JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

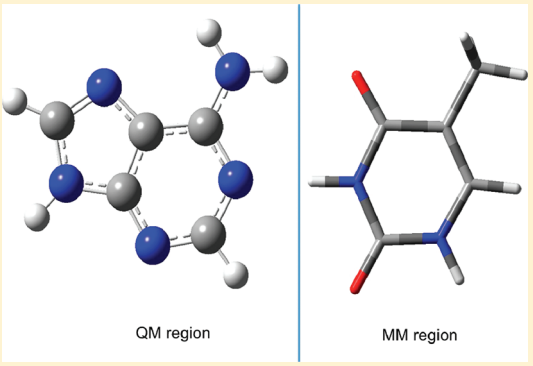
Assessment of Weak Intermolecular Interactions Across QM/MM Noncovalent Boundaries

Sadhana Kumbhar, Frank D. Fischer, and Mark P. Waller*

Theoretische Organische Chemie, Organisch-Chemisches Institut der Universität Münster, Corrensstrasse 40, D-48149 Münster, Germany

Supporting Information

ABSTRACT: An assessment of a number of quantum mechanical/molecular mechanical (QM/MM) combinations was performed for weak intermolecular interactions across noncovalent QM/MM 'boundaries'. The popular S22 data set, comprising of a number of weak hydrogen-bonded, dispersion-bound and complexes with mixed interactions was used for the assessment. A range of QM methods was combined with a number of popular MM force fields. The single-point interaction energies, at reference geometries, are presented as deviations to accurate CCSD(T)/CBS reference values. This investigation employed both additive and subtractive QM/MM schemes. The density functional has only a negligible effect; the choice of basis set was also negligible in terms of accuracy. The importance of selecting the most appropriate MM force field for accurately describing interactions at the noncovalent 'boundary' region has a dramatic effect on the accuracy.



INTRODUCTION

Large biomolecules (>1000 atoms) such as proteins or long polymeric DNA are simply too large to study them using density functional theory (DFT), although significant progress is being made within the linear scaling DFT community. The quantum mechanical (QM)/molecular mechanical (MM) method is one approach to circumvent the computational demands required to study extremely large systems. For this and other reasons, QM/MM is an attractive and reliable method that has become popular and applied effectively to solve present-day problems. In addition to the scalability of the solution, there exists a wealth of information such studies can reveal, e.g., mechanistic details and spectroscopic parameters.

The QM/MM scheme is available in a number of different "flavors", with a number of good reviews being published.[1,2] The range of methods can be broadly grouped into two families: (1) subtractive and (2) additive.

Subtractive. One such subtractive method was proposed by Morokuma and given the acronym ONIOM[3] (our own N-layered integrated molecular orbital and molecular mechanics) and has been widely utilized.[4] It employs the following scheme

$$E(\text{ONIOM, real}) = E(\text{low, real}) + E(\text{high, model}) - E(\text{low, model})$$

where "high" represents the QM method and "low" represents an MM method, the model system is treated at the QM level, while the real (entire) system is treated with the MM method, and the MM contributions from the model system subtracted. This subtractive scheme is available within the popular Gaussian 09 suite, with the atom assignment being easily carried out using the GaussView GUI.[5] Other examples of software employing the subtractive scheme are QMMM[6] and COBRAMM.[7]

Additive. The additive scheme is generally given by the following expression

$$E_{\text{Total}} = E_{\text{Q.M}} + E_{\text{MM}} + E_{\text{Q.M/MM}}$$

The E_{QM} term is the electronic energy (i.e., expectation value of the selected Hamiltonian), the E_{MM} energy is the strain energy (for that particular forcefield), and $E_{\text{QM/MM}}$ is the interaction energy between the QM and the MM regions.

The coupling term between the QM and the MM regions ($E_{\text{QM/MM}}$) has been investigated using a number of different schemes of varying form and complexity. In a simple scheme

$$E_{\text{Q.M/MM}} = E_{\text{elect}} + E_{\text{vdW}}$$

where E_{elect} is the interaction between a set of point charges (generally fixed) representing the MM region and the QM density (or QM density represented as point charges (PCs)) and E_{vdW} is typically computed using a standard Lennard−Jones (LJ) 6-12 potential. This type of approach for modeling the boundary region is perhaps justified based upon molecular mechanics investigations that have been shown to reasonably describe weak interactions with a relatively simple set of functions, see below.[8]

Received: August 31, 2011
Published: December 21, 2011

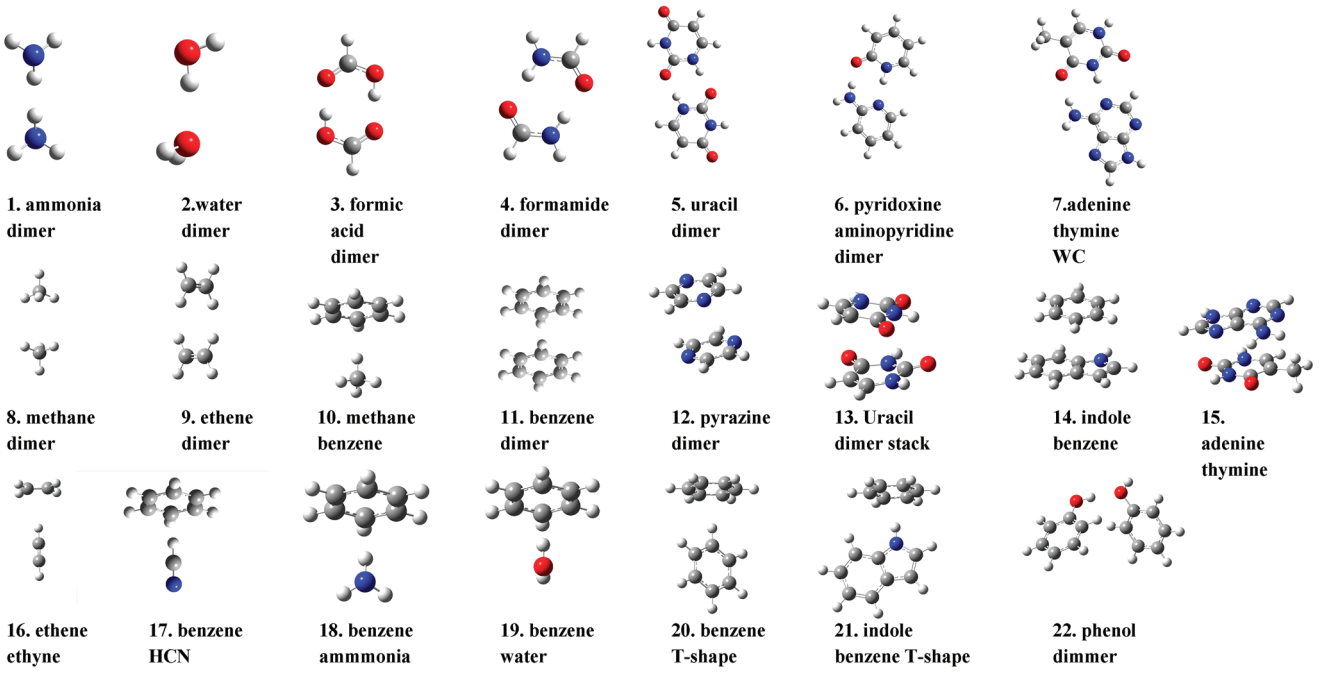


Figure 1. S22 dimer data set, including the molecular Mol IDs used throughout. Note: The first molecule name is in the QM region, and the second molecule for heterodimers was placed in the MM region.

The additive QM/MM approach is available in the ChemShell suite,[9] AMBER,[10] or with the PUPIL package.[11]

Within the well-established approach of QM/MM modeling, the primary region of interest (e.g., active site) is placed into a high layer (treated with quantum chemical methods) and the remaining system is placed into the low region (molecular mechanics). The 'primary' region in the higher layer should not be too large as the quantum chemical methods can be computationally expensive. That leaves the larger component of the system in the lower layer(s) to act as mechanical and electrostatic perturbations to the high(er) layer. Typically, great care must be taken at the border region where a covalent bond is "cut". If such a cut across a 'covalent' boundary is unavoidable, relatively nonpolar bonds should be cut preferentially.

Weak intermolecular (noncovalent) interactions between molecules in the gas phase are of popular interest. These interactions are essential for the quantitative description and understanding of complex molecular aggregates in physics (e.g., surface science), chemistry, and molecular biology. The same interactions also occur in an intramolecular fashion between atoms (or groups) in a molecule. Herein, we consider only intermolecular cases, more specifically dimers. However, one should keep in mind that many of the conclusions would transfer to intramolecular cases, e.g., for the structure of DNA which is a delicate interplay between intrastrand $\pi\cdots\pi$ stacking and interstrand hydrogen bonding.

The data set to be investigated, commonly referred to as S22,[12,13] comprises 22 dimers composed of typical organic molecules. The S22 data set, see Figure 1, has turned out to be a very useful tool in computational chemistry. It has been used to assess SCS-MP2[14] and create/refine new empirical corrections for density functional theory (DFT-D),[15,16] and even has been used with the quantum Monte Carlo method.[17] An additional advantage of using the S22 set of 22 dimers is that the issues regarding link atoms need not be considered, subsequently simplifying the analysis of the QM/MM boundary. The data set can be separated naturally into three subcategories: (a) H bonding, (b) dispersion bound, and (c) mixed complexes. Therefore, the S22 data set provides a natural way to separate the relative strength and weaknesses of the methods for these fundamentally different types of noncovalent interactions.

COMPUTATIONAL DETAILS

The data set containing the geometries of the S22 data set was taken from the Benchmark Energy & Geometry Database (BEGDB).[13] ChemShell was utilized as a QM/MM software suite, which employs an additive coupling scheme.[18] Turbomole[19] with the BP86,[20−22] BLYP, and B3LYP[23,24] functionals and the SVP, TZVP, and QZVP basis sets was applied to the QM regions using the resolution of the identity approximation together with suitable auxiliary fitting functions from the Turbomole library.[25] DL_POLY[26] was used to compute the MM energy and gradients using the CHARMM27 force field.[27,28] Either an electrostatic embedding scheme or mechanical embedding scheme was applied. The subtractive QM/MM calculations were carried out using the Gaussian 09 suite.[29] Morokuma's implementation of the ONIOM method was used. The QM and MM methods were used as implemented in the standard G09 suite using the same functionals and basis sets as above.[30] The interaction energy is simply defined as the difference between the total QM/MM energy and the sum of the QM and MM energies

$$\Delta E_{\mathrm{int}} = E_{\mathrm{Q\,M/MM}}(\mathrm{A,B}) - (E_{\mathrm{Q\,M}}(\mathrm{A}) + E_{\mathrm{MM}}(\mathrm{B}))$$

where A and B are the monomers within the S22 data set. A list of atom assignments, layer assignments, and point charges (PCs) is provided in the Supporting Information.

The number of possible QM and MM combinations and permutations is too large; therefore, an exhaustive search of all possible combinations is not realistic. Furthermore, we strongly

dx.doi.org/10.1021/ci200406s |J. Chem. Inf. Model. 2012, 52, 93–98

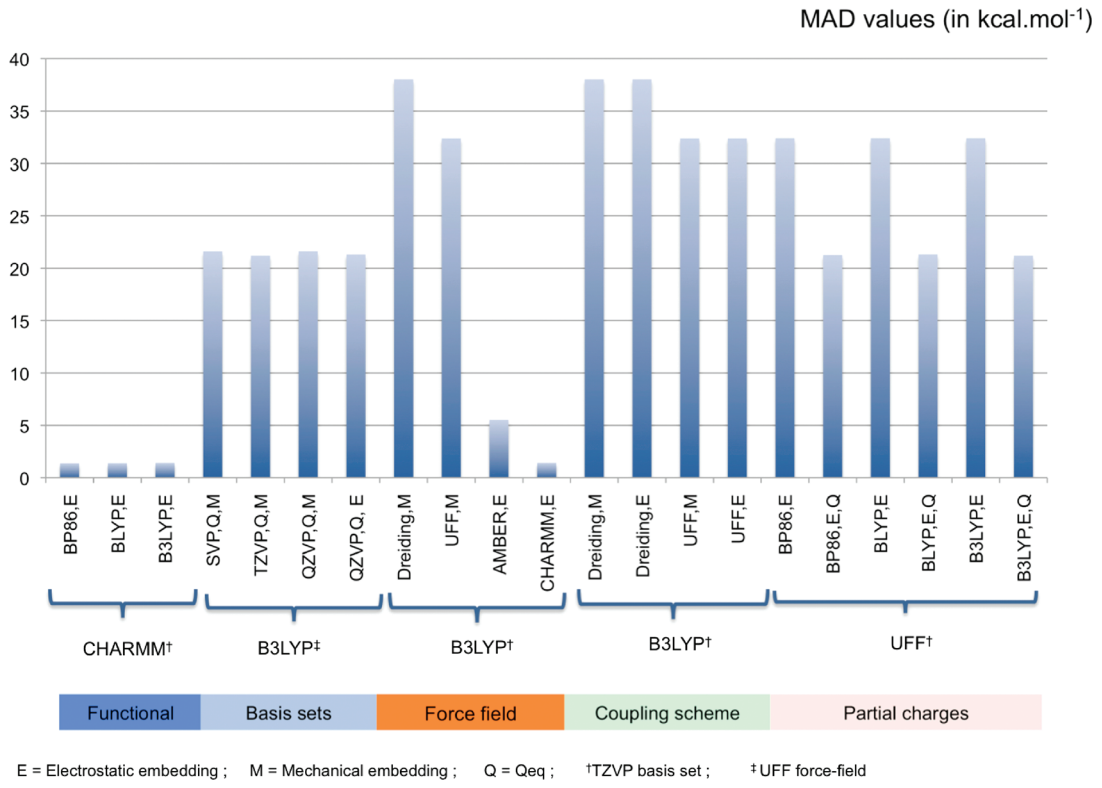


Figure 2. MAD values of different methods according to the types of QM/MM variables.[32]

believe that validation studies are needed for any specific application. The goal of this work is to simply present the general trends and highlight some spectacular failings of some example combinations to act as a warning for future studies.[31] One extremely important point is that the following results do not present results that include geometry optimization. This is because the benchmark values are computed at a consistent geometry, and this would not provide a fair comparison to the existing collated literature results.

RESULTS

The performance of all the QM/MM methods in Figure 2 is compared to the reference data CCSD(T)/CBS method.[12] The CCSD(T)/CBS level of theory is commonly considered to be the "Gold standard" for computing the interaction energies of weak intermolecular interactions.

FUNCTIONALS

The BP86, BLYP, and B3LYP functionals with a triple-ζ basis set (TZVP) were individually combined with the CHARMM force field using an additive scheme (including electrostatic embedding) to compute the interaction energies for the entire S22 data set. These three functionals were chosen as they are among the more popular choices from the lowest three rungs of the Jacob's ladder classification scheme. The individual deviations in interaction energies to the benchmark values of Hobza et al. are listed in Tables S2 and S3 of the Supporting Information, along with the overall mean absolute deviation (MAD) and maximum absolute deviation (MAX). There is a rather uniform spread of errors across the S22 data set, that is, no systematic problem is observed as evidenced by excessively high values for any particular type of interaction (i.e., any individual subset of the S22 data set). For reasons of brevity and clarity, only the MADs are presented in Figure 2. One can conclude that the choice of density functional has only a negligible influence on the MADs, much smaller than the observed changes when one uses pure DFT calculations.

BASIS SETS

The double-ζ (SVP), triple-ζ (TZVP), and quadruple-ζ (QZVP) basis set series with the B3LYP functional along with the UFF force field are shown in Figure 2 to have very little influence on the MAD for the QM/MM methods for the S22 data set. The fact that there is so little difference between the basis sets is an important finding. Adoption of an incomplete basis set is a notorious problem in purely quantum mechanical calculations. For such QM calculations, a number of extrapolation energy schemes[33] have been recommended in conjunction with basis sets that are designed to converge correctly to the CBS limit.[34] Importantly, the QM/MM interaction energy across the boundary is obviously free from basis set superposition error. Addition of polarization functions (see Supporting Information) also does not significantly improve the agreement between QM/MM interaction energies and reference values. Furthermore, this effect is irrespective of whether an electrostatic or mechanical embedding scheme is employed.

FORCE FIELDS

In Figure 2 there are a number of force fields (i.e., UFF, Dreiding, AMBER, and CHARMM) all combined with the B3LYP-TZVP level of theory. The biologically orientated CHARMM and AMBER force fields clearly outperform the UFF and

dx.doi.org/10.1021/ci200406s |J. Chem. Inf. Model. 2012, 52, 93–98

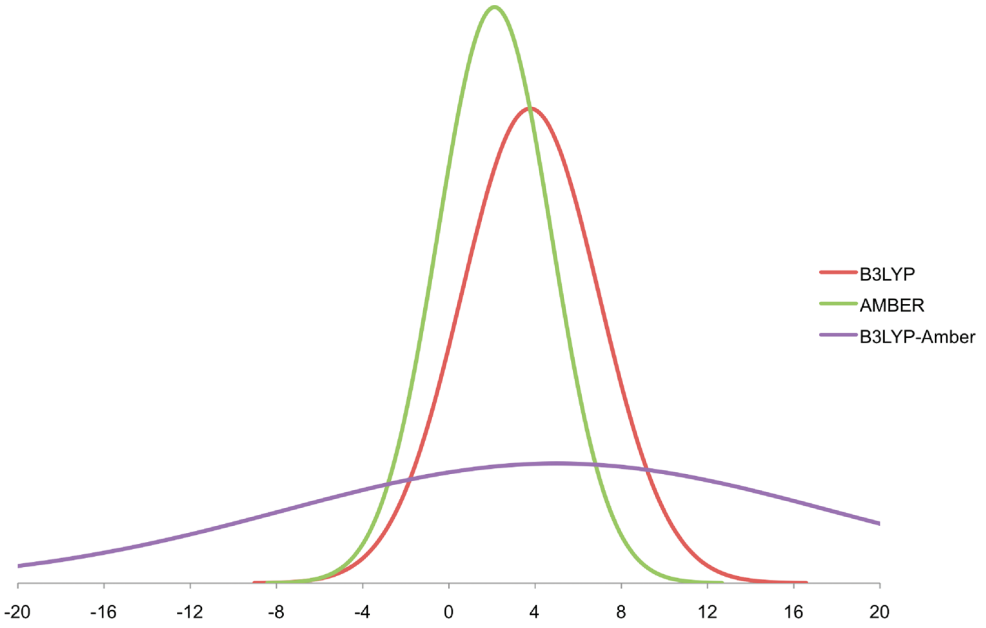


Figure 3. Normal error distribution for a number of commonly applied levels of theory.[8,37]

Dreiding force fields when considering the MADs for the S22. This is perhaps to be expected based on their intended areas of application and given the constituents of the S22 data set.

The MADs for CHARMM with BP86/TZVP, BLYP/TZVP, and B3LYP/TZVP are ∼1.38−1.43 kcal·mol^{-1}. These MADs are actually higher than observed for most pure DFT-D (S22 MADs ≈ 0.3−0.7 kcal.mol^{-1}).[16] The MADs for the different types of interactions are (a) hydrogen bonded > (b) dispersion dominated > (c) mixed interactions. The reason that the hydrogen-bonded systems (Mol id 1−7) are too strongly bound can at least partially be understood given the known problems of force fields describing hydrogen-bonding interactions. The QM/MM implementations used in this study contained no explicit corrections for hydrogen bonding across the QM/MM boundary, although future corrections may be explored in this direction (see, for example, Korth et al. 2010[35]). The dispersion-dominated complexes (Mol id 8−15) are in good agreement with the CCSD(T) benchmark values, i.e., MAD ≈ 1.5 kcal·mol^{-1}. These MADs are similar to the values obtained by Paton and Goodman using a number of popular force fields, which is unsurprising as the Lennard−Jones (LJ) potential used here in the QM/MM scheme is also employed in many of the popular force fields. Of note are the MADs obtained for the mixed systems (0.6−0.7 kcal·mol^{-1}), which are the lowest of the three subsets. However, the average error expressed as a percentage of the total reference interaction energies is 15% for hydrogen bonding, 31% for dispersion-bound complexes, and finally 16% for mixed complexes. Overall, we may conclude that when applying the QM/MM methodology with the CHARMM force field one can expect an average error of around 20% for the interaction energies of weak intermolecular dimers across the noncovalent boundary region. Other force fields such as UFF and Drieding overestimate hydrogen-bonded dimers and dispersion-bound and mixed systems.

EMBEDDING SCHEMES

In Figure 2 it is immediately clear that the MADs are not dramatically affected by the choice of embedding scheme. The study trialed both mechanical and electrostatic embedding schemes. Importantly, all of the hetero- and homodimers of the S22 data set are small organic molecules. In the mechanical embedding approach one must partition the QM molecular density into atomic contributions in order to compute a set of partial atomic charges (PAC). These partial atomic charges are subsequently used in the electrostatic embedding scheme, and these PACs would be expected to be more accurate for small organic dimers than for metallic systems. Therefore, the generality of the above conclusions about the mechanical vs electrostatic embedding scheme is perhaps limited to small organic molecules similar to those in the S22 data set. An important practical consideration is the significant extra computational cost of including the point charges of the MM region into the QM Hamiltonian when adopting the electrostatic embedding.

PARTIAL CHARGES

A particular set of partial charges employed within a force field are a result of painstaking parametrization, which is reliant on a delicate balance between the forces when a particular force field is developed. Herein, we are interested in the QM/MM boundary and are therefore no longer subject to conforming to such a balance as were are interested in the interaction across the regions. To investigate this possibility, we trialed replacing the pre-existing parametrized partial charges from a force field with a semiempirical scheme such as the QEq method within Gaussian. It is clear from Table S2, Supporting Information, that the MAD values are lower with QEq charges than those in the cases without QEq. Although one can rationalize this as lifting the constraints that the partial charges need to be "internally" consistent within a force field, one must keep in mind that the set of semiempirical (SE) charges are themselves not parametrized to fit a plethora of density functionals. Therefore, we conclude that although there is a decrease in the MAD across the noncovalent QM/MM border, one should not encourage their adoption for QM/MM applications without prior validation. This is because the set of

QEq would almost certainly impose a detrimental effect on the underlying force field (i.e., within the intra MM region).

LAYER ASSIGNMENT

The results presented in Figure 2 employed either the subtractive scheme as proposed by Morokuma's ONIOM method or the additive scheme within CHEMSHELL. In both cases, the layer assignment was simply the first molecule from the S22 data set being placed into the QM region and the second molecule in the MM layer. To ensure our results are not too biased by our arbitrary layer assignment, we switched assignments of all the heterodimers. The effect was very small (see S1 of Supporting Information), and therefore, we believe that this has little influence on our conclusions.

DISCUSSION

For quantum mechanical methods (including DFT) there exists a good correlation between the scaling of a many popular methods and the accuracy of the method for the S22 data set, that is, the MADs are typically high-level ab inito methods < perturbation theory < density functional theory < Hartree−Fock < semiempirical < force fields. Importantly, the explicit corrections for dispersion and hydrogen bonding made for DFT, HF, and SE methods have naturally lead to improvements above and beyond the standard method and generally at little extra computational cost. The reliability of a number of force fields was investigated by Paton and Goodman in 2009,[8] who showed that a number of popular force fields underestimate the strength of the hydrogen-bonding interactions. Interestingly, the energetics of the dispersion-bound complexes were found to be in good agreement with the benchmark values. The AMBER force field gave a MAD of 2.12 kcal·mol^{-1} and outperforms 'traditional' DFT calculations, i.e., DFT, which is uncorrected for the well-known dispersion problem, over the entries of S22 employing very large basis sets. However, the MADs are lower for the QM/MM calculations (1.37−1.43 kcal·mol^{-1}) than the BHandH functional (2.10 kcal·mol^{-1}) for example. This result is interesting as the dispersion is effectively modeled using a LJ potential in the QM/MM calculations, while the fundamental (well-known) dispersion problem for traditional DFT was shown to be alleviated to some extent within BHandH (MAD 0.7 kcal.mol^{-1}).[36] However, the high level of exact exchange has a detrimental effect on describing the hydrogen-bonded systems (MAD 4.8 kcal·mol^{-1}), therefore resulting in inferior performance of BHandH over the QM/MM calculations. To pictorially depict the magnitude of the differences between QM methods (B3LYP), the QM/MM method (B3LYP/Amber), and the MM method (Amber) the normal error distributions are presented in Figure 3.[8,37]

CONCLUSIONS

Herein we investigated the accuracy of QM/MM methods across noncovalent boundary regions. The ultimate goal of this investigation was not to conduct a systematic study on all possible QM/MM combinations available within the literature but rather to probe the general performance of hybrid QM/MM methods for weak intermolecular interactions between small organic molecules. We make the following general observations from the results presented above.

a The choice of the density functional has only a negligible effect on the intermolecular interaction energies.
b The size of the basis set, within reason, does not heavily influence the MADs for the S22 data set either with or without electrostatic embedding.
c The choice of force field has a very large impact on the accuracy, and force fields that were developed to investigate biological systems, e.g., AMBER and CHARMM, in general out perform the more generally applicable force fields, e.g., UFF and Dreiding, for the S22 data set.
d Electrostatic embedding has little effect on the accuracy of a QM/MM approach when investigating weak intermolecular interactions of the type represented in S22 data set.
e The overall accuracy of the QM/MM methods trialed in this study for the S22 data set is below the accuracy of pure MM calculations.

Therefore, in light of the above observations, when investigating molecular systems with a QM/MM approach, care must be taken when interpreting the interaction energies across QM/MM boundaries given the magnitude of errors noted above.

ASSOCIATED CONTENT

Supporting Information. Detailed listing of the results for the S22 data set with reversed QM/MM assignments, along with a number of different functional/basis set combinations, and list of AMBER atom types and soft parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*Phone: (+49)-2518336514. E-mail: m.waller@uni-muenster.de.

ACKNOWLEDGMENT

This work was supported by the Deutsche Forschungsgemeinschaft in the framework of the SFB 858. S.K. would like to gratefully acknowledge the SFB858 for a doctoral stipend.

REFERENCES

(1) Senn, H. M.; Thiel, W. QM/MM studies of enzymes. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182–187.
(2) Senn, H. M.; Thiel, W. QM/MM-Methoden für biomolekulare Systeme. *Angew. Chem.* **2009**, *121*, 1220–1254. Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
(3) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. A New ONIOM Implementation in Gaussian 98. 1. The Calculation of Energies, Gradients and Vibrational Frequencies and Electric Field Derivatives. *J. Mol. Struct. (THEOCHEM)* **1999**, *462*, 1–21.
(4) For example:(a) Tschumper, G. S.; Morokuma, K. Gauging the applicability of ONIOM (MO/MO) methods to weak chemical interactions in large systems: hydrogen bonding in alcohol dimmers. *J. Mol. Struct. (THEOCHEM)* **2002**, *592*, 137–147. (c) Morokuma, K. ONIOM and Its Applications to Material Chemistry and Catalyses. *Bull. Korean Chem. Soc.* **2003**, *24*, 797–801. (d) Lundberg, M.; Morokuma, K. Protein Environment Facilitates O2 Binding in Non-Heme Iron Enzyme. An Insight from ONIOM Calculations on Isopenicillin N Synthase (IPNS). *J. Phys. Chem. B* **2007**, *111*, 9380–9389.
(5) http://www.gaussian.com/g_prod/gv5b.htm (accessed Oct 14, 2011).
(6) Lin, H.; Zhang, Y.; Truhlar, D. G. QMMM 1.3.8, Feb 11, 2011, http://comp.chem.umn.edu/qmmm/ (accessed Oct 14, 2011)

(7) Altoè, P.; Stenta, M.; Bottoni, A.; Garavelli, M. A tunable QM/MM approach to chemical reactivity, structure and physico-chemical properties prediction. *Theor. Chem. Acc.* **2007**, *118*, 219–240.

(8) Paton, R. S.; Goodman, J. M. How Reliable are Force Fields? *J. Chem. Inf. Model.* **2009**, *49*, 944–955.

(9) Sherwood, P.; deVries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schafer, A.; Lennartz, C. QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 1–28.

(10) Walker, R. C.; Crowley, M. F.; Case, D. A. The implementation of a fast and accurate QM/MM potential method in Amber. *J. Comput. Chem.* **2008**, *29*, 1019–1031.

(11) Torras, J.; de M. Seabra, G.; Deumens, E.; Trickey, S. B.; Roitberg, A. E. A Versatile AMBER-Gaussian QM/MM Interface Through PUPIL. *J. Comput. Chem.* **2008**, *29*, 1564–1573.

(12) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(13) http://www.begdb.com/ (accessed Oct 14, 2011)

(14) Antony, J.; Grimme, S. Is Spin-Component Scaled Second-Order Møller-Plesset Perturbation Theory an Appropriate Method for the Study of Noncovalent Interactions in Molecules? *J. Phys. Chem. A* **2007**, *111*, 4862–4868.

(15) Antony, J.; Grimme, S. Density functional theory including dispersion corrections for intermolecular interactions in a large benchmark set of biologically relevant molecules. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287–5293.

(16) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(17) Korth, M.; Lüchow, A.; Grimme, S. Toward the exact solution of the electronic Schrödinger equation for noncovalent molecular interactions: Worldwide distributed Quantum Monte Carlo calculations. *J. Phys. Chem. A* **2008**, *112*, 2104–2109.

(18) Corresponding to model B in:Bakowies, D.; Thiel, W. Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

(19) Ahlrichs, R.; Bär, M.; Häser, M.; Hornand, H.; Kölmel, C. Electronic structure calculations on workstation computers: the program system TURBOMOLE. *Chem. Phys. Lett.* **1989**, *162*, 165–169.

(20) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(21) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822–8824.

(22) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *34*, 7406.

(23) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(24) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti conelation energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(25) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Gaussian basis sets of quadruple zeta valence quality for atoms H-Kr. *Theor. Chem. Acc.* **1997**, *97*, 119–124.

(26) Smith, W.; Forester, T. DL_POLY_2.0: a general-purpose parallel molecular dynamics simulation package. *J. Mol. Graphics* **1996**, *14*, 136–141.

(27) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(28) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25*, 1400–1415.

(29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.1; Gaussian, Inc., Wallingford CT, 2009.

(30) Hehre, W. J.; Radom, L.; Schleyer, P. V. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley-Interscience: New York, 1986.

(31) On a technical note, the available software severely limits the possible combinations that are available "out of the box". Therefore, not all combinations are available without significant overhead of implementation time.

(32) It must be noted that new (highly accurate) values are available for the S22 dataset (Sherrill *J. Chem. Phys.* **2010**, *132*, 144104, which could potentially be used as reference values. The original values of Hobza et al. were used for consistency with the existing literature for QM and MM methods, alike.

(33) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

(34) For example, seeDunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(35) Korth, M.; Pitonak, M.; Rezac, J.; Hobza, P. A Transferable H-Bonding Correction for Semiempirical Quantum-Chemical Methods. *J. Chem. Theory Comput.* **2010**, *6*, 344–352.

(36) Waller, M. P.; Robertazzi, A.; Platts, J. A.; Hibbs, D. E.; Williams, P. A. Use of DFT for prediction of π-stacking interactions: Applications to benzenes, pyridines, and DNA bases. *J. Comput. Chem.* **2006**, *27*, 491–504.

(37) Goerick, L.; Grimme, S. A General Database for Main Group Thermochemistry, Kinetics, and Noncovalent Interactions - Assessment of Common and Reparameterized (meta)GGA Density Functionals. *J. Chem. Theory Comput.* **2010**, *6*, 107–126.