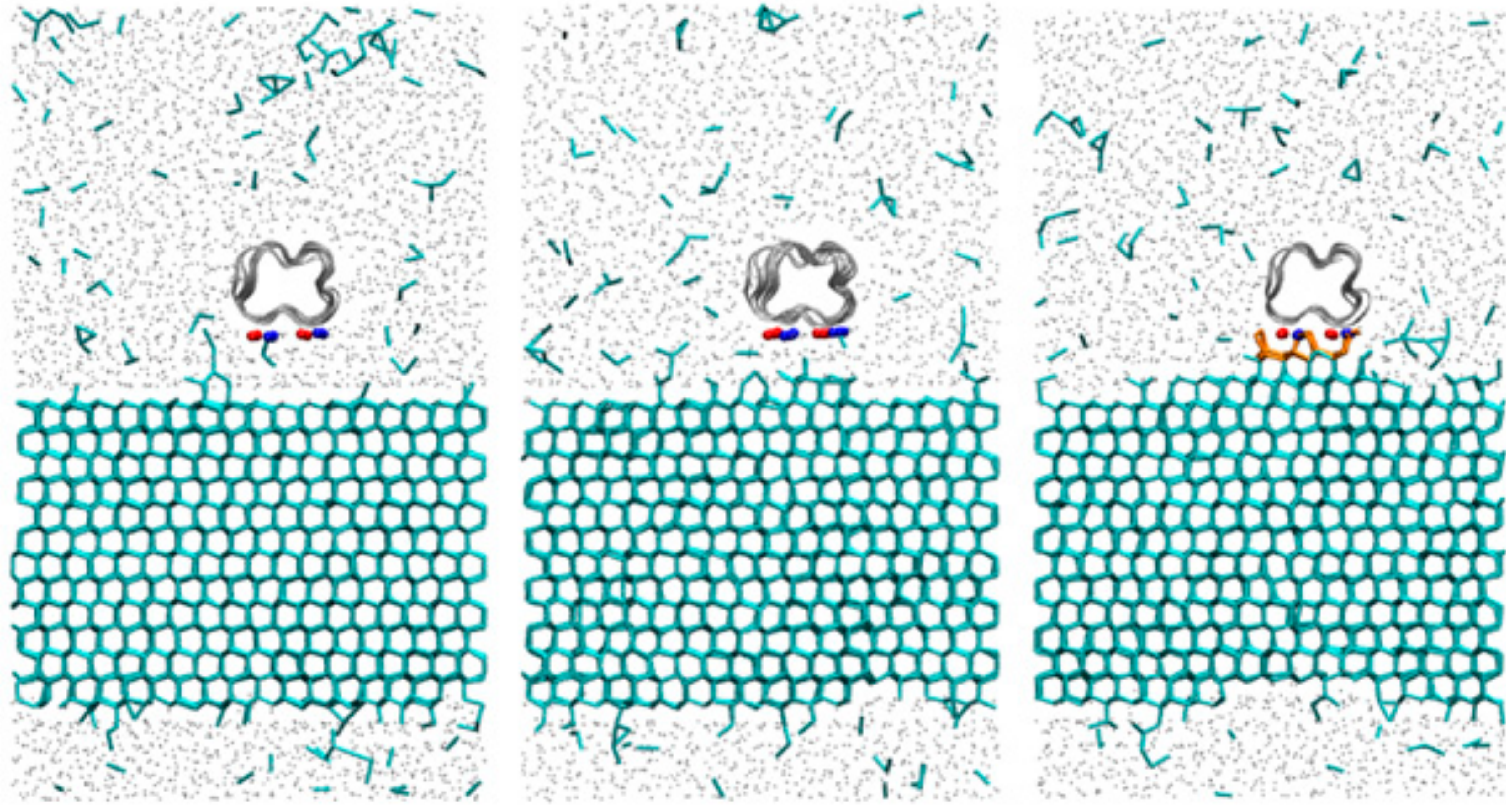
From Water to Ice…

Homogeneous ice nucleation



Moberg, Becker, Dierking, Zurheide, Bandow, Buck, Hudait, Molinero, FP & Zeuch, Proc. Natl. Acad. Sci. U.S.A., in press https://doi.org/10.26434/chemrxiv.8856053.v1

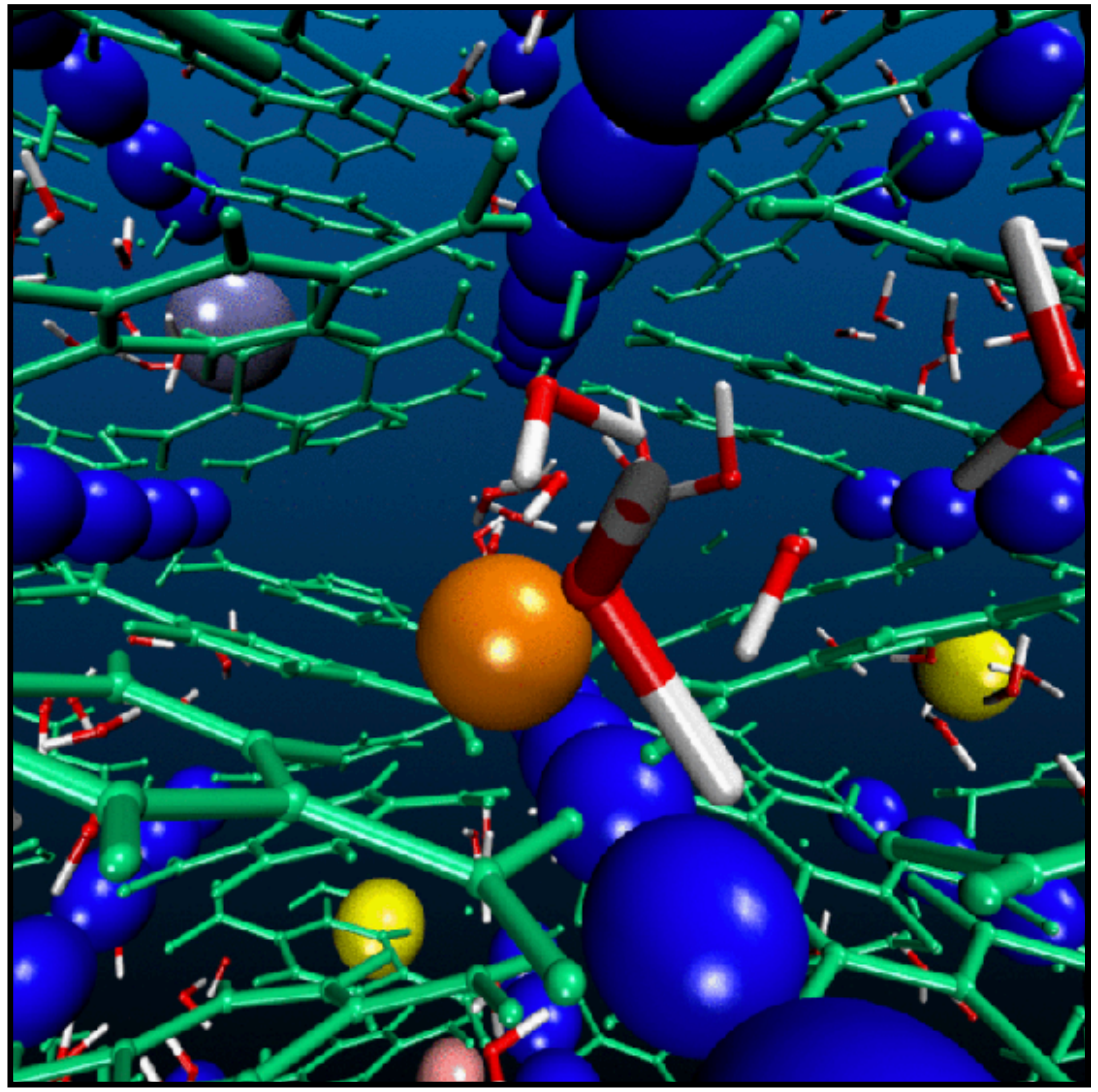
Heterogeneous ice nucleation



Hudait, Moberg, Qiu, Odendahl, FP & Molinero Proc. Natl. Acad. Sci. U.S.A. 115, 8266 (2018) ongoing collaboration with Peter Frazier (Cornell University)

… to Aqueous Solutions in Confinement and Fluid Mixtures

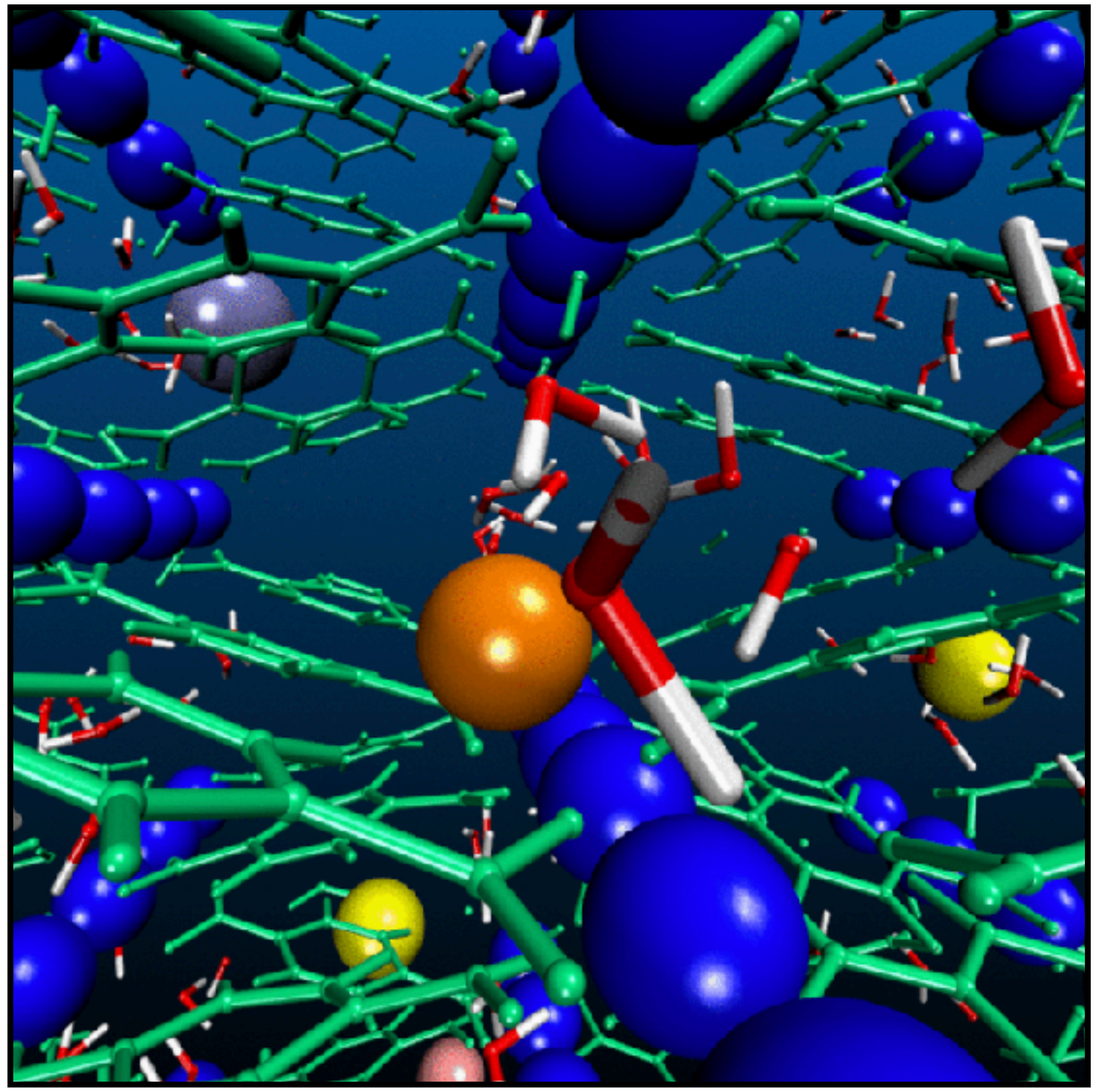
Water harvesting and desalination in MOFs



Rieth, Hunter, Dinča & FP, Nat. Commun. 10, 4771 (2019)
ongoing collaboration with Sicun Gao (UC San Diego)

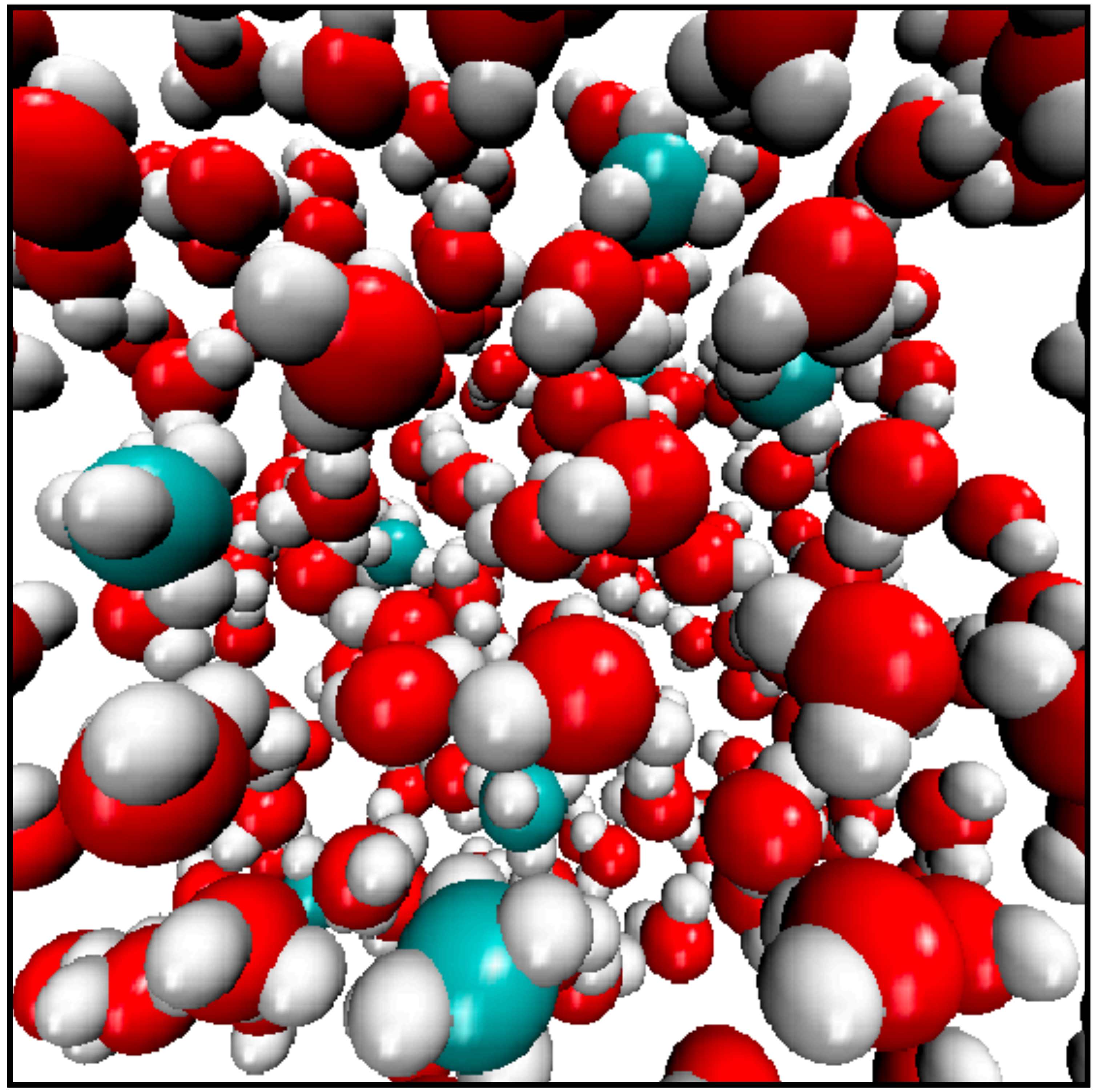
… to Aqueous Solutions in Confinement and Fluid Mixtures

Water harvesting and desalination in MOFs



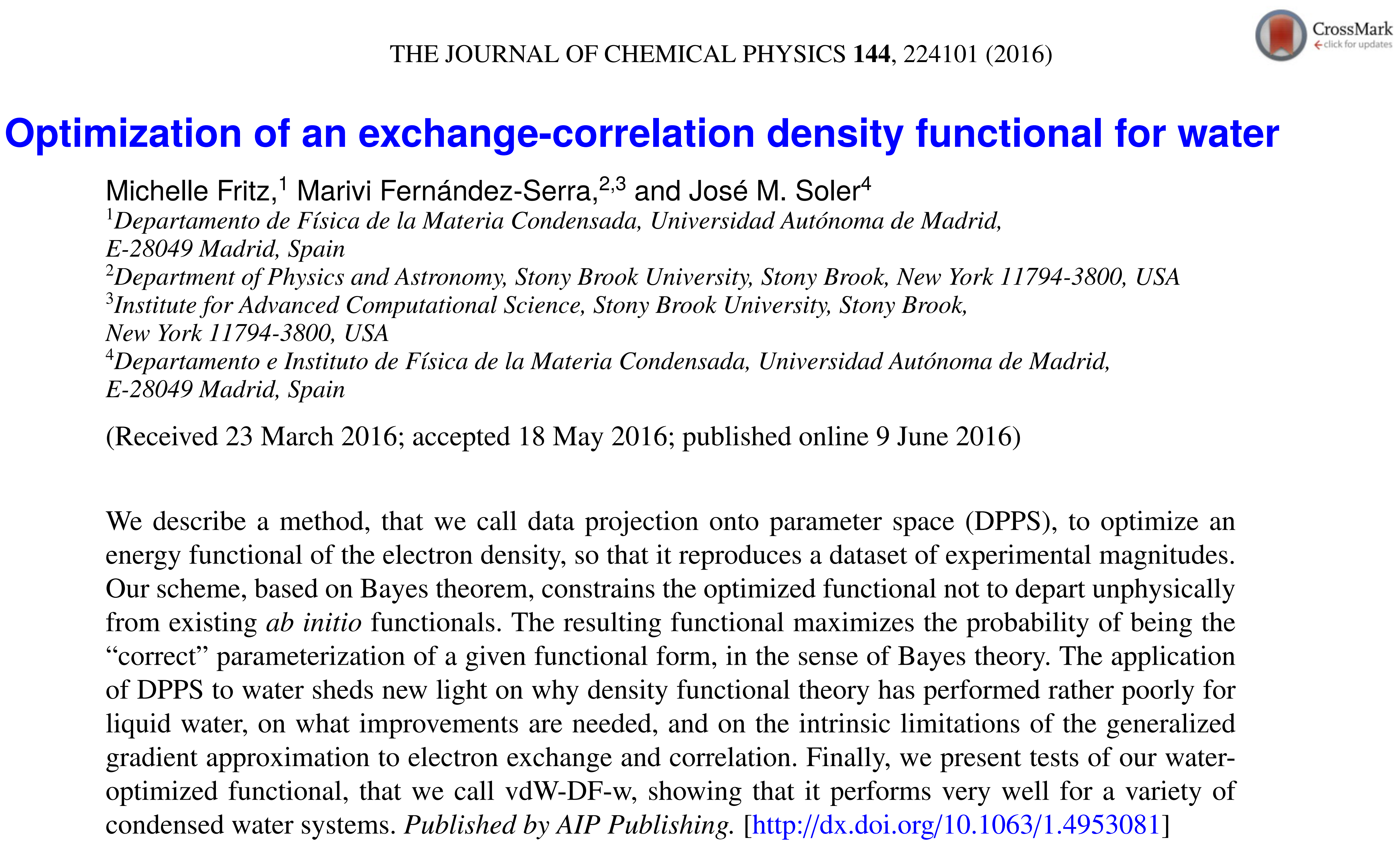
Rieth, Hunter, Dinča & FP, Nat. Commun. 10, 4771 (2019)
ongoing collaboration with Sicun Gao (UC San Diego)

Complex fluid mixtures



Riera, Hirales, Ghosh & FP, in progress

Many-Body Models as a Reference for ML DFT Models



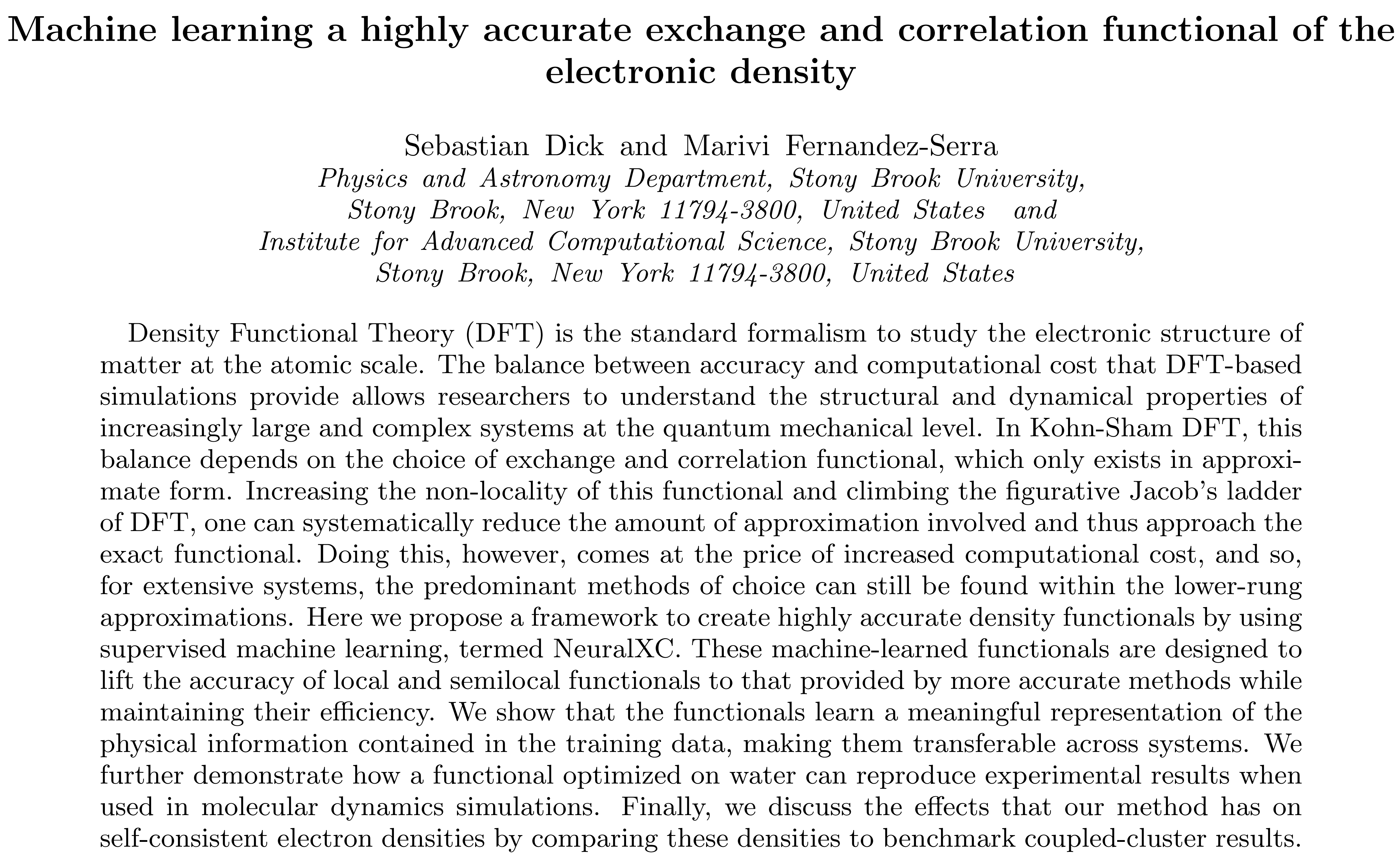
THE JOURNAL OF CHEMICAL PHYSICS **144**, 224101 (2016)

Optimization of an exchange-correlation density functional for water

Michelle Fritz,[1] Marivi Fernández-Serra,[2,3] and José M. Soler[4]
[1]*Departamento de Física de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
[2]*Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794-3800, USA*
[3]*Institute for Advanced Computational Science, Stony Brook University, Stony Brook, New York 11794-3800, USA*
[4]*Departamento e Instituto de Física de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*

(Received 23 March 2016; accepted 18 May 2016; published online 9 June 2016)

We describe a method, that we call data projection onto parameter space (DPPS), to optimize an energy functional of the electron density, so that it reproduces a dataset of experimental magnitudes. Our scheme, based on Bayes theorem, constrains the optimized functional not to depart unphysically from existing *ab initio* functionals. The resulting functional maximizes the probability of being the "correct" parameterization of a given functional form, in the sense of Bayes theory. The application of DPPS to water sheds new light on why density functional theory has performed rather poorly for liquid water, on what improvements are needed, and on the intrinsic limitations of the generalized gradient approximation to electron exchange and correlation. Finally, we present tests of our water-optimized functional, that we call vdW-DF-w, showing that it performs very well for a variety of condensed water systems. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4953081]



Machine learning a highly accurate exchange and correlation functional of the electronic density

Sebastian Dick and Marivi Fernandez-Serra
Physics and Astronomy Department, Stony Brook University,
Stony Brook, New York 11794-3800, United States and
Institute for Advanced Computational Science, Stony Brook University,
Stony Brook, New York 11794-3800, United States

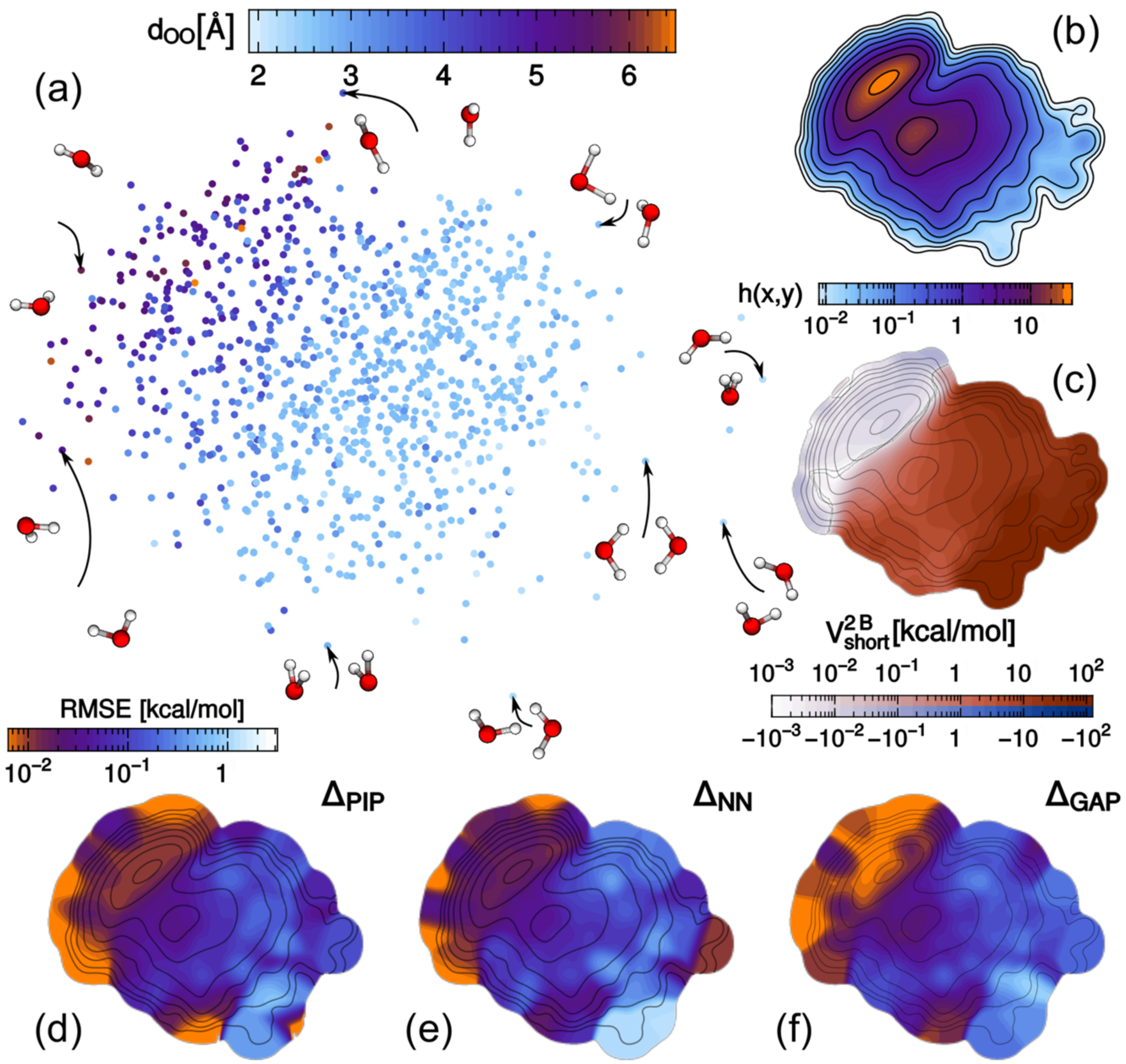
Density Functional Theory (DFT) is the standard formalism to study the electronic structure of matter at the atomic scale. The balance between accuracy and computational cost that DFT-based simulations provide allows researchers to understand the structural and dynamical properties of increasingly large and complex systems at the quantum mechanical level. In Kohn-Sham DFT, this balance depends on the choice of exchange and correlation functional, which only exists in approximate form. Increasing the non-locality of this functional and climbing the figurative Jacob's ladder of DFT, one can systematically reduce the amount of approximation involved and thus approach the exact functional. Doing this, however, comes at the price of increased computational cost, and so, for extensive systems, the predominant methods of choice can still be found within the lower-rung approximations. Here we propose a framework to create highly accurate density functionals by using supervised machine learning, termed NeuralXC. These machine-learned functionals are designed to lift the accuracy of local and semilocal functionals to that provided by more accurate methods while maintaining their efficiency. We show that the functionals learn a meaningful representation of the physical information contained in the training data, making them transferable across systems. We further demonstrate how a functional optimized on water can reproduce experimental results when used in molecular dynamics simulations. Finally, we discuss the effects that our method has on self-consistent electron densities by comparing these densities to benchmark coupled-cluster results.

"As prior information for the Bayesian optimization, we have used the 15 enhancement factors shown in Fig. 1. For the reference energies, we use the MB-pol force field, a sophisticated fit to highly accurate quantum chemical calculations of water monomers and clusters."

"It is important to emphasize that, although fitted only to monomers, dimers, and trimers, the MB-pol force field has been shown to reproduce accurately the structural and thermodynamic properties of condensed phases of water."

Representability of Low-Order Many-Body Interactions

2B interaction energies



sketch map

training point density

distribution of 2B energies in the training set

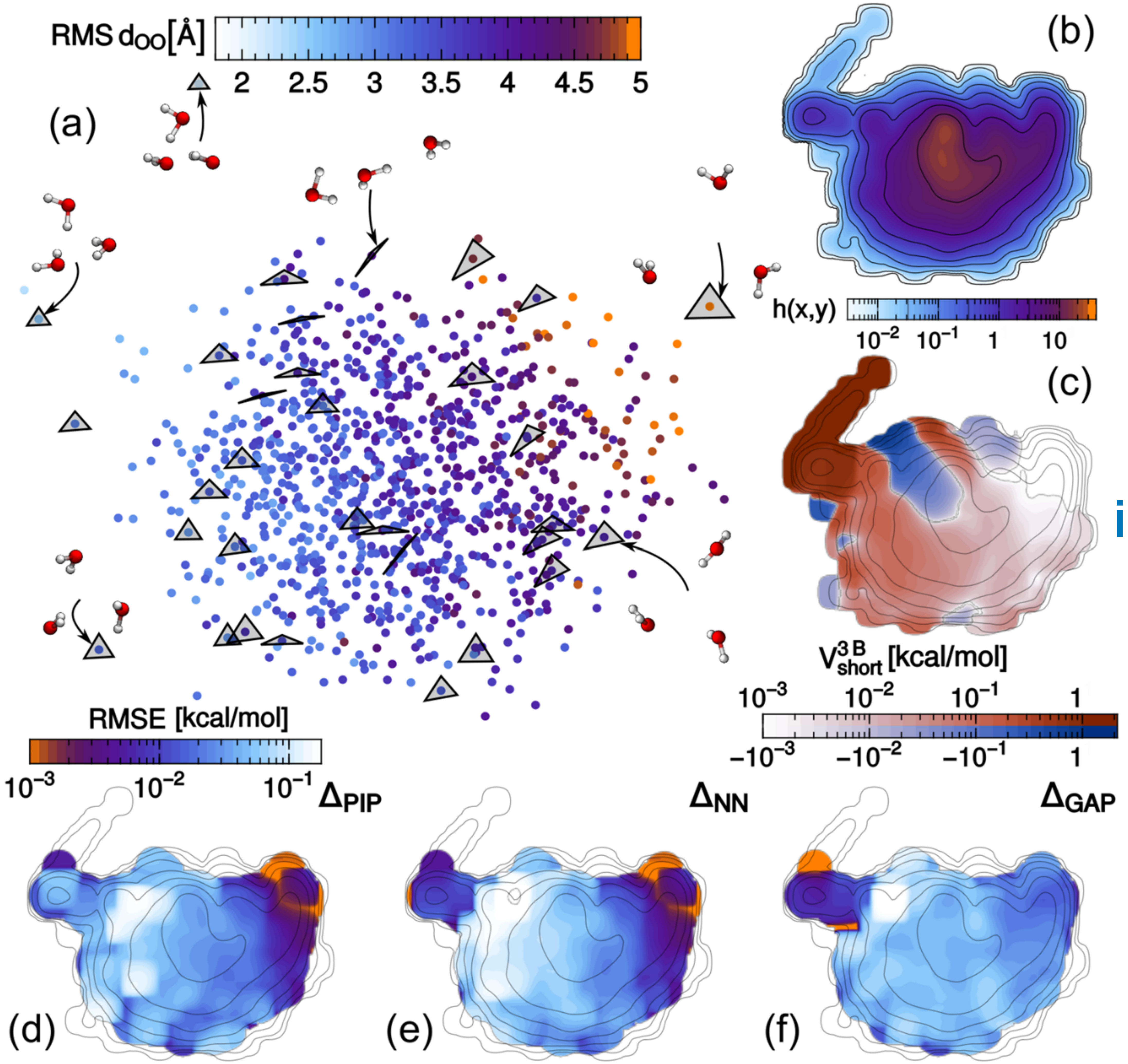
RMSE for PIP, NN, GAP 2B energies

Nguyen, Székely, Imbalzano, Behler, Csányi, Ceriotti & FP, J. Chem. Phys. 148, 241725 (2018)

Representability of Low-Order Many-Body Interactions

3B interaction energies

sketch map



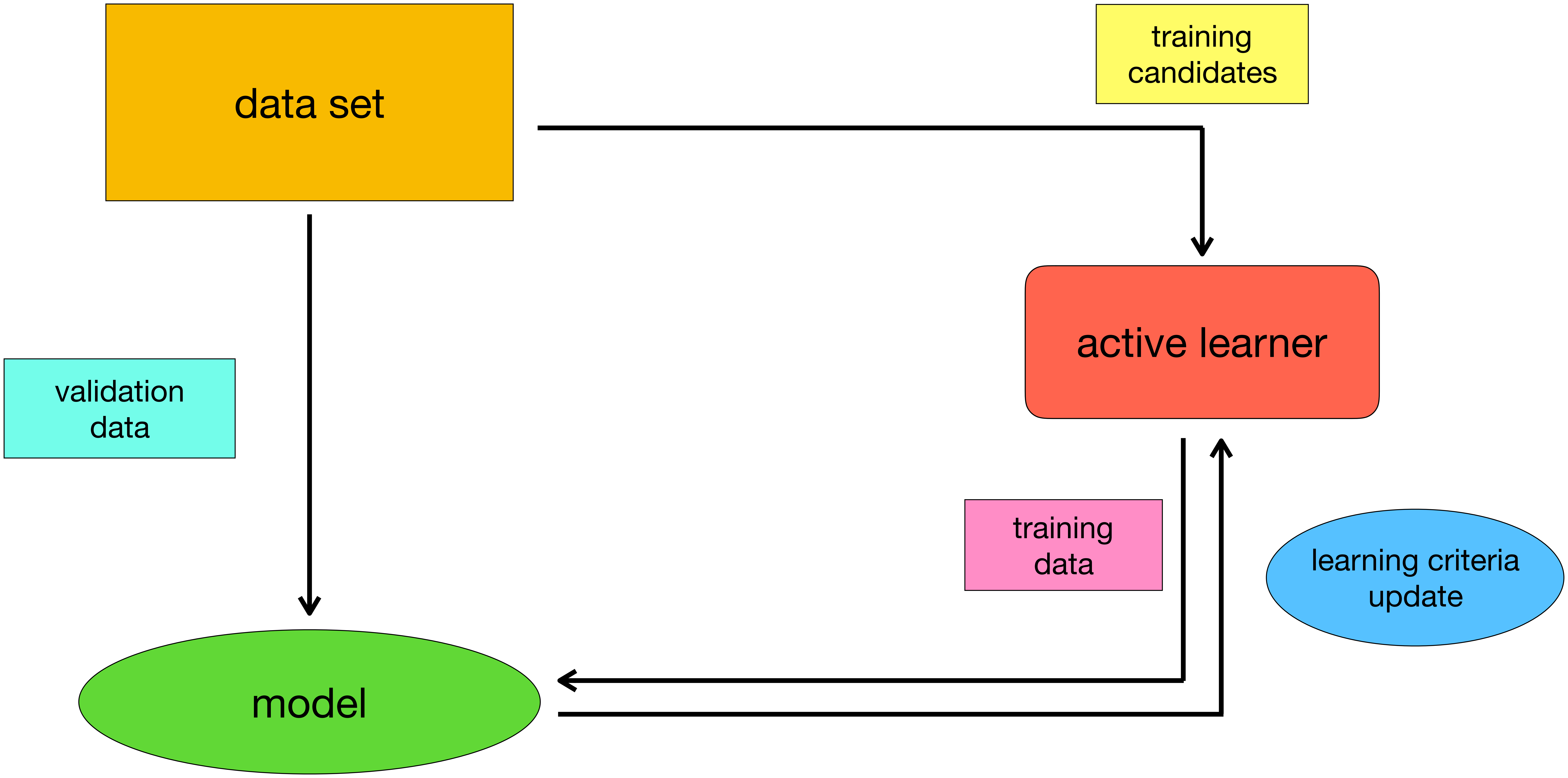
training point density

distribution of 2B energies in the training set

RMSE for PIP, NN, GAP 2B energies

Nguyen, Székely, Imbalzano, Behler, Csányi, Ceriotti & FP, J. Chem. Phys. 148, 241725 (2018)

Active Learning of Low-Order Many-Body Interactions



data set

training candidates

active learner

validation data

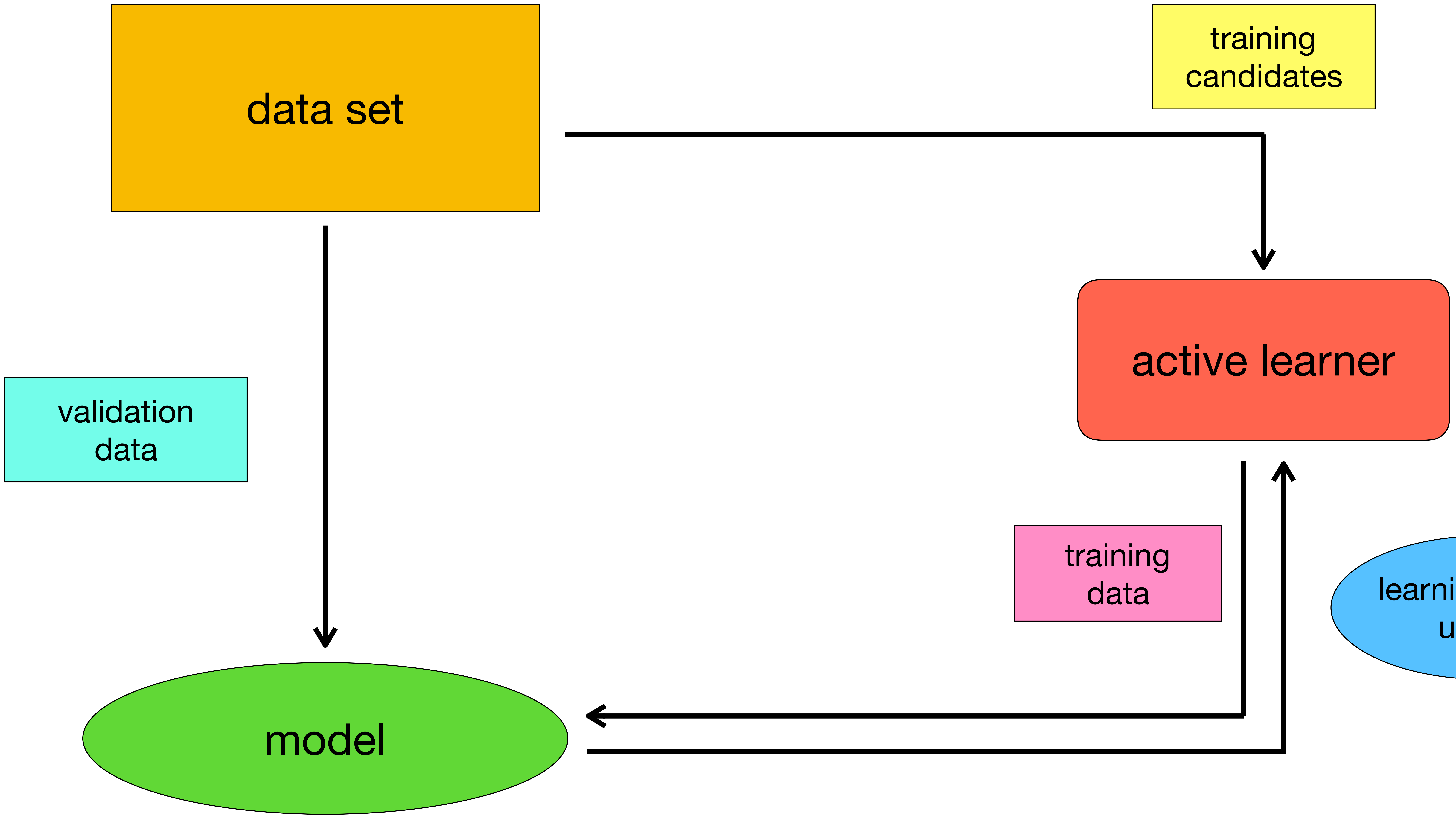
training data

learning criteria update

model

Zhai, Caruso, Gao & FP, in progress

Active Learning of Low-Order Many-Body Interactions



data set

training candidates

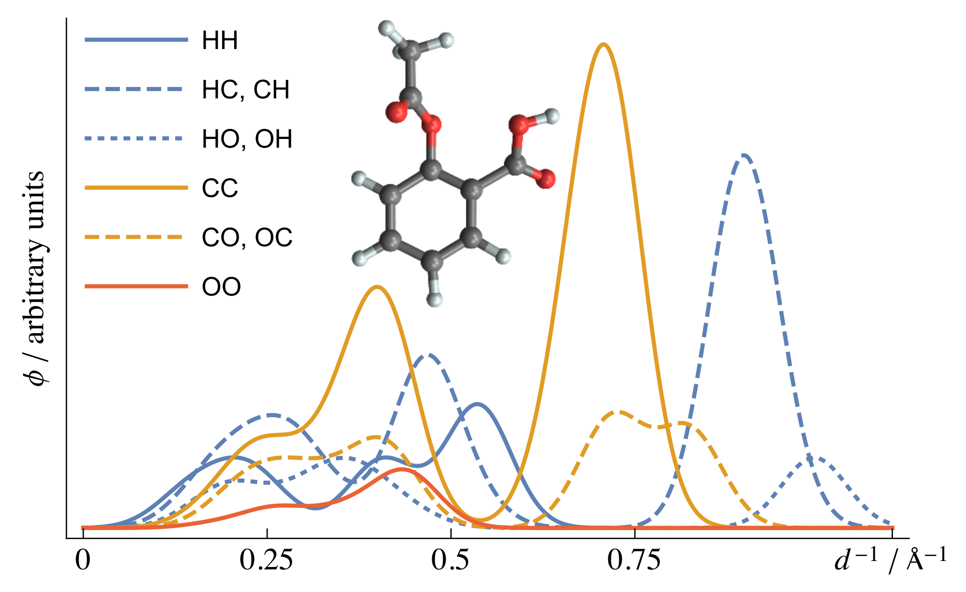
active learner

validation data

training data

learning criteria update

model

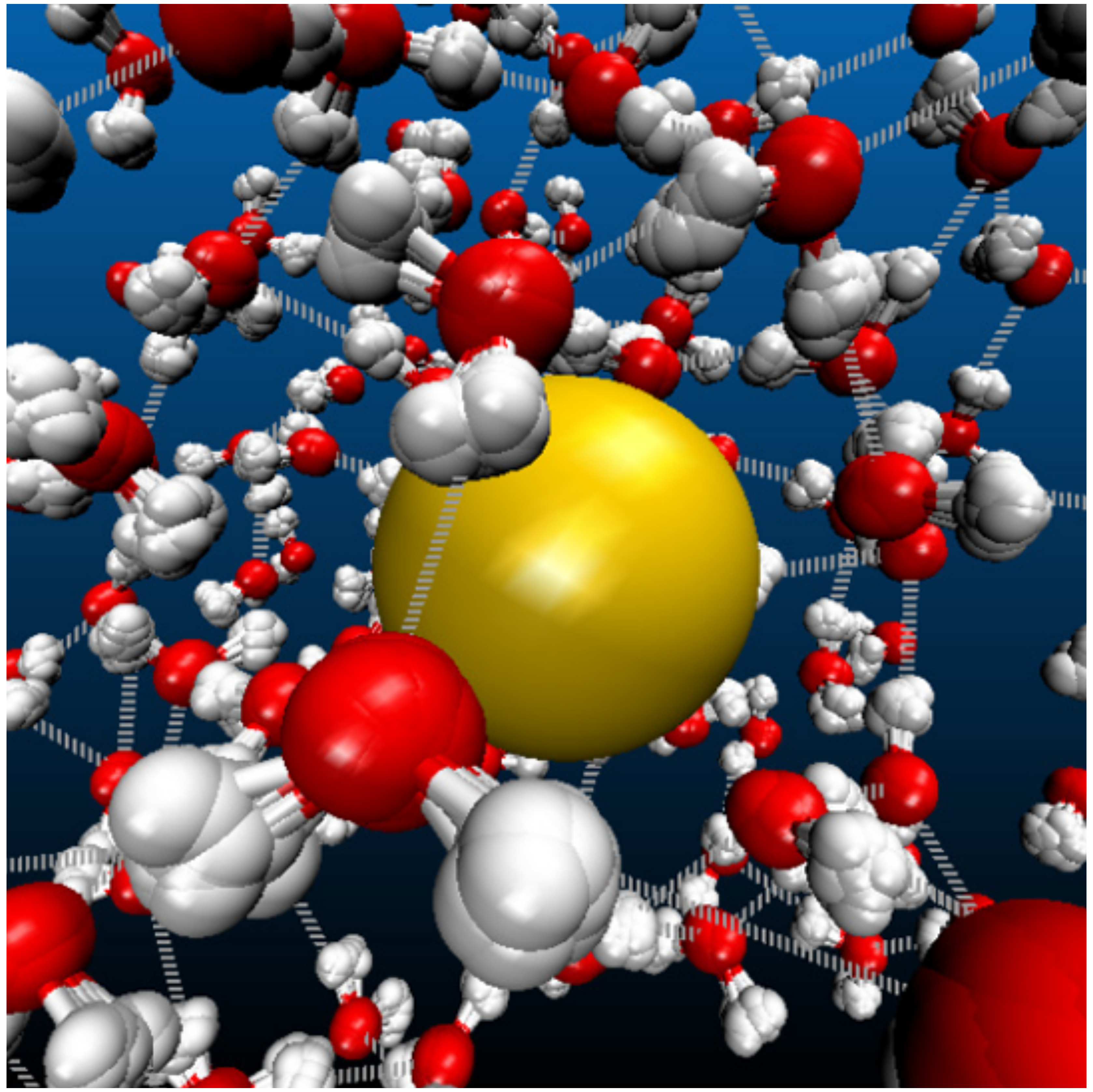


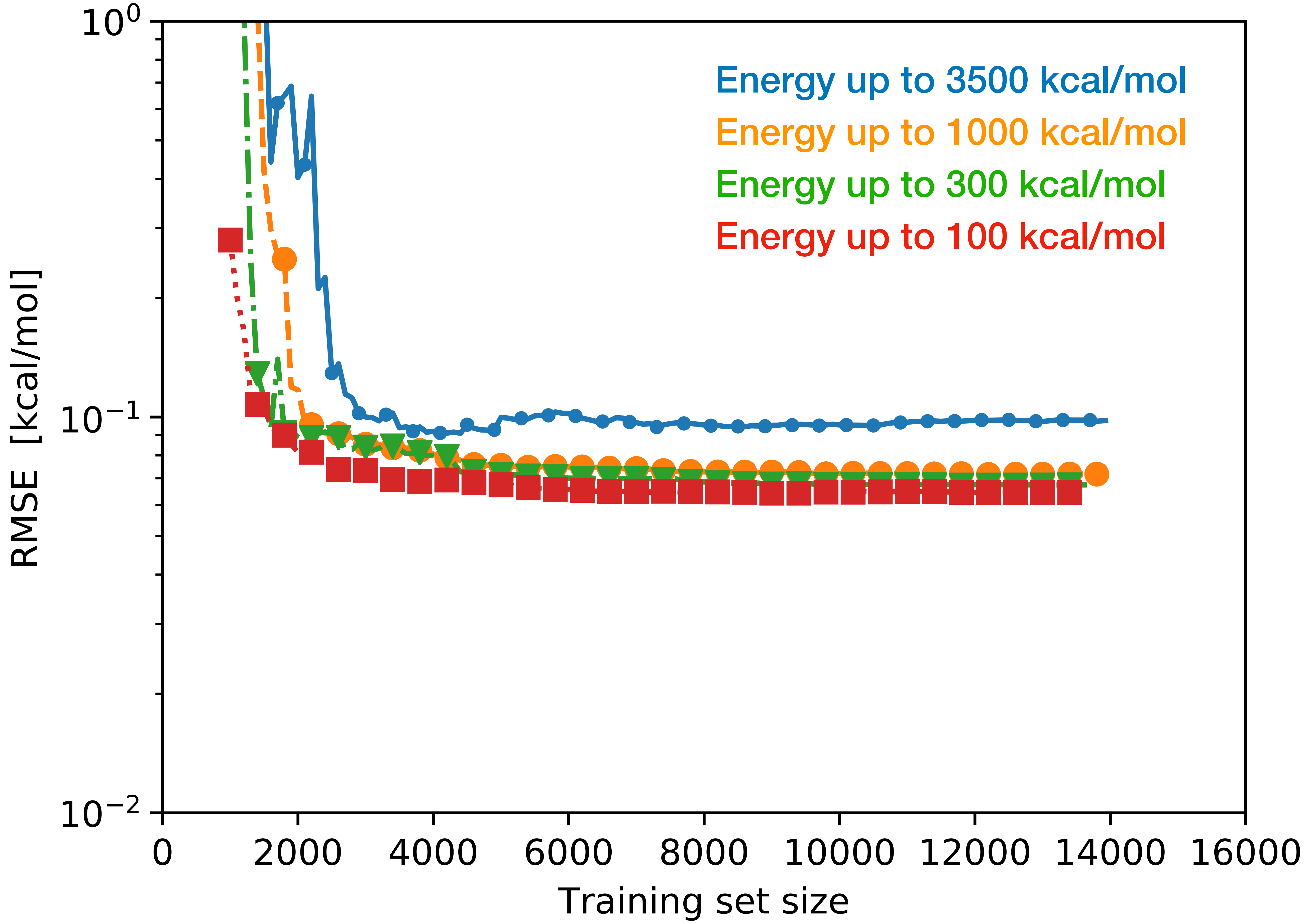
active learning:
many-body tensor representation
Huo & Rupp,
arxiv:1704.06439

Zhai, Caruso, Gao & FP, in progress

Generating "Optimal" Many-Body Training Sets

Actively learning water-Cs^+ many-body interactions

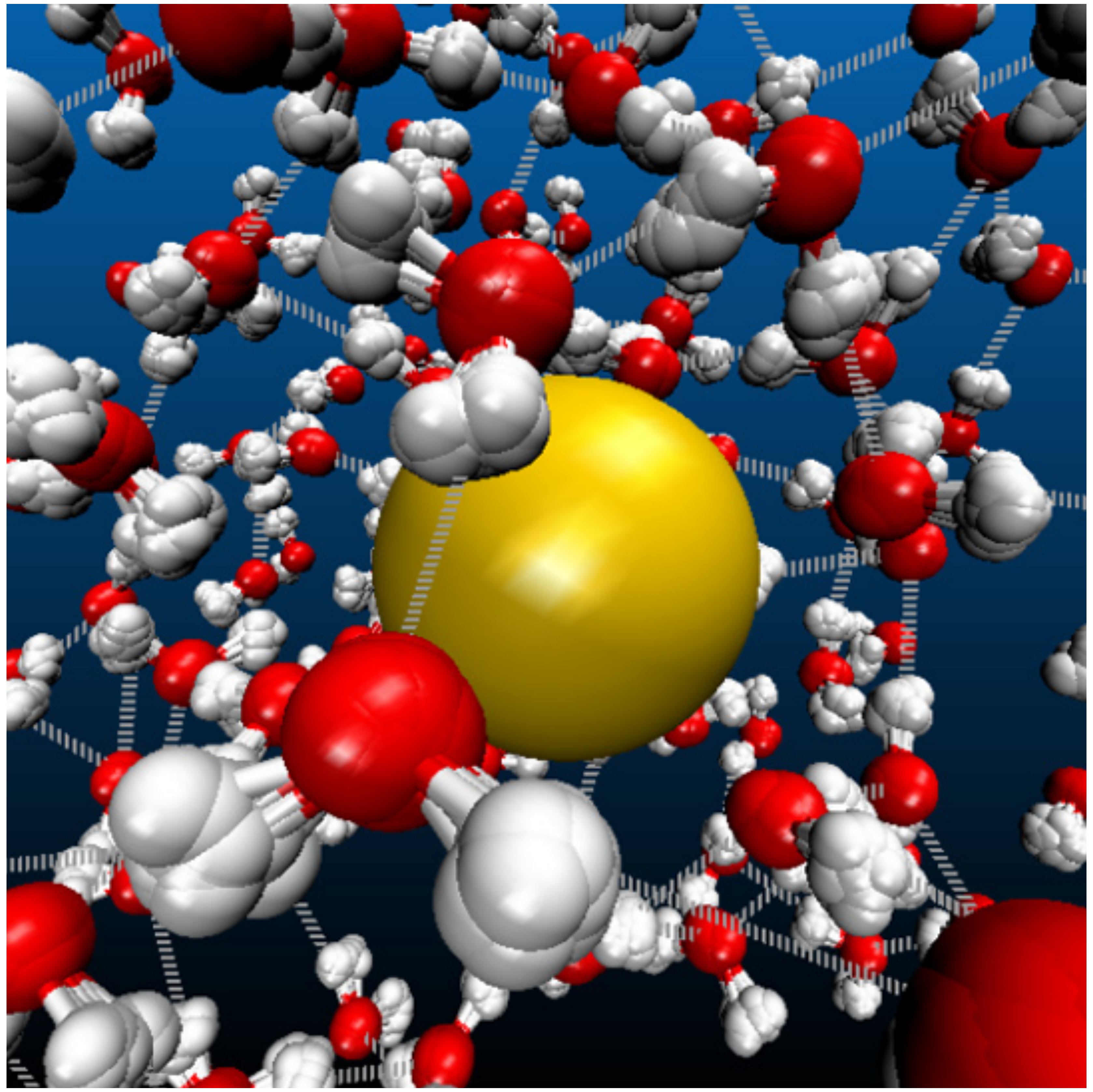


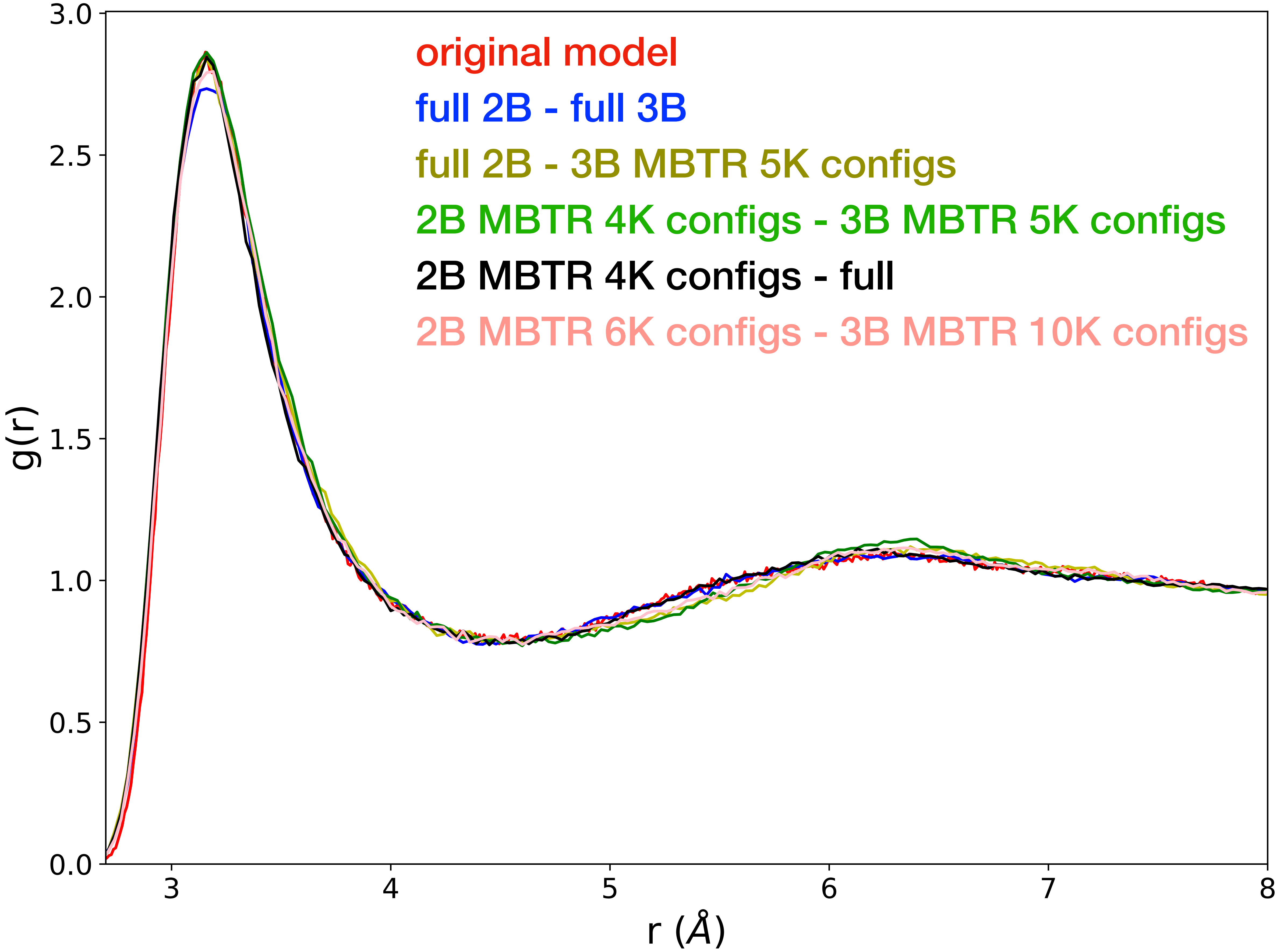


Zhai, Caruso, Gao & FP, in progress

Generating "Optimal" Many-Body Training Sets

Actively learning water-Cs^+ many-body interactions



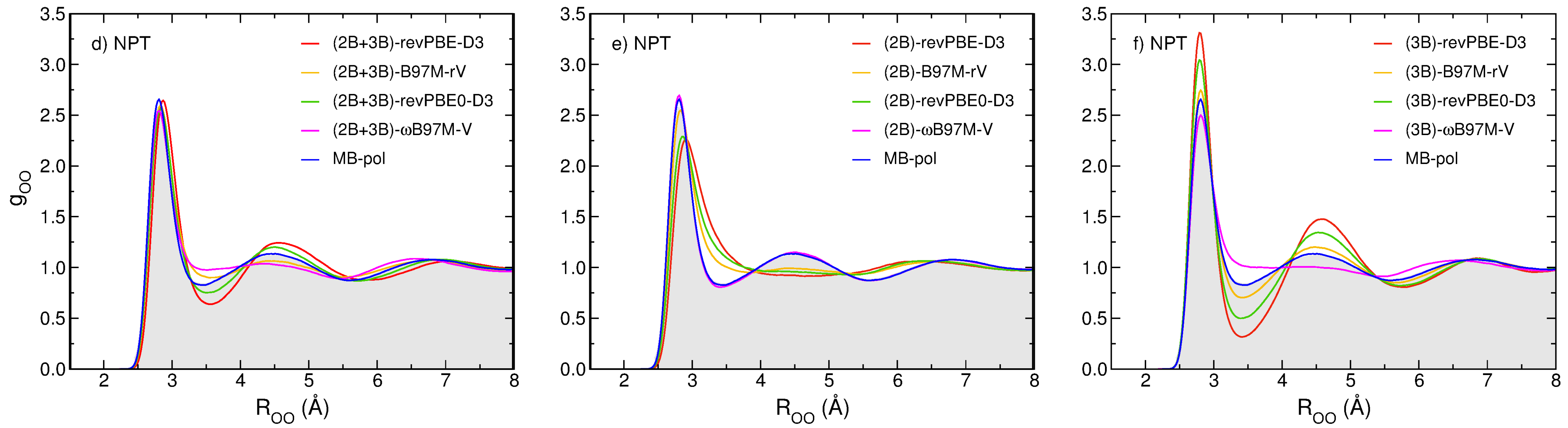


Zhai, Caruso, Gao & FP, in progress

Learning Physics from Many-Body Models

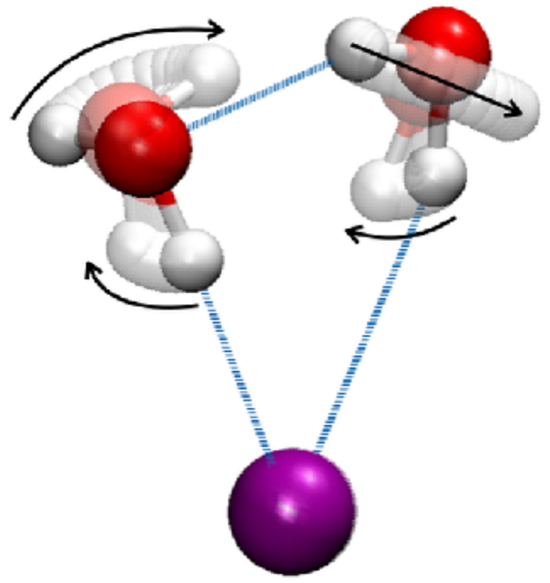
What makes water water?

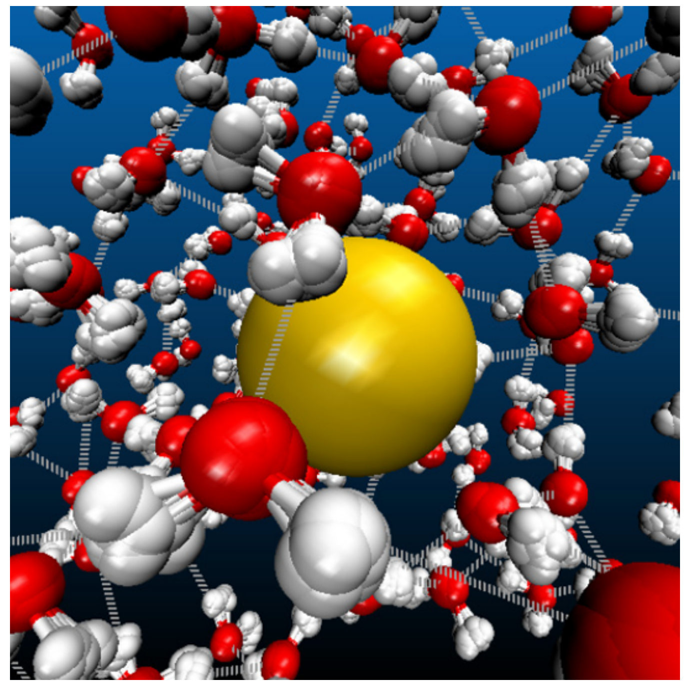
- Low-order many-body interactions determine the local structure of water
- Delocalization errors make current DFT models unreliable for water

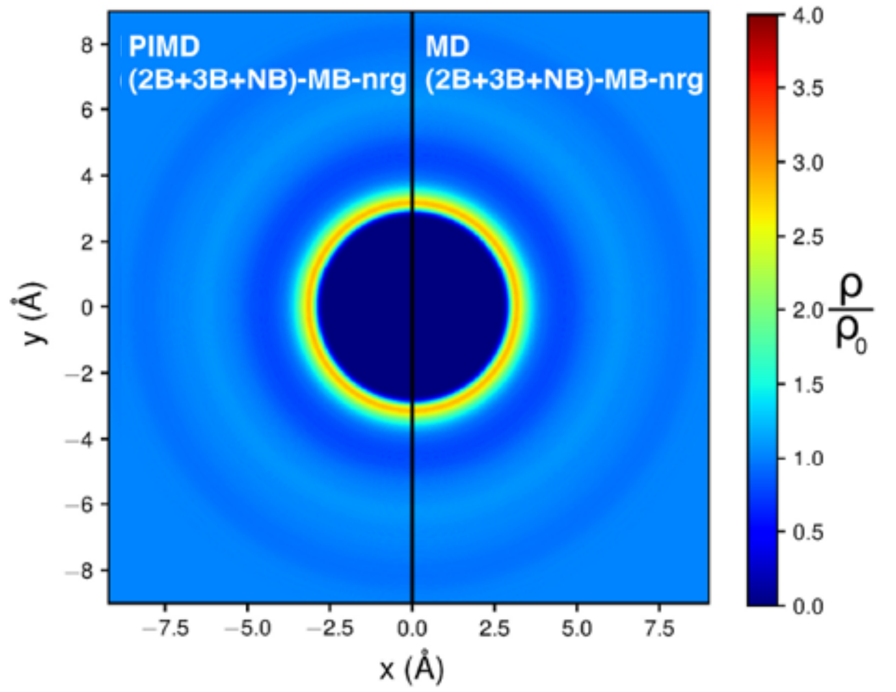


How good are simulations of hydration?

- Many-body interactions at both short and long ranges determine the hydration structure of monovalent ions
- Common force fields are unable to correctly predict the hydration structure of monovalent ions

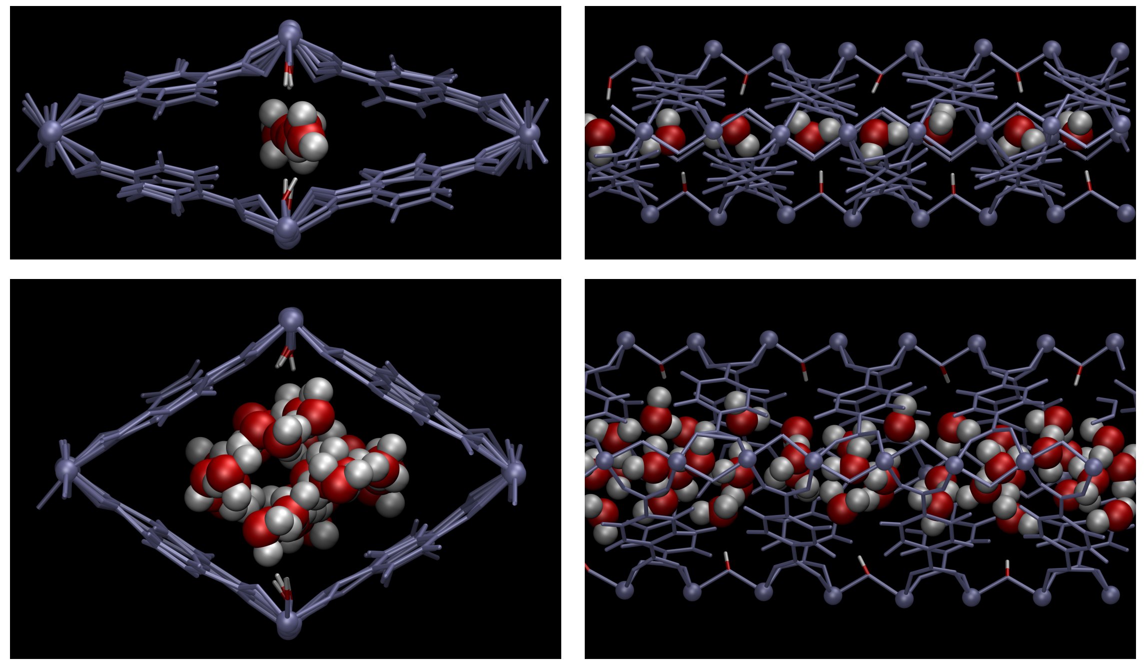


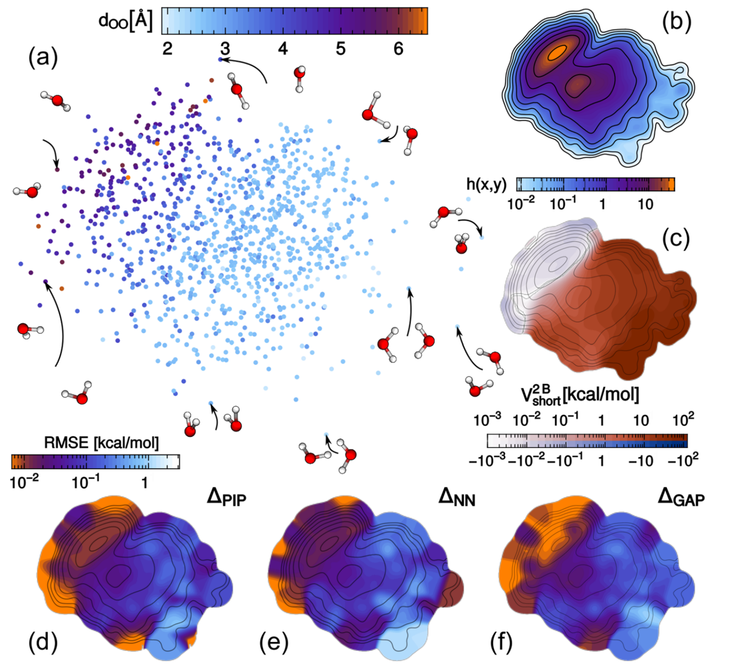




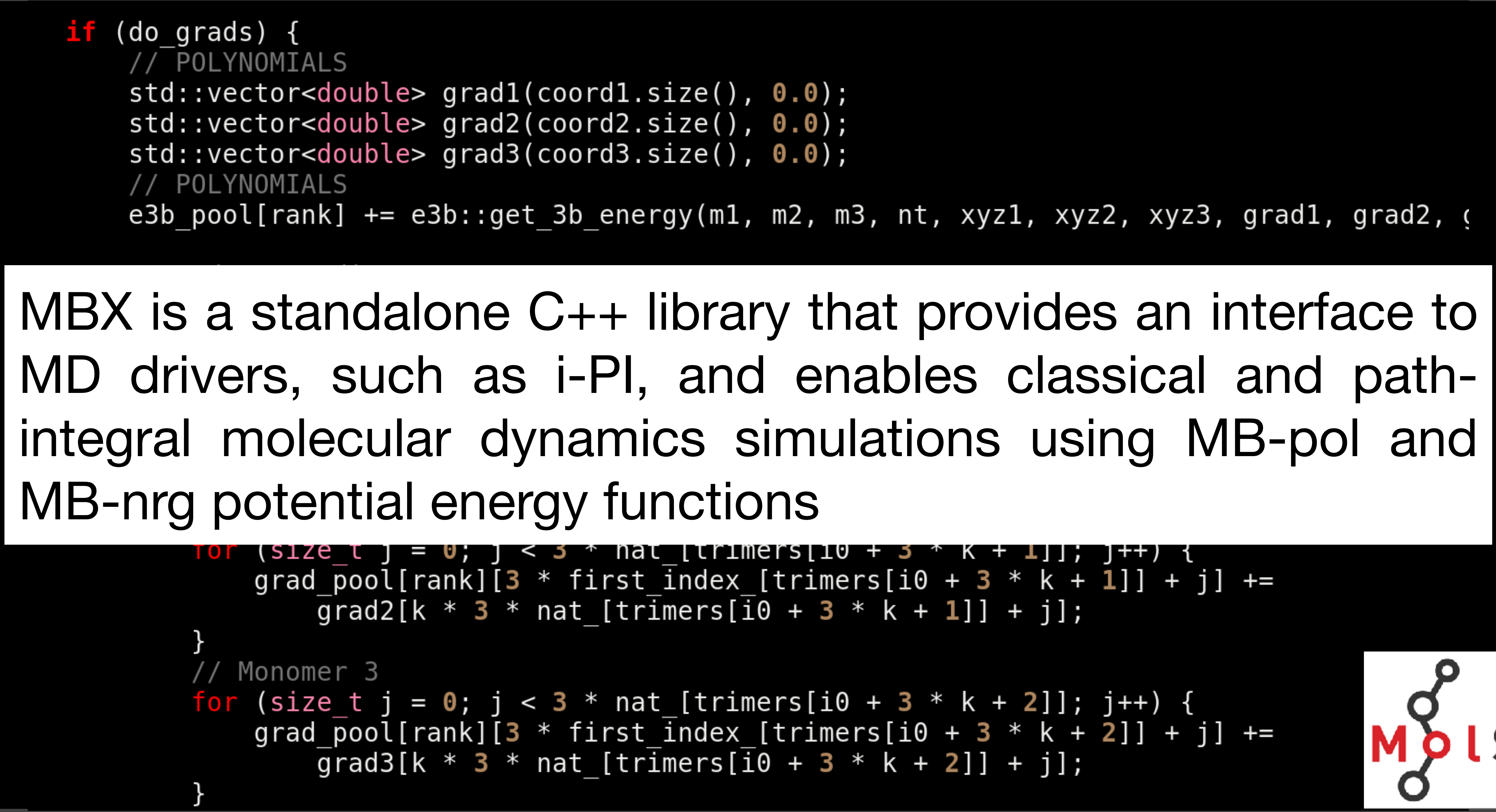
Data-driven many-body models: Quo vadis?

- Materials discovery for: anti-icing, water harvesting and desalination, separation
- Reference for more accurate DFT models
- Improved approaches for many-body training sets





MBX: C++ Library for Data-Driven Many-Body Models



```
if (do_grads) {
    // POLYNOMIALS
    std::vector<double> grad1(coord1.size(), 0.0);
    std::vector<double> grad2(coord2.size(), 0.0);
    std::vector<double> grad3(coord3.size(), 0.0);
    // POLYNOMIALS
    e3b_pool[rank] += e3b::get_3b_energy(m1, m2, m3, nt, xyz1, xyz2, xyz3, grad1, grad2,
```

```
        for (size_t j = 0; j < 3 * nat_[trimers[i0 + 3 * k + 1]]; j++) {
            grad_pool[rank][3 * first_index_[trimers[i0 + 3 * k + 1]] + j] +=
                grad2[k * 3 * nat_[trimers[i0 + 3 * k + 1]] + j];
        }
        // Monomer 3
        for (size_t j = 0; j < 3 * nat_[trimers[i0 + 3 * k + 2]]; j++) {
            grad_pool[rank][3 * first_index_[trimers[i0 + 3 * k + 2]] + j] +=
                grad3[k * 3 * nat_[trimers[i0 + 3 * k + 2]] + j];
        }
```

MBX is a standalone C++ library that provides an interface to MD drivers, such as i-PI, and enables classical and path-integral molecular dynamics simulations using MB-pol and MB-nrg potential energy functions

Riera, Smith, Simmonett & FP, http://paesanigroup.ucsd.edu/software/mbx.html

On the Quality of the Reference Data

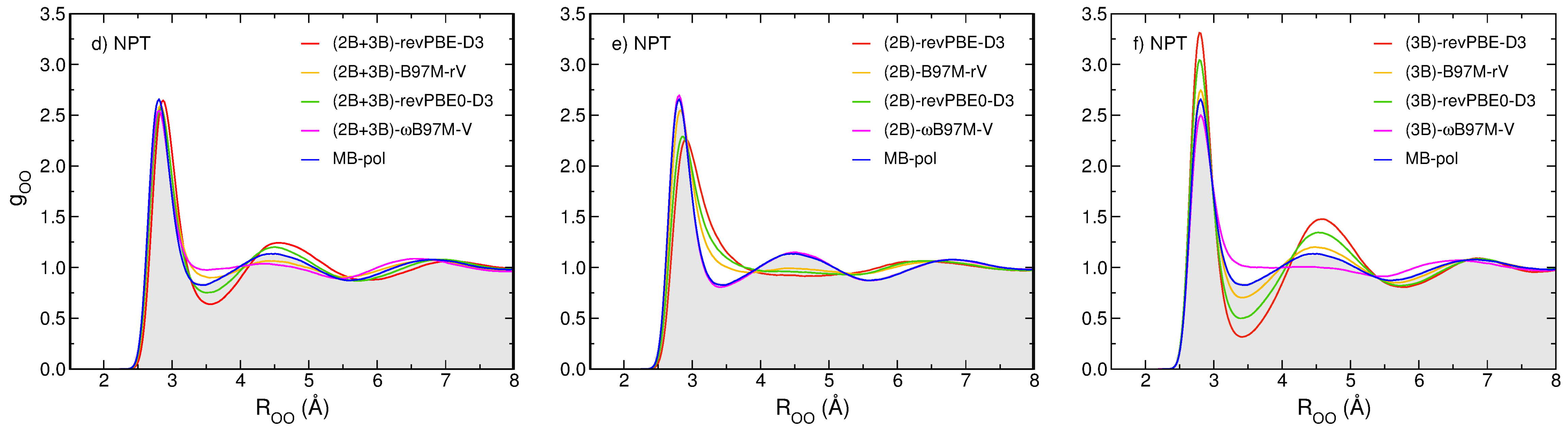
On the Quality of the Reference Data

Greetings from Las Vegas!

Learning Physics from Many-Body Models

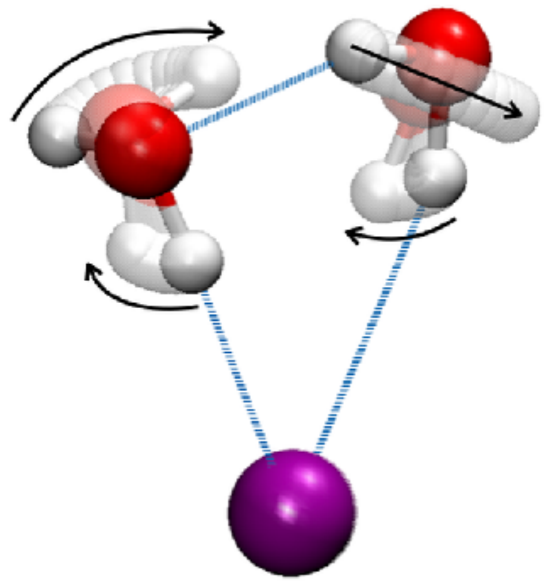
What makes water water?

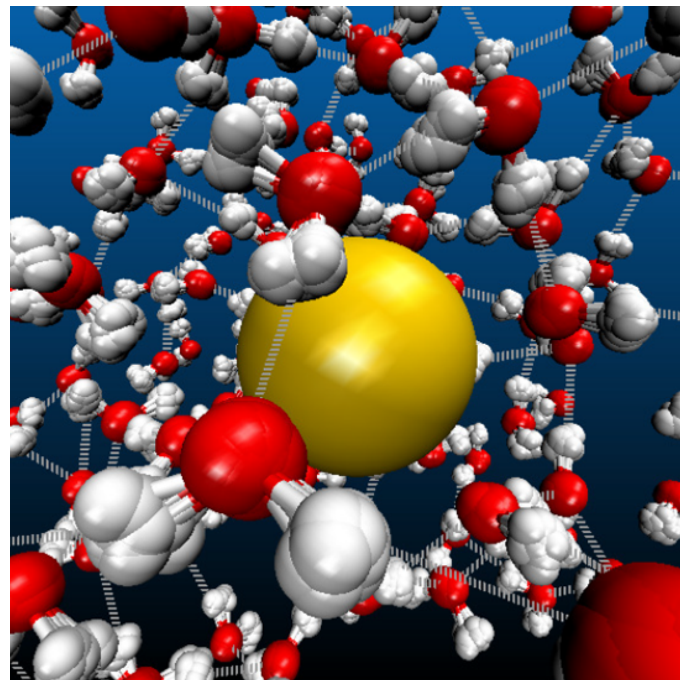
Acc. Chem. Res. 49, 1844 (2016)
J. Am. Chem. Soc. 138, 3912 (2016)
J. Am. Chem. Soc. 139, 7082 (2017)
Nat. Commun. 9, 247 (2018)
Phys. Rev. Lett. 121, 137401 (2018)
Chem. Sci. 10, 8211 (2019)

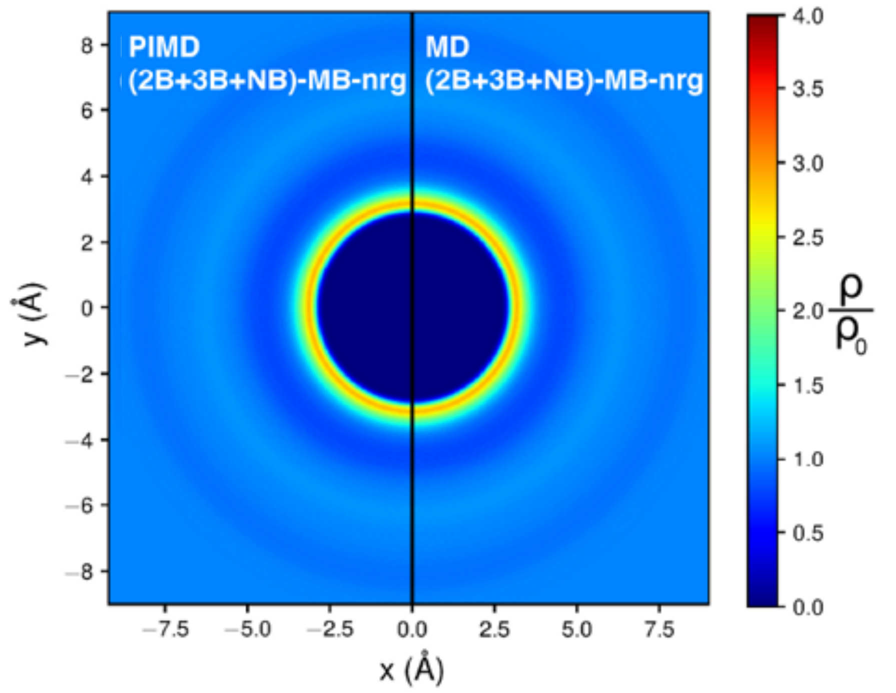


How good are simulations of hydration?

J. Chem. Theory Comput. 12, 2698 (2016)
J. Chem. Phys. 147, 161715 (2017)
J. Chem. Theory Comput. 15, 2983 (2019)
J. Phys. Chem. Lett. 10, 406 (2019)
J. Phys. Chem. Lett. 10, 2823 (2019)
Nat. Chem. 11, 367 (2019)







Data-driven many-body models: Quo vadis?

Proc. Natl. Acad. Sci. U.S.A. 115, 8266 (2018)
Nat. Commun. 10, 4771 (2019)
Proc. Natl. Acad. Sci. U.S.A., in press. https://doi.org/10.26434/chemrxiv.8856053.v1
J. Chem. Phys. 144, 224101 (2016)
J. Chem. Phys. 148, 241725 (2018)