NEW FUNCTIONAL FORMS AND PARAMETERIZATION METHODS FOR AB INITIO, INTERMOLECULAR FORCE FIELD DEVELOPMENT

by

Mary J. Van Vleet

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN–MADISON

2017

Date of final oral examination: 08/15/17

The dissertation is approved by the following members of the Final Oral Committee:
J.R. Schmidt, Associate Professor, Chemistry
Clark R. Landis, Professor, Chemistry
Qiang Cui, Professor, Chemistry
Arun Yethiraj, Professor, Chemistry
Reid Van Lehn, Assistant Professor, Chemical and Biological Engineering

© Copyright by Mary J. Van Vleet 2017
All Rights Reserved

Soli Deo gloria.

ACKNOWLEDGMENTS

It is customary for authors of academic books to include in their prefaces statements such as this: "I am indebted to ... for their invaluable help; however, any errors which remain are my sole responsibility." Occasionally an author will go further. Rather than say that if there are any mistakes then he is responsible for them, he will say that there will inevitably be some mistakes and he is responsible for them....

Although the shouldering of all responsibility is usually a social ritual, the admission that errors exist is not — it is often a sincere avowal of belief. But this appears to present a living and everyday example of a situation which philosophers have commonly dismissed as absurd; that it is sometimes rational to hold logically incompatible beliefs.

— David C. Makinson (1965)

Above is the famous "preface paradox," which illustrates how to use the `wbepi` environment for epigraphs at the beginning of chapters. You probably also want to thank the Academy.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

NEW FUNCTIONAL FORMS AND PARAMETERIZATION METHODS FOR AB INITIO, INTERMOLECULAR FORCE FIELD DEVELOPMENT

Mary J. Van Vleet

Under the supervision of Professor J.R. Schmidt
At the University of Wisconsin-Madison

FIXME: basically a placeholder; do not believe

I did some research, read a bunch of papers, published a couple myself, (pick one):

1. ran some experiments and made some graphs,
2. proved some theorems

and now I have a job. I've assembled this document in the last couple of months so you will let me leave. Thanks!

J.R. Schmidt

ABSTRACT

FIXME: basically a placeholder; do not believe

I did some research, read a bunch of papers, published a couple myself, (pick one):

1. ran some experiments and made some graphs,
2. proved some theorems

and now I have a job. I've assembled this document in the last couple of months so you will let me leave. Thanks!

PUBLISHED WORK AND WORK IN PREPARATION

---

[1] Van Vleet, M. J.; Misquitta, A. J.; Stone, A. J.; Schmidt, J. R. *J. Chem. Theory Comput.* **2016**, *12*, 3851–3870.

Part I

Introduction

1.1 The Importance of Molecular Simulation

This ref[2] is super cool!

What is molecular simulation? What types of problems can it solve? How does molecular simulation work? (Be sure to include solving Newton's EQs of motion and relevant details on the partition function and interaction energies!)

2.1 Molecular Mechanics and the Theory of Intermolecular Forces

What is a force field? What are the important components of a force field, and how do we model them?

2.1.1 The Many-Body Expansion

How do we break apart a force field into manageable pieces? Why does it make sense to break a force field into 2- and many-body components?

2.1.2 Energy Decomposition Schemes

Intramolecular Interactions

Brief commentary on the non-intermolecular portions of a force field

Electrostatics

Conceptual description of electrostatics: long-range multipoles and charge penetration

Exchange

Quantum-mechanically-based Pauli Exclusion. Theoretical grounds for exponential behavior

Induction

Charge transfer. Polarization. Polarization Damping.

Dispersion

Theoretical Formulation. Damping.

2.2 Ab-Initio Force Field Development

2.2.1 Electronic Structure Benchmarks

SAPT

General SAPT methodology. DFT-SAPT.

Coupled-Cluster Methods

CCSD(T). CCSD(T)-f12.

2.3 ISA-based methods for force field development

What is ISA? How can ISA be used to generate parameters for intermolecular force field development? What progress has been made from this approach?

Part II

Published Work

3 ISOTROPIC AB INITIO FORCE FIELDS

4 ANISOTROPIC AB INITIO FORCE FIELDS

Part III

Unpublished Work

5 AB INITIO FORCE FIELDS USING LMOEDA

5.1 Preface

The preceding sections have been devoted to a development of various methodologies for ab initio intermolecular force field development, all generally assuming that Symmetry-Adapted Perturbation Theory (SAPT) can be used as a benchmark electronic structure theory. Critically, and especially given the developments discussed in Chapter 4, we can now usually expect our model force field energies to be within ~1 kJ/mol of the SAPT reference values! In spite of this success, this high precision between the model and SAPT energies can only lead to experimentally-accurate molecular simulation provided that the SAPT energies themselves are accurate, either with respect to the exact underlying potential energy surface (PES) or (in practice) with respect to gold-standard CCSD(T) calculations. Indeed, for systems where SAPT and CCSD(T) disagree by several kJ/mol, there is little point in developing SAPT-based force fields with sub- kJ/mol accuracy! This limitation raises to two fundamentally important questions. First, for what types of systems might we expect SAPT to be inaccurate? Second, for the systems where SAPT and the exact PES are in disagreement, how must we modify our typical methodology for ab-initio force field development?

The purpose of this chapter is to partially address these questions, all within the specific context of force field development for Metal-Organic Frameworks (MOFs). Note that the results presented here were gathered from 2013–2015, so some important advances (namely those presented in Chapters 3 and 4) haven't been incorporated into the force fields presented here. This is probably to the determinent of the accuracy and transferability that might be possible with the LMO-EDA-based methodology, and (should this project be picked up in the future) it may be necessary to refit these force fields to the functional forms and monomer-based parameters discussed in Chapter 4.

5.2 Introduction

Metal-Organic Frameworks (MOFs) are an increasingly important class of compounds, fundamentally defined as porous materials comprised of inorganic nodes connected by organic linkers. Within this general motif, more than 20,000 compounds have been reported and studied,[3] and this vast diversity of MOF materials shows great promise for chemical customization and optimization. Within the past two decades, a huge body of research has been devoted to the design and study of MOFs, and current applications range from gas separation and storage to catalysis and biomedical imaging.[3]

Somewhat recently, it has been discovered that so-called Coordinatively-Unsaturated (CUS) MOFs can be created by activation of solvent-coordinated inorganic nodes to yield exposed (or 'open') metal sites.[4–6] These CUS-MOFs have been shown to exhibit exhibit excellent uptakes and selectivities in a number of gas separation and storage problems,[4,5,7] making this family of compounds an excellent target for future investigation and materials design. Owing to the vast scope of hypothetical CUS-MOF materials, however, and the number of chemically-distinct targets for gas separation/storage, it is unlikely that experiment alone can be used to screen for new and promising CUS-MOF materials.[8] Rather, a combination of experiment and computational modeling will be required to find (or possibly even rationally design) optimal CUS-MOFs.[7–9]

Despite the utility of computational studies, it remains challenging to develop molecular models for CUS-MOFs.[6] Because the strong binding between metal and adsorbate leads to chemical environments substantially different from typical coordinatively-saturated MOFs, many standard force fields (such as UFF and DREIDING) which yield good predictions for these MOFs frequently (and substantially!) underpredict adsorption in CUS-MOFs, especially at low pressures.[8–10] While CUS-MOFs can sometimes be studied using quntum mechanical means,[9,11] clearly new and improved force fields will be required to perform in-depth simulations and large-scale screenings of these materials.

The goal of the present chapter is to present a general methodology for devel-

oping accurate and transferable force fields for CUS-MOFs. The current study is limited to a discussion of the MOF-74 series (a prototypical and well-studied CUS-MOF), however it is expected that the methods presented herein might also be applicable to other systems. After outlining this methodology (Sections 5.3 and 5.4), we next show how our force fields can be applied to accurately predict CO_2 adsorption isotherms in Mg-MOF-74. At the present time, we do not have results for other compounds in the M-MOF-74 series (M = Co, Cr, Cu, Fe, Mn, Ni, Ti, V, and Zn), largely as a result of techical challenges in the force field parameterization itself. We discuss these technical challenges in some detail, and conclude with our perspective on the challenges and opportunities associated with developing transferable force fields for these and other CUS-MOF systems.

5.3 Background and Motivation

Prior work in our group has shown how, at least for coordinatively-saturated MOFs, accurate and transferable force fields can be generated for a wide variety of systems by fitting force field parameters on a component-by-component basis to reproduce an ab initio SAPT energy decomposition.[12,13] While full details for this force field development methodology can be found in refs. 13, 14, a short overview is as follows:

1. Generate a representative cluster model from which interaction parameters can be determined for each (new) pairwise interaction. An example cluster, used to parameterize Mg–CO_2 interactions in Mg-MOF-74, is shown in Fig. 5.1.

2. Compute, using DFT-SAPT (a variant of SAPT with monomer densities given by Density Functional Theory (DFT)), a series of representative dimer interaction energies for the model cluster. For the cluster model in Fig. 5.1, representative dimers were generated by varying the position of CO_2 with respect to the MOF cluster, and the corresponding DFT-SAPT total interaction energies are shown for a subset of representative points.

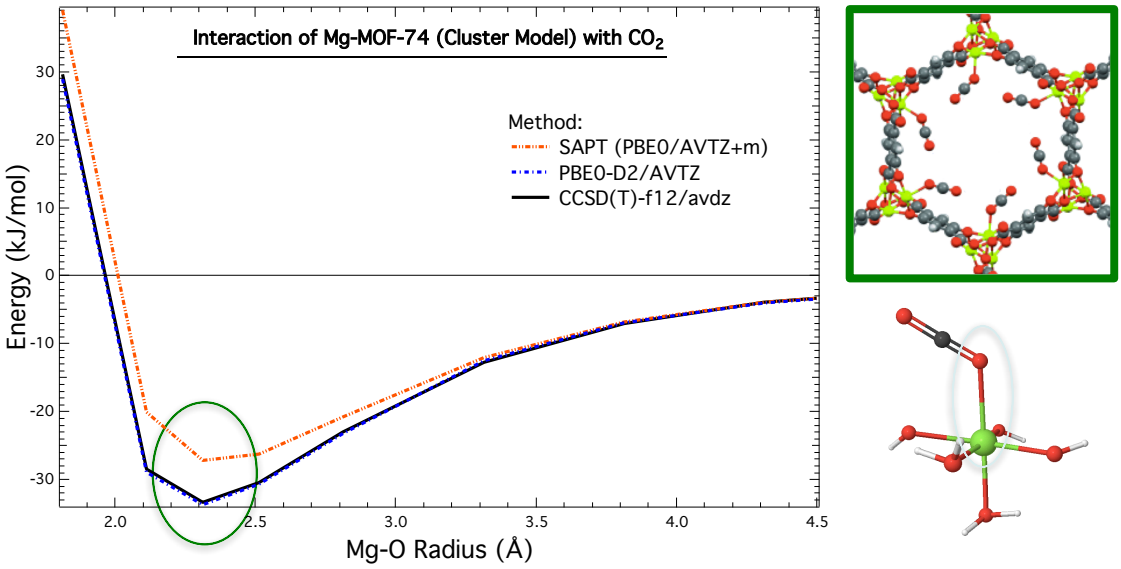


Figure 5.1: Model PES for interactions between CO_2 and Mg-MOF-74. (Left) Interaction energies between CO_2 and a cluster model of Mg-MOF-74 (shown bottom right), computed at a CCSD(T)-f12 (black), SAPT (orange), and/or PBE0-D2 (blue) level of theory. Discrepancies between SAPT and CCSD(T)-f12 in the minimum-energy region of the potential have been highlighted. (Top right) The structure of CO_2-bound Mg-MOF-74. (Bottom right) The structure of the cluster model used for Mg-MOF-74, where the circled atom pair indicates the relevant Mg-O radius from the x-axis in the leftmost figure.

3. To determine partial charges for the system, generate representative clusters (as described in Section 5.5) for each the organic ligand and the inorganic node, and perform a Distributed Multipole Analysis (DMA) analysis on each cluster to determine charges for the overall periodic system.

4. For each component of the DFT-SAPT interaction energy, parameterize the relevant functional forms (as detailed in ref. 13 and Section 5.4) to reproduce the DFT-SAPT component energy.

Once parameterized, these SAPT-based MOF force fields can be used for calculating individual adsorption isotherms or even for high-throughput screening.[14]

In the generation of force fields for CUS-MOFs, we expect that many of the advantages of the above workflow for coordinatively-saturated MOFs (such as the component-by-component based parameterization and method for partial charge determination) will also translate well to CUS-MOF materials. Nevertheless, there are two reasons why a SAPT-based methodology cannot be used to generate such force fields. First, and as shown in Fig. 5.1 for a representative Mg-MOF-74 cluster model, we have empirically found SAPT to be in error for CUS-MOF-like systems compared to benchmark CCSD(T)-f12 calculations. DFT-SAPT has been known to struggle with highly ionic systems (relative to CCSD(T) or DFT methods),[15,16] and so this error is perhaps not surprising. (Possible sources of the discrepancy between SAPT and CCSD(T)-f12 will be discussed in **??**.) Nevertheless, and in the absence of fortuitous error cancellation, predictions from an ab initio force field can only be as good as the level of theory that they are parameterized against. Consequently, because SAPT underbinds CO_2 by a full 6 kJ/mol compared to CCSD(T)-f12, we would not expect to see good predictions for the CO_2 adsorption isotherm with a SAPT-based methodology, and a new strategy will be required.

As a second barrier to using a SAPT-based methodology, many of the compounds in the M-MOF-74 series are open-shell. Though this poses no fundamental issue, in practice most implementations of SAPT (aside from the seldom-used SAPT 2012 package developed in Krzysztof Szalewicz's group at Delaware) do not allow for computations of open-shell systems, and indeed SAPT-based studies of open-shell compounds are very rare.[17] For these reasons, a new electronic structure benchmark is highly preferable.

5.4 New Methods for CUS-MOF force fields

Based on the results for Mg-MOF-74, it is clear that, at least for CUS-MOFs, a new methodology is required which simultaneously keeps the important advantages of the old workflow (especially the component-by-component based parameterization, which is essential for generating transferable force fields) while overcoming the limitations of SAPT itself. Put differently, for CUS-MOFs we should seek a

new electronic structure theory benchmark and associated energy decomposition analysis with the following qualities:

1. High accuracy with respect to to CCSD(T)-f12 benchmark energies
2. Physically-meaningful energy decomposition into (at least) electrostatics, exchange, induction, and dispersion
3. For systems where SAPT and CCSD(T)-f12 agree, a quantitative correspondence between the energy decompositions of SAPT and the new method

Assuming these three qualites are met, we expect to be able to generate force fields for CUS-MOFs that are both highly accurate and maximally-compatible with previous force fields developed for coordinatively-saturated MOF systems.

A substantial number of Energy Decomposition Analyses (EDAs) exist in the literature, and the interested reader is referred to ref. 16 for a review and comparison of various popular methods. Aside from SAPT, which is a perturbative method, most EDAs are 'variational', meaning that the various energy components are calculated in stages from a series of constrained relaxations of the monomer wavefunctions into the optimized dimer wavefunction. For this reason, all variational EDAs are guaranteed to have total energies that match the result from a supermolecular interaction energy calculation. Furthermore, these EDAs are often implemented for wavefunction and DFT methods, thus allowing for significant flexibility (compared to the SAPT EDA) in terms of finding an EDA whose total energy closely matches CCSD(T)-f12. Indeed, and as shown in Fig. 5.1, PBE0-D2 shows excellent agreement with CCSD(T)-f12 for a Mg-MOF-74 cluster model, and any DFT-compatible EDA will meet our first criteria from above.

Although all variational EDAs yield the same total interaction energy for a given level of theory, many EDAs can differ substantially in terms of how this total energy is decomposed into chemically-meaningful components. At the time this research was completed, only a handful of variational EDAs distinguished each electrostatics, exchange, induction, and dispersion. (Notably, the recent second-generation ALMO-EDA[18] now separates their 'frozen' energy term into electrostatic, exchange, and

dispersion components, and might thus might be worth re-investigation.) Of the popular EDA methods available as of 2014, we found that LMO-EDA,[19,20] GKS-EDA,[21] and PIEDA[22] decompose the total interaction in a manner philosophically similar to SAPT, and include each electrostatic, exchange, induction, and dispersion terms. These three methods thus meet our second criteria for an optimal energy decomposition scheme for CUS-MOFs, and complete formalisms and details for the methods can be found in refs. 19–22.

As for the last criterion, that of maximum correspondence between SAPT and a variational EDA, we have performed component-by-component analyses to compare SAPT to both LMO-EDA and GKS-EDA. PIEDA is known to overestimate the relative magnitude of the polarization energy, compared to SAPT, and thus was not considered in detail.[16] As for LMO-EDA and GKS-EDA (both of which are based on very similar theories, and tend to yield similar energy decompositions), we have in general found semi-quantitative to quantitative agreement with the SAPT energy decomposition, particularly for the electrostatic and exchange energies. Comparisons between LMO-EDA and SAPT are shown for the CO_2 dimer (Fig. 5.2) and for CO_2 interacting with a model Mg-MOF-74 compound (Fig. 5.3). GKS-EDA results are not shown, as the LMO-EDA and GKS-EDA results tend to be very similar, with the GKS-EDA results in slightly worse agreement with SAPT. For this reason, and because LMO-EDA does the best job of meeting our three criteria above, we choose in this work to use LMO-EDA as our new benchmark EDA for fitting CUS-MOF force fields.

In addition to describing the advantages of the LMO-EDA method, it is worthwhile to overview some of its relevant shortcomings, and limitations. As with most variational EDA methods,[16] and especially for DFT-based methods, it becomes difficult to precisely assign and separate out the true 'dispersion' energy for a system. This limitation is also true of LMO-EDA, where the dispersion energy is defined as the difference in correlation energy between the monomer and dimer wavefunctions. For density functionals employing Grimme's –D dispersion correction, this correction is also added to the LMO-EDA dispersion energy. For functionals that have a well-defined and theoretically-grounded distinction between the

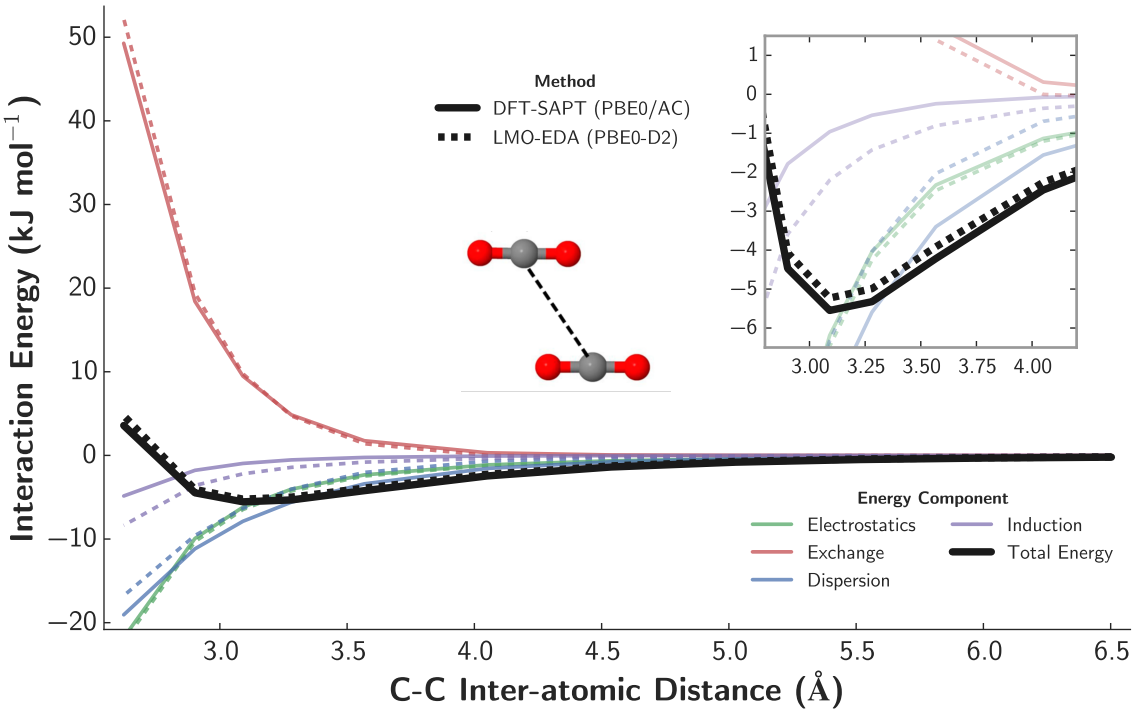


Figure 5.2:

exchange and correlation functionals, the LMO-EDA energies tend to agree well with SAPT, and we have found good agreement (for instance) between SAPT and LMO-EDA-PBE0-D2. With other functionals, such as with our tests using the M06 functional, there is no separation between the exchange and correlation functionals, and LMO-EDA gives unphysical values for both the exchange and dispersion energies in this case. (GKS-EDA attempts to rectify this issue by changing the LMO-EDA formalism for dispersion. While this leads to qualitative agreement between SAPT and GKS-EDA for a wider variety of functionals, the quantitative agreement for the PBE0-D2 functional is somewhat worsened for the systems studied herein, and we instead use LMO-EDA-PBE0-D2 for all results in this work.)

A second, and purely practical, limitation of LMO-EDA is its memory-intensive implementation in GAMESS. As will be discussed in detail later, calculations on a large (43 heavy atom) cluster model of Mg-MOF-74 were infeasible for us (using

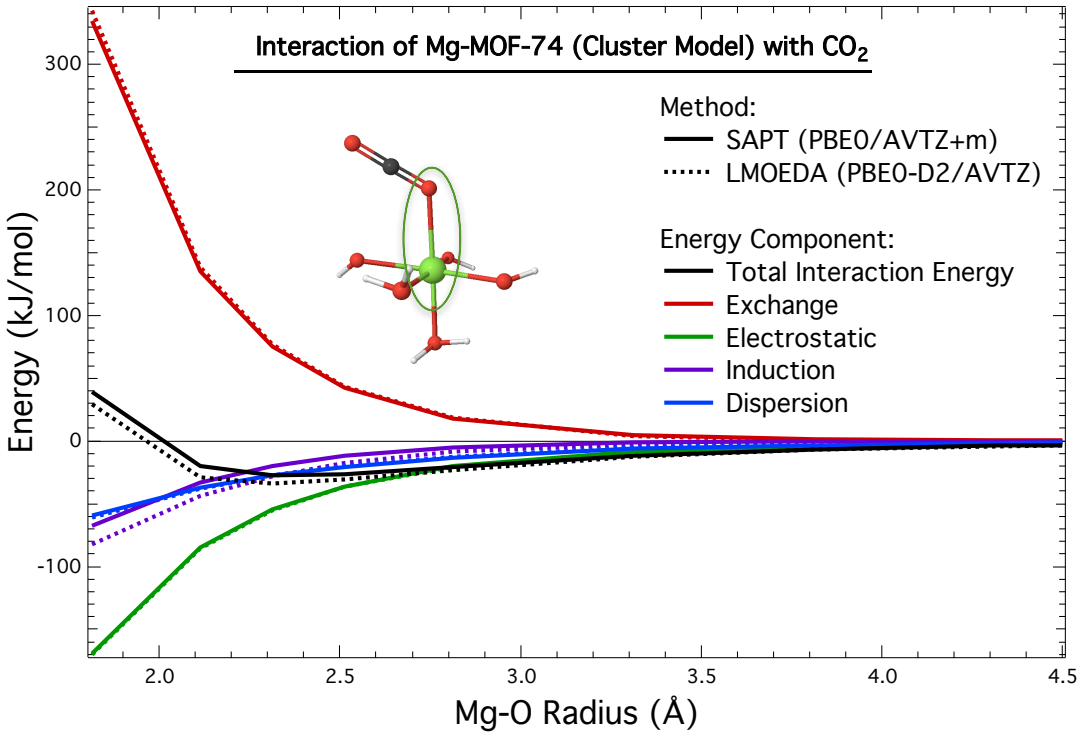


Figure 5.3:

the Phoenix cluster in 2014) in all but the smallest VDZ basis set, and calculations on an identical cluster model of Co-MOF-74 could not be completed at all. For this reason, the LMO-EDA method is practically restricted to studies of smaller systems and/or basis sets.

5.5 Computational Methods

Determining Partial Charges

Force Field Functional Forms

Force field fitting – single exponential form – double exponential form

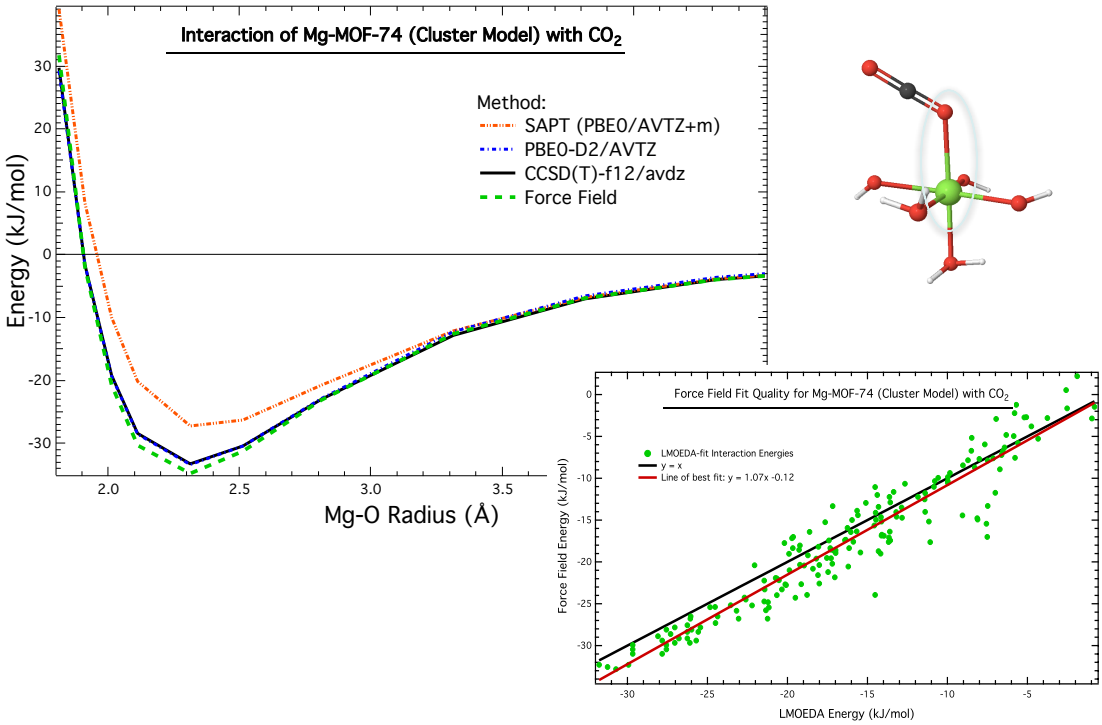


Figure 5.4:

5.6 Results

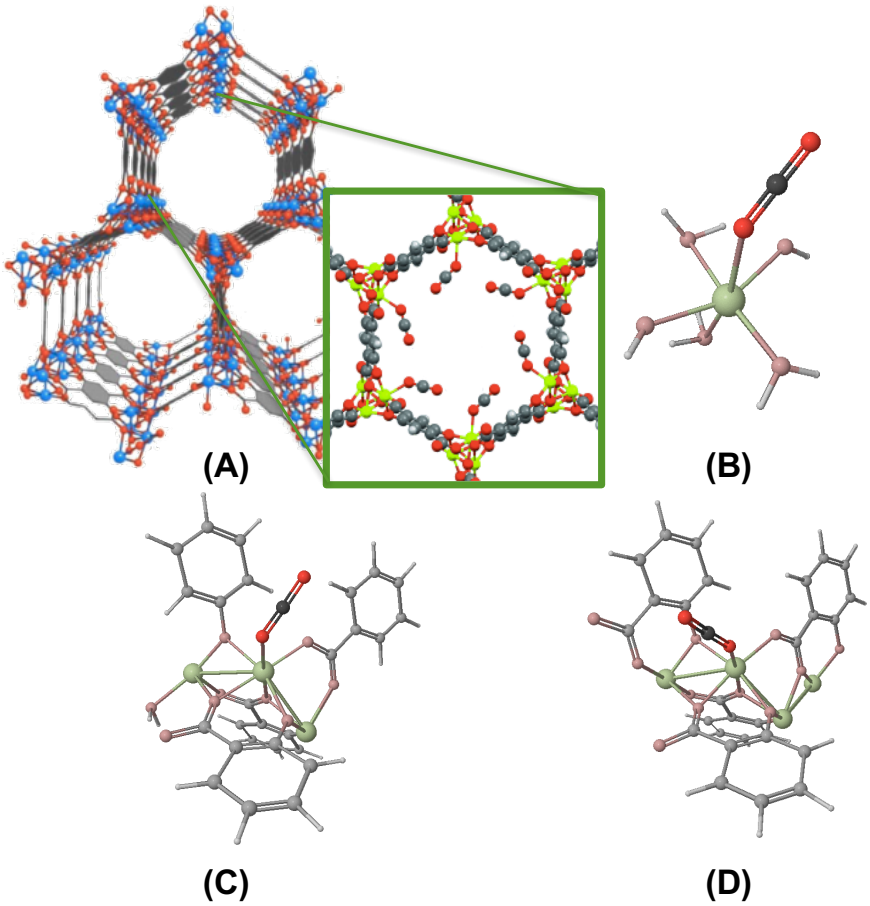
5.6.1 Initial Force Field and Cluster Model Analysis

Originally, we attempted to fit Mg parameters on the basis of a small, 6 heavy atom cluster ('Mg-MOF-74-small', see Fig. 5.4 for chemical structure), which we felt would be representative of the Mg environment in Mg-MOF-74. Using the functional forms discussed in Section 5.5, force field parameters were fit to reproduce LMO-EDA-PBE0-D2 energies for a variety of CO_2/Mg-MOF-74-small interactions, with results shown in Fig. 5.4. Though some interaction energies disagree by several kJ/mol between LMO-EDA-PBE0-D2 and the force field energies, overall the agreement is reasonable, and the force field correctly reproduces trends in the interaction energies without significant systematic error.

Based on the agreement between PBE0-D2 and the force field, as well as between PBE0-D2 and CCSD(T)-f12, we expected to obtain good CO_2 adsorption isotherm predictions for the Mg-MOF-74 system itself. By contrast, our computed isotherm substantially underpredicts the experimental adsorption at low pressures, where Mg–CO_2 interactions are known to dominate. This underprediction strongly suggests that we had originally underestimated the magnitude of the Mg–CO_2 binding, a result which we were then able to attribute to our choice of cluster model (vida infra).

Cluster models for the M-MOF-74 series have been investigated by several groups, and it has been found in general that computed binding energies are sensitive both to the size of the cluster model as well as the treatment of geometry relaxation effects.[24,25] Consequently, we calculated the CO_2 binding energies and geometries of both our original Mg-MOF-74-small cluster as well as for two larger clusters developed in refs. 6, 23. These latter two clusters, respectively denoted Mg-MOF-74-Yu and Mg-MOF-74-Dzubak, are the same size (each with 60 atoms), but have made different choices as to the number of included magnesium atoms (three vs. four) and number/location of capping atoms. To test the influence of model cluster on the CO_2 binding energy/geometry, we performed two sets of optimizations of the Mg-MOF-74-Yu and Mg-MOF-74-Dzubak clusters: one in which only the CO_2 position was optimized, and one in which the exposed MgO_5 pocket was additionally relaxed. Binding geometries were relatively insensitive to the geometry relaxation, though binding energies varied by 2-5 kJ/mol, in agreement with other studies that have tested geometry relaxation effects.[24] Results for the CO_2 + MgO_5 relaxation are shown in Fig. 5.5.

Of the three studied cluster models, both Mg-MOF-74-small and Mg-MOF-74-Yu correctly reproduce the Mg–O interatomic distance and Mg–O–C tilt angle. These geometrical parameters arise primarily from electrostatic interactions between CO_2 and the MgO_5 pocket,[11] suggesting that both of these models capture such important interaction features. By contrast, the Mg-MOF-74-Dzubak model predicts a substantially shorter binding distance and increased tilt angle, both in contrast to results from the periodic system. In part, these deficiencies can be attributed



Model	CO_2 Binding Energy (kJ/mol)	Mg–O Interatomic Distance (Å)	Mg–O–C Tilt Angle (°)
A[11]	**-41.5**	**2.31**	**129**
B	-23.3	2.31	122
C	-31.4	2.28	123
D	-41.7	2.20	149

Figure 5.5: Various cluster models for Mg-MOF-74 interacting with CO_2. (A) Full periodic Mg-MOF-74 structure with inset showing adsorbed CO_2 positions. (B) Mg-MOF-74-small cluster, containing 6 heavy atoms (not including CO_2). (C) Yu et al. cluster model for Mg-MOF-74, denoted in text as Mg-MOF-74-Yu. (D) Dzubak et al. cluster model for Mg-MOF-74, denoted in text as Mg-MOF-74-Dzubak. All cluster models as shown with optimized CO_2 positions, and bond lengths and angles for adsorbed CO_2 are given in the bottom table. Data for (A) was taken from Valenzano et al. using a B3LYP-D level of theory,[11] whereas data for (B-D) was computed in this work using PBE0-D2. Finally, note that the binding energy for (A) includes framework geometry relaxation effects, whereas (B-D) were computed using rigid cluster geometries and only optimizing the CO_2 position.

to spurious CO_2 interactions with the exposed carbonyl capping groups in the Mg-MOF-74-Dzubak model which are not present in the periodic system or the other two cluster models. Additionally, a Mulliken charge analysis of the Mg-MOF-74-Dzubak cluster yields larger partial charges for the surrounding Mg atoms as compared to the Mg-MOF-74-Yu model, which may help explain the increased binding and shortened Mg–O contact in the Mg-MOF-74-Dzubak model. As for binding energies, there are substantial differences in binding energies between the various cluster models, with Mg-MOF-74-small severely underbinding These results for the Mg-MOF-74-small cluster indicate the inadequacy of such a small model, and likely explain the underprediction of the CO_2 adsorption isotherm from above. The Mg-MOF-74-Dzubak model shows best energetic agreement with the periodic system. Nevertheless, some of the Mg-MOF-74-Dzubak binding energy arises from truncation effects (as described above), and the energetic agreement is thus due (at least in part) to error cancellation. Indeed, some of the binding energy in the periodic system arises from (attractive) long-range interactions, and thus we should expect to see a cluster model somewhat underpredict the binding energy. Primarily for its good agreement in binding geometries, and reasonable agreement in binding energy, we opt to use the Mg-MOF-74-Yu cluster model for the remainder of this work.

5.6.2 Final Mg-MOF-74 CO_2 Adsorption Isotherm

Using our new Mg-MOF-74-Yu cluster model, we next attempted to fit new force field paramters for Mg. As discussed earlier, and because of the size of this new cluster (60 atoms), LMO-EDA-PBE0-D2 calculations became cost prohibitive in all but the smallest VDZ basis set and for a limited set of points. Starting from the minimum energy configuration shown in Fig. 5.5, we fit Mg parameters to a 12-point scan along the Mg-O bond vector, with fit results shown in Fig. 5.6. Interestingly, though the functional form used in this fit was sufficient to accurately paramterize the interaction energies in the Mg-MOF-74-small cluster, the same force field methodology proved unsuccessful in paramterizing Mg-MOF-74-Yu

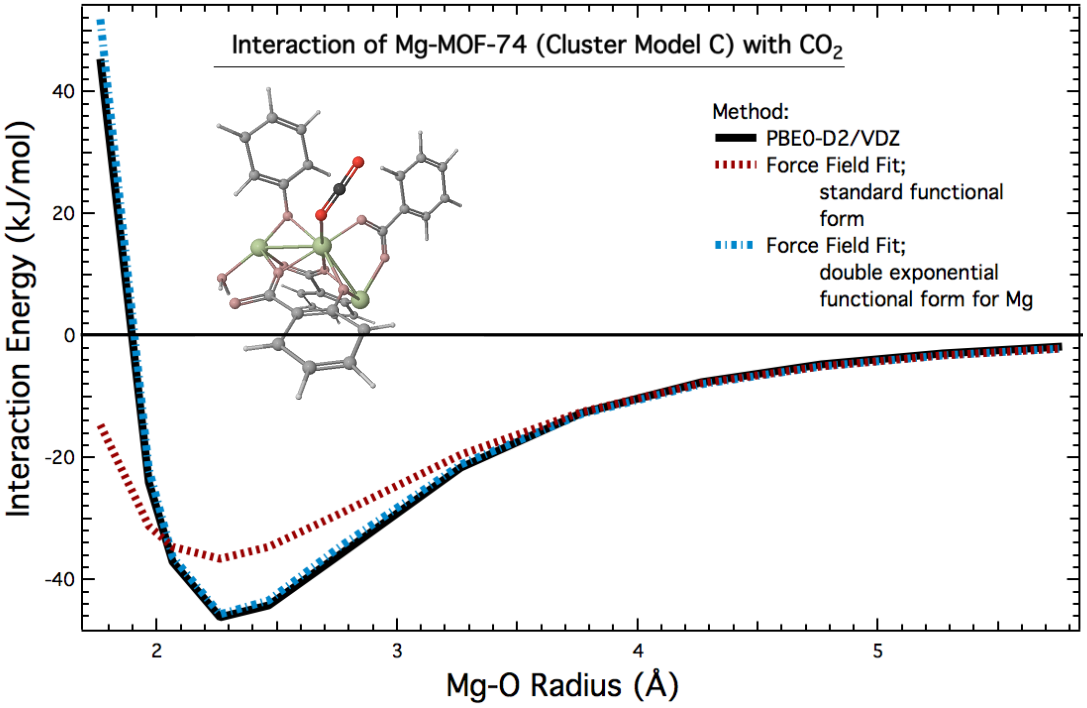


Figure 5.6:

interactions. We knew at the time that this inaccuracy was probably a consequence of uncertainties in correctly parameterizing the Mg short-range exponent. (See Chapter 3 for a full discussion of new methods for parameterizing the short-range potential.) Nevertheless, because the Slater-ISA methodology for short-range interactions had not yet been developed, we opted instead to fit the Mg interactions to a double exponential functional form, with each exponent corresponding to the ionization potential for either Mg^{+} or Mg^{2+}, the two chemical environments that we thought might correctly represent the Mg cation. As shown in Fig. 5.6, this form could excellently reproduce the Mg-MOF-74-Yu model PES.

Using the double exponential functional form from above, we recomputed the Mg-MOF-74 CO_2 adsorption isotherm. Before comparing to experiment, and as recommended by others,[26] we scaled the experimental isotherm in order to account for the pore blocking effects that are common in the M-MOF-74 series. Using this

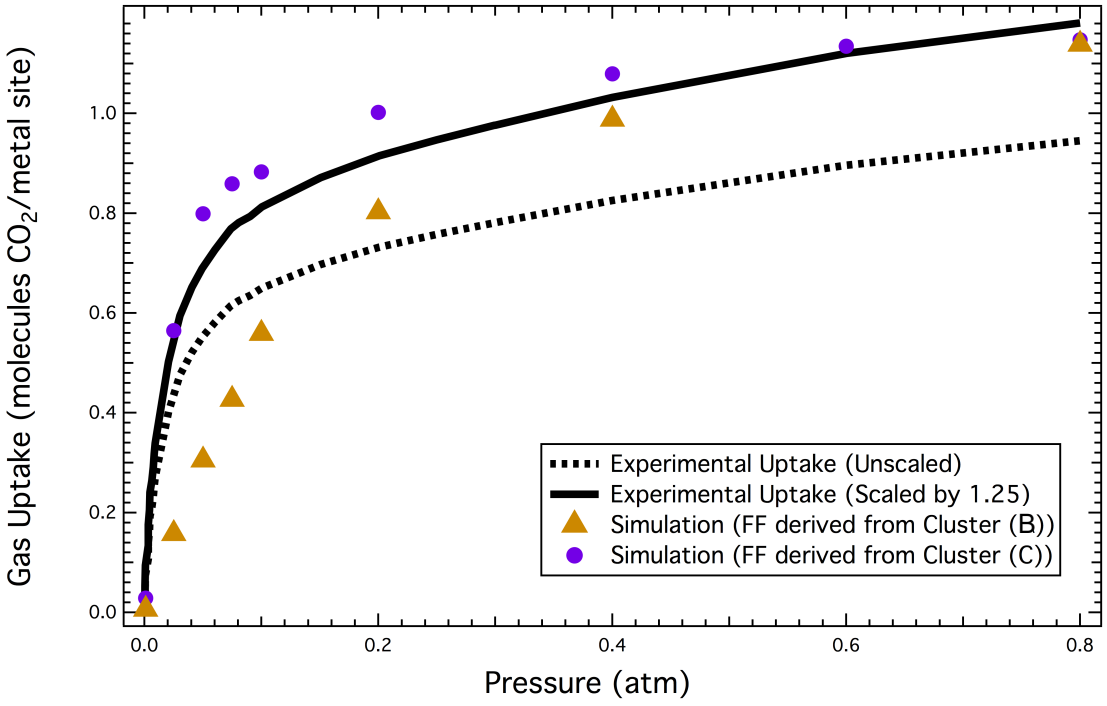


Figure 5.7:

scaled isotherm, we then obtain excellent agreement between our model potential and experiment (Fig. 5.7). Crucially, this accuracy is seen both at low- and high-pressure ranges, indicating the accuracy of the force field in modeling both the strong Mg–CO_2 binding as well as the weaker physisorption regime.

5.6.3 Transferability to Other Adsorption Isotherms

In addition to using our Mg paramters to compute the CO_2 adsorption isotherm, we also used our Mg force field in conjunction with the N_2 parameters developed by Yu et al.[27] to predict the N_2 adsorption isotherm. These predictions were generally poor, and results are not shown. Nevertheless, the poor N_2 results suggest a lack of transferability of our Mg parameters, possibly (and as discussed in the section on Future Work) due to unphysical double-exponential functional form used to

parameterize Mg.

5.6.4 Transferability to Other M-MOF-74 systems

As a second test of transferability, we also attempted to develop force fields for other compounds in the M-MOF-74 series, starting with Co-MOF-74. Unfortunately, the open-shell nature and increased electron count of Co-MOF-74 made LMO-EDA calculations computationally prohibitive for any reasonable basis set, and these systems were not investigated further.

5.7 Conclusions

5.8 Future Work

Practical issues – don't use GAMESS – need more memory

Theoretical issues: – Functional form was bad – Hard to choose an appropriate cluster model

Fixes: – try second-generation ALMO-EDA? – try SAPT correction scheme? – for this system, maybe s2 approximation for the induction energy is to blame

5.A Force Field Parameters for CO_2 and Mg-MOF-74

Final force field parameters, fit using the double-exponential functional form above and the Mg-MOF-74-Yu cluster model, for CO_2 and Mg-MOF-74. These parameters should be read in as input into our group's lattice simulation code, see `http://schmidt.chem.wisc.edu/montecarlosimulationcodes` for details.

```
lennard_jones_type 1 "1 for buckingham, 2 for lennard jones"

" parameters are listed as charge, A,B,C, polarizability for buckingham, or as charge, epsilon,sigma, junk, polarizability
" units are A:kJ/mol,  B:A^-1,  C: KJ/mol*A^6   epsilon:KJ/mol,   sigma:A,  polarizability: A^3 "
" polarizability is defined as q^2/k, spring constant is set to .1*1.8897^3 e^2/A^3"
```

```
solute_species
atom_type_parameters  ( q, Aexch, Aelec, Aind, Adhf, Adisp, C6, C8, C10, alpha )
2
C0      0.6573800       95510.43       -27846.98       -13425.1  2065.044
        0.0     6.891E02 0.0 0.0 1.1926153
O0      -0.328690       521902.066     -163908.84      -4475.8095       -26042.04
        0.0     1.8341E03 0.0 0.0   0.9009290

solute parameters for framework cross terms  ( Aexch, Aelec, Aind, Adhf, Adisp, B, C6, C8, C10, C12 )
C0 74376.65 24130.18     12513.37          795.36
        0.0 3.4384        1147.41867     6329.41038    29659.50100    183546.714
O0 354373.13 108208.5      2544.89        -18178.7
        0.0 3.7795   867.27598     4266.54582    28761.10636  132581.301

solute dhf cross terms (check code for input format if more than one cross term)
-6124.0

solute-solute exponents ( Bii, Bij, Bjj)
3.5105206 3.6993494 3.9288490


framework_species
atom_type_parameters   ( q, Aexch, Aelec, Aind, Adhf, Adisp, B, C6, C8, C10, C12, alpha )
9
C1  -0.1639350  612892.611  229850.679   -6511.529  -60036.930
        0.000       3.438     1628.820002     6821.530007    44464.989999   193602.980000  0.0
H1   0.2637500    8538.651    1678.771    -612.739    -502.639
        0.000       3.778      129.439978      679.640001     4995.299998        0.000000  0.0
C2  -0.3191850  612892.611  229850.679   -6511.529  -60036.930
        0.000       3.438     1628.820002     6821.530007    44464.989999   193602.980000  0.0
C3   0.4964850  612892.611  229850.679   -6511.529  -60036.930
        0.000       3.438     1628.820002     6821.530007    44464.989999   193602.980000  0.0
O3  -1.0339500    3398.424    1965.168    -182.412    -178.025
        0.000       2.457     2237.635879    29956.090890   561056.184030  7451461.601923  0.0
C4   0.9468300  263600.161  112896.479     -11.681   -2837.170
        0.000       3.438      772.870024     2349.180008    27539.189998   102366.260000  0.0
O4  -0.8903225  656757.170  174054.351  -45410.640  -33954.271
        0.000       3.779     1799.560008    11576.089993    50164.639999        0.000000  0.0
Mg   1.5906500     917.037    2417.463  -12542.799       0.000
        0.000       2.834      630.723467        0.000000        0.000000        0.000000  0.0
Du   0.0000000   29176.333       0.000  142260.481       0.000
        0.000       3.973        0.000000        0.000000        0.000000        0.000000  0.0
```

5.B Simulation Parameters CO_2 Adsorption in Mg-MOF-74

These parameters should be read in as input into our group's lattice simulation code, see `http://schmidt.chem.wisc.edu/montecarlosimulationcodes` for details.

```
Simulation Methodology
energy_decomposition yes        ! yes for our force fields, no for UFF LJ, etc.
solute_cross_parameter_set yes    ! this should be set to yes if using different solute parameters for solute-solute and so
C8_10_dispersion_terms yes    ! set to yes if using C8, C10 dispersion terms as in our force fields
C12_dispersion yes
electrostatic_type pme    ! either "pme" for particle-mesh ewald, "cutoff", or "none"
lj_comb_rule ZIFFF    ! "opls" or "standard" for lj, "standard" or "ZIFFF" for bkghm

Simulation Parameters
temperature 296.0 ! temperature in Kelvin
too_close 1.8 ! reject move if molecules are within this separation in Angstroms.  helpful to avoid unnecessary energy calc
lj_bkghm 1 ! 1 for bkghm force field, 2 for lj
screen_type 1 ! screening for coulomb potential: 0 = no screening, 1 = Tang-Toennies type screening for our force fields
springconstant 0.1 ! spring constant for drude oscillators (au). set to 0.1 for our CO2/N2 force fields
thole 2.0 ! thole parameter for intra-molecular drude oscillator screening.  Set to 2.0 for our CO2/N2 force fields.
drude_simulation 1 ! set to 1 if drude-oscillators are being used, 0 otherwise
pme_grid                        100             ! size of the pme grid
alpha_sqrt                      0.6             ! alpha sqrt for the electrostatic interactions
lj_asqrt                        0.6             ! alpha sqrt for the pme dispersion
lj_cutoff                       7.5             ! cutoff for long range LJ or C6,C8,C10 dispersion interactions
ewald_cutoff                    5.0             ! cutoff for real space pme
cav_grid_a                      30
cav_grid_b                      30
cav_grid_c                      30
na_nslist                       30              ! neighbour list searching grid
nb_nslist                       30              ! neighbour list searching grid
nc_nslist                       30              ! neighbour list searching grid
orientation_try                 2000            ! max number of orientation samplings
REL_THRSH                       0.05            ! sampling threshold
ABS_THRSH                       3.0
BZ_CUTOFF                       100.0
```

6 BENCHMARK DATABASE FOR AB INITIO FORCE FIELD DEVELOPMENT

Part IV

Practical Matters

7 WORKFLOW FOR INTERMOLECULAR FORCE FIELD DEVELOPMENT

7.1 Overview

7.2 Codes

7.2.1 Molpro

7.2.2 CamCASP

7.2.3 Scripts

7.3 Geometry Generation

7.4 SAPT Calculations

7.5 Monomer-Based Parameterization

7.5.1 Multipoles

7.5.2 ISA Exponents

7.5.3 Polarization Charges

7.5.4 Dispersion Coefficients

Jesse's Method

Alston's Method

7.6 Dimer-Based Parameterization

Refer to POInter Code section

8 POINTER: A PROGRAM FOR INTERMOLECULAR FORCE FIELD OPTIMIZATION

8.1 Overview

8.2 Documentation

8.3 Examples

Part V

Conclusions and Future Work

9 FUTURE WORK

10 CONCLUSIONS

BIBLIOGRAPHY

[2] Stone, A. J. *The Theory of Intermolecular Forces*, 2nd ed.; OUP Oxford, 2013.

[3] Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. *Science (80-.).* **2013**, *341*, 1230444–1230444.

[4] Millward, A. R.; Yaghi, O. M. *J. Am. Chem. Soc.* **2005**, *127*, 17998–17999.

[5] Dietzel, P. D. C. et al. *J. Mater. Chem.* **2009**, *19*, 7362.

[6] Dzubak, A. L.; Lin, L.-C.; Kim, J.; Swisher, J. a.; Poloni, R.; Maximoff, S. N.; Smit, B.; Gagliardi, L. *Nat. Chem.* **2012**, *4*, 810–816.

[7] Czaja, A. U.; Trukhan, N.; Müller, U. *Chem. Soc. Rev.* **2009**, *38*, 1284.

[8] Krishna, R.; van Baten, J. M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10593–10616.

[9] Getman, R. B.; Bae, Y.-s.; Wilmer, C. E.; Snurr, R. Q.; Carlo, M. *Adsorpt. J. Int. Adsorpt. Soc.* **2012**, 703–723.

[10] Yazaydin, a. O.; Snurr, R. Q.; Park, T.-H.; Koh, K.; Liu, J.; Levan, M. D.; Benin, A. I.; Jakubczak, P.; Lanuza, M.; Galloway, D. B.; Low, J. J.; Willis, R. R. *J. Am. Chem. Soc.* **2009**, *131*, 18198–9.

[11] Valenzano, L.; Civalleri, B.; Chavan, S.; Palomino, G. T.; AreaÌ'n, C. O.; Bordiga, S. *J. Phys. Chem. C* **2010**, *114*, 11185–11191.

[12] McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. C* **2012**, *116*, 14031–14039.

[13] McDaniel, J. G.; Yu, K.; Schmidt, J. R. *J. Phys. Chem. C* **2012**, *116*, 1892–1903.

[14] McDaniel, J. G.; Li, S.; Tylianakis, E.; Snurr, R. Q.; Schmidt, J. R. *J. Phys. Chem. C* **2015**, *119*, 3143–3152.

[15] Lao, K. U.; Schaeffer, R.; Jansen, G.; Herbert, J. M. *J. Chem. Theory Comput.* **2015**, 150417132228001.

[16] Pastorczak, E.; Corminboeuf, C. *J. Chem. Phys.* **2017**, *146*, 120901.

[17] Żuchowski, P. *Chem. Phys. Lett.* **2008**, *450*, 203–209.

[18] Horn, P. R.; Head-gordon, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23067–23079.

[19] Su, P.; Li, H. *J. Chem. Phys.* **2009**, *131*, 014102.

[20] Chen, Y.; Li, H. *J. Phys. Chem. A* **2010**, *114*, 11719–24.

[21] Su, P.; Jiang, Z.; Chen, Z.; Wu, W. *J. Phys. Chem. A* **2014**, *118*, 2531–42.

[22] Fedorov, D. G.; Kitaura, K. **2006**,

[23] Yu, K.; Kiesling, K.; Schmidt, J. R. **2012**,

[24] Verma, P.; Xu, X.; Truhlar, D. G. **2013**,

[25] Valenzano, L.; Civalleri, B.; Sillar, K.; Sauer, J. **2011**, 21777–21784.

[26] Haldoupis, E.; Borycz, J.; Shi, H.; Vogiatzis, K. D.; Bai, P.; Queen, W. L.; Gagliardi, L.; Siepmann, J. I. *J. Phys. Chem. C* **2015**, *74*, 150616135429005.

[27] Yu, K.; McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. B* **2011**, *115*, 10054–10063.

ACRONYMS

C | D | E | M | P | S

C

CUS Coordinatively-Unsaturated. iv, 10, 11, 13–15

D

DFT Density Functional Theory. 11, 13–15

DFT-SAPT Density Functional Theory Symmetry-Adapted Perturbation Theory. 11–13

DMA Distributed Multipole Analysis. 12

E

EDA Energy Decomposition Analysis. 14, 15

M

MOF Metal-Organic Framework. iv, 9–15

P

PES potential energy surface. vii, 9, 12, 22, 38

S

SAPT Symmetry-Adapted Perturbation Theory. 9, 11–16, 24, *Glossary:* SAPT

GLOSSARY

C | L | S

C

CCSD(T) Coupled Cluster methods including singles, doubles, and perturbative triples excitations. CCSD(T). Given a sufficiently large (aVQZ or better) basis set, can be used as a 'gold-standard' estimate of the exact potential energy surface. 9, 13

CCSD(T)-f12 Explicitly-correlated CCSD(T). Given a sufficiently large (aVDZ or aVTZ) basis set, used throughout this work as a 'gold-standard' estimate of the exact potential energy surface. 12–14, 19

L

LMO-EDA FILL . 9, 15–18, 21, 24

S

SAPT Symmetry-Adapted Perturbation Theory, a perturbative treatment of intermolecular interactions which is pretty cool. 9