PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Dipole Moments of *trans*- and *cis*-(4-methylcyclohexyl)methanol (4-MCHM): Obtaining the Right Conformer for the Right Reason

Nathan J. DeYonker*, Katherine A. Charbonnet, and William A. Alexander*

Department of Chemistry
213 Smith Chemistry Building
The University of Memphis, Memphis, TN 38152

ABSTRACT

Accurate computational estimates of fundamental physical properties can be used as input in the myriad extant models employed to predict toxicity, transport, and fate of contaminants. However, as molecular complexity of contaminants increases, it becomes increasingly difficult to determine the magnitude of the errors introduced by ignoring the 3D conformational space averaging within group-additivity and semi-empirical approaches. The importance of considering 3D molecular structure is exemplified for the dipole moments of *cis* and *trans* isomers of (4-methylcyclohexyl)methanol (4-MCHM). When 10,000 gallons of 4-MCHM was spilled into the Elk River in January 2014, a lack of toxicological data and environmental partitioning coefficients hindered the immediate protection of human health and the local water supply in West Virginia, USA. Post-spill analysis of the contaminants suggested that the *cis* and *trans* isomers had observably different partitioning coefficients and solubility, and thus differing environmental fates. Obtaining high-quality dipole moments using *ab initio* quantum chemical methods for the isomeric pair was crucial in validating their experimental differences in solubility [*Environ. Sci. Technol. Lett.*, **2**, 127 (2015)]. The use of first principles electronic structure theory is further explored here to obtain accurate conformer relative energies and dipole moments of *cis*- and *trans*-4-MCHM. Overall, the MP2 aug-cc-pVDZ level of theory affords the best balance between accuracy and computational cost.

KEYWORDS

Environmental fate, conformational averaging, ab initio, dipole moment, (4-methylcyclohexyl)methanol, isomerization, solubility

I. Introduction

First-responders on the scene of an environmental or industrial disaster must quickly assess the presence of chemical, biological, or nuclear hazards and then begin the process of identifying the hazards. This process must occur while the first-responders are simultaneously protecting their own health and safety, attempting to minimize or prevent casualty to property and population, and preparing for decontamination.[1] The appropriate first-response is predicated on the availability of information about the contaminants. Yet comprehensive data regarding the toxicity and environmental fate and transport of many substances does not exist, and experimental determination of relevant properties is not feasible in a time-sensitive first-response situation. With more than 83,000 entities in the Toxic Substances Control Act (TSCA) Chemical Substance Inventory, there exist many compounds with little to no information on their molecular properties publicly or privately available.[2] This situation was highlighted by the January 2014 Elk River chemical spill in West Virginia, during which in excess of 10,000 gallons of a mixture of coal-washing chemicals were allowed to enter the Elk River, which serves as the source water of the largest domestic water system in the state, leaving over 300,000 residents without safe, clean water for drinking, bathing, or cooking for many days and weeks.[3,4,5]

The main contaminant of the 2014 Elk River spill, (4-methylcyclohexyl)methanol (or 4-MCHM) was patented twenty-five years ago for use in cleaning and purifying coal,[7] though the synthesis of 4-MCHM was introduced in the literature over a century ago.[8] Despite the compound's long history, little information was available regarding its physicochemical properties. In part, this is because crude 4-MCHM was grandfathered under the TSCA and regulatory reporting requirements did not apply. This scarcity of publicly available information led to much fear and confusion on the part of the public, and highlighted the need for rapid, on-demand methods for obtaining chemical information about properties of emerging contaminants in the environment and water supply. It was surprising to the analytical teams involved in the

"second response" to discover that two isomers of 4-MCHM exist in an approximately 2 (*trans*):1 (*cis*) ratio in crude MCHM from the manufacturer,[9] and that the two isomers have rather different water solubilities, air odors, and odor thresholds.[10, 11] It is thus expected that the two isomers of 4-MCHM also differ in their environmental fates and transport.[12-15]

It is clear that for many substances, especially those compounds used only for industrial purposes, gathering molecular property data *via* thorough experimental/analytical means once the crisis has become public knowledge may not be expedient. Concerns about proprietary information can muddle the situation further, e.g. trade secret additives to fossil fuels, fracking propant mixtures, etc. Beyond the classic "chemical spill", there are other conceivable situations where quick and reliable predictions of physicochemical molecular property data will be essential to protecting humanity and the environment. Due to time constraints during environmental incident response, various molecular properties are currently determined using quantitative structure-property relationships, or more simply estimated with atom/fragment group contribution methods. These methods are often heavily parameterized, and details of the calibration with expected levels of statistical accuracy and precision can be unclear, especially for proprietary or commercial packages. In emergency situations where quantitatively rigorous data must be obtained, it is unsettling that the validity of these models can be difficult to assess. However, tools such as the Estimation Programs Interface Suite (EPI Suite™), constructed and peer-reviewed by the US Environmental Protection Agency (EPA), can provide valuable assessments of immediate toxicological impact, as well as estimations of long-term atmospheric or environmental fate and transport. Detailed information about the benchmarking of the EPI Suite™ modules and comparison to other packages is available in both the published EPA Science Advisory Board final review,[6] as well as within the "help menu" of the software package itself. One glaring deficiency of EPI Suite™ (which is by no means unique to this package) presented itself in the aftermath of the Elk River chemical spill: it is only able to consider two-dimensional structure of contaminants.

Ab initio computational chemistry accounting for electron correlation is the most accurate and systematically improvable methodology for determining molecular structure and properties, and three-dimensional conformational structure is explicitly treated with quantum chemical techniques. However, the non-linear scaling [typically $O(N^3)$ and higher], the hefty memory and

disk space requirements, and the multi-layered approximations necessary to incorporate bulk or solvated corrections typically prevent density functional theory (DFT) or the more expensive/accurate coupled cluster theory from becoming a reasonable go-to tool in a time-critical contamination disaster. However, as part of the "second response" team, quantum chemistry results were extremely useful for understanding the environmental transport properties of *trans* and *cis* 4-MCHM. We previously reported our "best" results for the dipole moments and relative conformational energies of *trans*-4-MCHM and *cis*-4-MCHM at the MP2(all electron) aug-cc-pwCVDZ level of theory, which validated observed trends in aqueous solubility and carbon loading.[13] In that work, it was found that the *trans* isomer was less soluble in water and was better sorbed to activated carbon than the *cis* isomer, which was also consistent with the *cis* isomer having a lower octanol-water partitioning coefficient (K_{ow}). These trends were supported by the *cis* isomer having a higher computed dipole moment than the *trans* isomer, as general physicochemical trends would predict: species with larger electric dipole moments are generally more soluble in polar media. In this contribution, we will further explore the dipole moment differences and identity of the lowest energy conformers between *trans*-4-MCHM and *cis*-4-MCHM and show that the data have a subtle and unexpected methodological dependence. These results will lead to further calibration of the properties of cyclohexane derivatives and will eventually lead to refinement of macro-environmental fate and transport models. Most importantly, we explore herein the need to consider three-dimensional conformational averaging when attempting to predict physical properties.

II. Methods

Stochastic conformational search and energy cutoff criteria

Initial conformations of *trans*- and *cis*-4-MCHM were obtained via the stochastic search program in the *MOE 2013.08* software package.[16] The stochastic search algorithm in MOE, based on the Random Incremental Pulse Search,[17] generates structures by perturbing rotatable bonds of the initial system then performing a geometry relaxation. By default, the energy minimization technique in MOE utilizes the MMFF94x force field and incorporates a distance-dependent solvation effect with a dielectric constant of 80 for the "exterior" or solvent-exposed parts of the system. Conformers with relative energies higher than 7.0 kcal mol^{-1} from the lowest energy conformer are discarded.

To our knowledge, there has been very little calibration of stochastic conformational searches in the context of thermochemistry and spectroscopic properties. A previous calibration of the MOE stochastic conformational search in the context of small molecule drug discovery has provided some general guidelines for improving the reproduction of bioactive conformers.[18] These improvements included incorporating the generalized Born solvation model and increasing the "energy window" from 7 kcal mol^{-1} to 15 kcal mol^{-1}, where a geometry-optimized conformer is thrown out if its MMFF94x energy is larger than the energy cutoff. For the test case reported here, we observed that both the Born and default implicit solvation models in MOE found the identical set of conformers with the same degeneracies. Relative energies of conformers with the two solvation models were also qualitatively similar. Thus, only results with the default distance-dependent solvation model are reported. It is also understood that bioactive conformations of potential therapeutics can have significantly different structure and can be much higher in energy than those that are found in a room temperature solution phase ensemble, hence the necessity to increase the energy window for greater accuracy in Ref. 18. In their continuing computational study of enzyme mechanisms related to glycolysis, Rovira and coauthors have provided many elegant examples of sugar inhibitors where boat and twist-boat conformations play an important role.[19-22] For the case of MCHM, the higher-energy twist-boat conformers will contribute negligible weight to the Boltzmann distribution. For example, at our recommended level of theory (*vida infra),* all twist-boat conformers have a relative free energy of 5.9 kcal mol^{-1} or greater compared to the lowest energy conformer. The dipole moments of the twist-boat conformers will contribute less than 0.0007% to the total Boltzmann weighting. However, it is likely (but not confirmed for a variety of molecules) that the default energy window of 7 kcal mol^{-1} for saved unique conformers in the MOE stochastic search is reasonable; therefore, we employ this cutoff in our methods.

Considerations for conformational searching and conformational acceptance criteria

While knowledge of all conformers, even those at poorly-experimentally-accessible high energies, are useful for comprehensive understanding of the system, it is more crucial that in our utilization of a stochastic conformational search algorithm that *all* low-energy *cis* and *trans* conformers are located in a single search. In order to reliably identify all unique conformers for 4-MCHM, the other contaminants identified in the Freedom Industries chemical spill, as well as

a larger calibration set of cyclohexane and benzene derivatives with experimentally known dipole moments, two non-default options needed to be used. The first non-default option is to allow hydrogen atoms to be included in the RMSD calculation for detecting duplicates within the search. For example, the two conformers of *trans*-4-MCHM shown in Figure 1 have a nearly degenerate energy, but the difference in their dipole moment is 0.23 Debye. The only structural difference in these two conformers is the orientation of their hydroxyl protons. With a RMSD of only 0.023 Å for the 9 heavy atoms (versus a RMSD of 1.091 Å for all atoms), one of these two conformers would be discarded with the default settings of the MOE stochastic search. The impact of missing or neglecting any low-energy 4-MCHM conformers is surprisingly profound, and will be discussed below. When H-atom positions are factored into the RMSD of the stochastic search, the number of unique 4-MCHM conformers located went from 14 to 31. The second non-default option is to turn off the "rejection limit", where the entire stochastic search terminates if (by default) 100 perturbations of the guess structure fail to find a unique conformer. Surprisingly, this search expands the number of unique conformers to 42 and captures two low-lying *cis* conformers that were missing from our previous examination[13] of the 4-MCHM dipole moment. From these 42 conformers, many were found to be doubly-degenerate due to symmetry of the various methyl and hydroxyl rotamers. As previously mentioned, in the computation of ensemble-averaged dipole moments, all *trans* and *cis* twist-boat conformations were discarded. Overall, nine unique *trans* conformers and nine unique *cis* conformers were obtained.

Lastly, it is not given that all conformers will be located in a particular stochastic search. However, we are confident that in this case study the complete set of low-energy unique conformers has been located due to multiple runs of the stochastic search when varying and testing the conformational search options and carefully inspecting the results. In a situation where a time-critical response is necessary, there is no guarantee that our recommendations will provide a definitively robust stochastic conformational search. Also, it is possible that the complete potential energy surface of 4-MCHM at the MMFF94x level of theory possesses local minima (unique conformers) that do not exist at higher levels of theory. In such a case, geometry optimizations at the higher levels of theory will collapse to non-unique conformers. Inclusion of non-unique conformers could cause double-counting of degeneracies in the calculation of the Boltzmann weighting if unnoticed. Similarly, if new local minima exist on the complete potential energy surface of the more sophisticated levels of theory but do not exist at the MMFF94x level

of theory, then these conformers may be entirely missed from the analysis. These issues are all important with regards to technically rigorous exploration and inclusion of the entire molecular conformational space; however, within the context of a mission-critical determination of physicochemical property estimates for disaster response applications, missing or double-counting a few conformations within an otherwise robust ensemble is likely to have minimal overall impact. It should be noted that in all reported computations, we have performed geometry optimizations of higher levels of theory from the *same* starting geometries – the optimized MMFF94x geometries of each unique conformer. While we believe this approach is sufficient to greatly improve upon extant estimation methods for time-critical situations, clearly there is much interesting further calibration to be performed.

Computational methodology for energies and dipole moments

All computations were performed using the Gaussian09 software package.[23] The unique conformations generated from MOE via MMFF94x and the distance-dependent dielectric solvation model were further optimized in both the gas phase and via implicit solvation with standard aqueous cavity parameters of the SMD model.[24] Equilibrium geometries and harmonic vibrational frequencies were obtained using semiempirical methods such as AM1, PM3, and PM6, DFT with the hybrid B3LYP functional,[25, 26] and *ab initio* methods such as MP2 and coupled cluster with single and double excitations (CCSD). The small "Pople-style" 6-31G(d') basis set[27,28] was used with B3LYP, MP2, and CCSD. For all MP2 computations, the double-zeta and triple-zeta correlation consistent basis sets were used (cc-pVDZ and cc-pVTZ). Further MP2 computations were then performed with larger correlation consistent basis sets systematically augmented to incorporate diffuse functions (aug-cc-pVDZ, d-aug-cc-pVDZ, aug-cc-pVTZ) and tight core functions (cc-pwCVDZ, aug-cc-pwCVDZ, d-aug-cc-pwCVDZ). The aug-cc-pwCVDZ basis set was also used with B3LYP for comparison. Spherical harmonic d functions were used throughout; i.e., there are five angular basis functions per d function. The Hessian of the energy was computed at all stationary points to confirm that they were a minimum on the potential energy surface. Zero-point energies (ZPE) and thermal enthalpy/free energy corrections were computed at 1 atm and 298.15 K.

To obtain conformational ensemble-averaged dipole moments reported throughout the text, we take the Boltzmann-average over the individual conformer-specific dipole moment

values. For each isomer having m conformers, the fractional abundance of individual conformer F_j is based on the free energy difference between conformer, j, and the most stable conformer for that isomer, ΔG_j , and is expressed as

$$F_j = \frac{\omega_j \exp(-\frac{\Delta G_j}{kT})}{\sum_{i=1}^{m} \omega_i \exp\left(-\frac{\Delta G_i}{kT}\right)}$$

where we have set T = 298.15 K, k is the Boltzmann constant, and ω_i is the degeneracy of conformer i. The final conformationally Boltzmann-averaged dipole moment, μ, is then calculated by taking the sum of the dipole moments of the individual conformers, μ_i, weighted by their fractional abundance.

$$\mu = \sum_{i=1}^{m} F_i \mu_i$$

The importance of conformational averaging in computational chemistry has been shown in some recent case studies.[29-35] Also, some investigations of cyclohexane and cyclohexane derivatives exist that employed high-quality levels of electronic structure theory.[36-39]

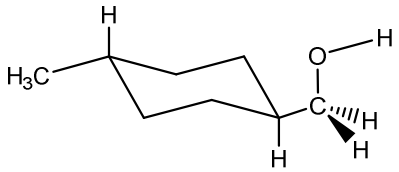
III. Results and Discussion

For the eighteen unique conformers of 4-MCHM (9 *trans* and 9 *cis*), the numbering scheme (which is different than the one presented in Ref. 13 and is now enumerated via the *trans/cis* energy ordering given from the MOE MMFF94x values rather than using the indexing from the conformational search) is given in Table 1 along with relative energies from the MMFF94x force field and a list of the 1,4 substituent orientations on the cyclohexane ring, and the degeneracy (ω_i in the above equation) of each conformer. 3D structures of *trans* and *cis* 4-MCHM conformers and Cartesian coordinates at all computed levels of theory are given in the Supporting Material.

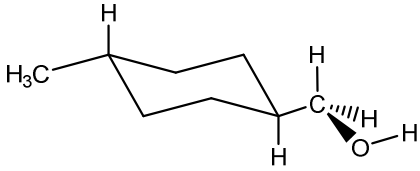
Identification of global minima conformers requires explicit consideration of hydroxyl rotamers

Trans isomer. One might easily envision the lowest energy *trans* conformer of 4-MCHM to have C_s point group symmetry with the oxygen atom in the plane of C1, C4, and the 4-methyl

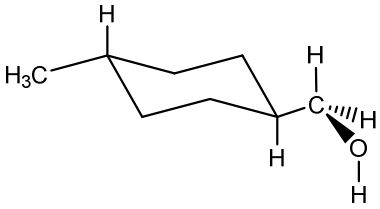
substituent, shown in Scheme 1 (labeled as **t3**). On the contrary, this is not the lowest energy conformer when using any methodology tested in this investigation. In the gas phase, the oxygen has increased stability when it is rotated out of the plane of the C1, C4, and the 4-methyl substituent. At the AM1, PM3, and PM6 levels of theory, the hydroxyl atom is pointed "inward" towards the ring (**t4**). With B3LYP, MP2 and CCSD and the 6-31G(d') basis sets, the lowest energy conformer found has its hydroxyl hydrogen oriented "downward" in the same direction as the axial C1 hydrogen (**t2**). With more sophisticated levels of theory, the lowest energy *trans* conformer has its hydroxyl hydrogen pointed "outwards" (**t1**).



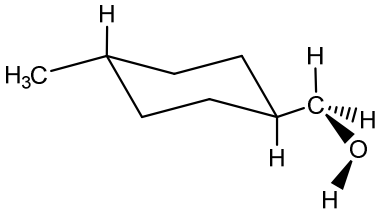
t3 - Symmetric form
Not the lowest-energy *trans* conformer



t1 - lowest energy conformer w/ gas phase MP2 + large basis sets



t2 - lowest energy conformer w/ gas phase *ab initio* and small basis sets



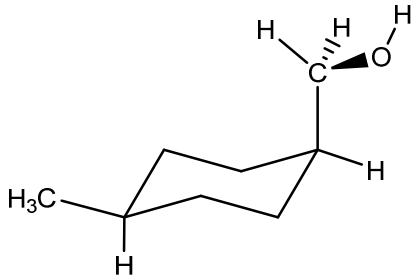
t4 - lowest energy conformer w/ gas phase semiempirical methods and all high-level methods when SMD is included

Scheme 1.

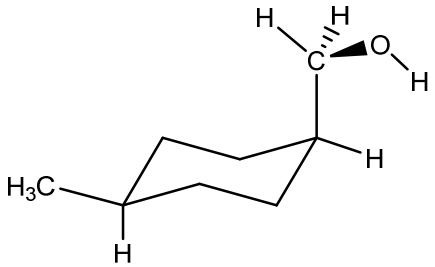
The discrepancies between the identities of the *trans* global minimum rotamer are partially overcome by incorporation of solvation effects. When implicit solvation is incorporated into the computational model, there is more agreement between methodologies for the lowest energy *trans*-4-MCHM conformer. Except for the AM1 and PM3 methods, where **t2** is the most stable conformer, all other levels of theory identify the **t4** conformer as the lowest energy structure.

Cis isomer. The disagreement between methods is similar for the *cis* isomer, but solvation does not seem to help settle the disagreement as much. In the gas phase, **c10** is the lowest energy conformer with the AM1 and PM6 methods, while **c4** is the lowest energy conformer with PM3. DFT and *ab initio* techniques with smaller basis sets identify **c4** as the lowest energy conformer,

while MP2 with basis sets augmented with diffuse functions identify **c1** as the lowest energy *cis* conformer. In solution phase, semi-empirical methods predict **c3**, **c8**, or **c9** to be the most stable *cis* conformer. From DFT, the **c3** conformer has the lowest in energy, while more robust *ab initio* methods show that **c4** is the lowest energy *cis* conformer (Scheme 2). Overall, the identity of the lowest energy *trans* and *cis* conformers for each level of theory is given in Table 2, as well as the free energy difference between the *trans* and *cis* conformers. These results provide evidence that conformations involving hydroxyl rotamers are especially challenging for predictions of proper energy ordering and should be carefully considered.



c1 - lowest energy *cis* conformer in gas phase w/ high level methods



c4 - lowest energy *cis* conformer in solution phase w/ high level methods

Scheme 2.

Determination of the optimal level of theory

In the following sections, we examine the influence of computational method, basis sets, and implicit solvation on the resulting predictions for thermal-averaged physicochemical properties. The results of these investigations are summarized within Figures 2-5, and S1. Relative free energies of *trans* and *cis* 4-MCHM conformers with respect to level of theory are presented respectively in Figures 2 and 3 for gas phase computations and Figures 4 and 5 for solution phase computations. Figure S1 is equivalent to Figure 2, but without MM and semi-empirical levels of theory to show the influence of basis set in finer detail. As shown by the gas phase relative free energies in Figures 2 and 3, there are a larger density of low-lying states in the ensembles of *cis*-4-MCHM conformers. This is also borne out in Figures 4 and 5 which depict the SMD relative free energies of the *trans*-4-MCHM and *cis*-4-MCHM conformers, respectively. Various aspects of the computational method influence are described in the sections below.

Semi-empirical and MM methods

Gas-phase. From Figure 2, it is immediately evident that the semi-empirical methods (AM1, PM3, and PM6) perform unreliably in both relative free energy and free energy ordering of the four lowest-lying *trans*-4-MCHM conformers. The same is true for the *cis* isomer, with the semi-empirical methods doing a poor job of reproducing the more accurate *ab initio* conformational orderings. Surprisingly, the MMFF94x force field performs admirably well for both isomers when compared to large basis set MP2 computations.

Solvated phase. Even though the MOE energy minimizer incorporates implicit solvation, it is interesting that the most stable conformer predicted with MMFF94x (**t1**) is the lowest free energy conformer in the gas phase, but not in solution phase. Rather, most of the *ab initio* methods used in this study (*vida infra*) that incorporate diffuse basis functions show **t4** to be the most stable conformer. The PM6 SMD level of theory performs admirably well for the *trans*-4-MCHM conformers, but provides quite unreliable energy differences and conformer orderings for solution phase *cis*-4-MCHM. For the *cis*-4-MCHM isomer, the MMFF94x force field and the other semi-empirical methods also show complete disagreement with large basis set MP2 results.

Influence of basis set size

Trans isomer. With small basis sets, B3LYP, MP2, and even CCSD computations show qualitatively different relative free energies compared to the converging MP2 computations with larger basis sets. The B3LYP results may initially seem to have a non-systematic relationship to basis set size. At the B3LYP aug-cc-pwCVDZ level of theory, the conformers have the same free energy ordering, and only the free energy difference between **t3** and **t1** ($\Delta\Delta G$ = 0.25 kcal mol^{-1}) is more than 0.05 kcal mol^{-1} from the MP2 aug-cc-pwCVDZ free energies. Shifts in relative free energies are not substantially altered by upgrading MP2 computations from the aug-cc-pVDZ basis sets to the aug-cc-pVTZ basis sets. The relative free energies for **t2**, **t4**, and **t3** change by -0.03, +0.01, and +0.06 kcal mol^{-1}, respectively compared to **t1**.

Cis isomer. Unfortunately, for the *cis*-4-MCHM conformers, B3LYP aug-cc-pwCVDZ relative free energies show no resemblance to the MP2 aug-cc-pwCVDZ relative free energies. There is again no qualitative difference between the MP2 aug-cc-pVDZ and MP2 aug-cc-pVTZ results, as conformer ordering differences in the tight cluster of **c5**, **c9**, **c3**, and **c8** will not have a

severe impact on the overall *cis* dipole moment. For all *trans* and *cis* conformations of 4-MCHM, the range of relative free energy shifts from MP2 aug-cc-pVDZ to MP2 aug-cc-pVTZ averages -0.03 kcal mol^{-1}. Introduction of solvation seems to aid the DFT method somewhat, as the B3LYP aug-cc-pwCVDZ level of theory gives a qualitatively correct state ordering, but conformers are destabilized compared to the lowest energy *trans*-4-MCHM conformers, by more than 0.2 kcal mol^{-1} in a few cases. In light of the poor performance of B3LYP for the *cis* isomer, MP2 with a modest basis set affords the best speed and accuracy.

Inclusion of diffuse functions

The improved reliability of B3LYP aug-cc-pwCVDZ compared to the B3LYP cc-pVTZ computations is due to the importance of including diffuse basis functions in dipole moment calculations. However, there is a noticeable difference in relative free energies and orderings when a second set of diffuse basis functions are added to the computation with the valence MP2 d-aug-cc-pVDZ and d-aug-cc-pwCVDZ estimates. The **t3** conformer is stabilized with the additional d-aug functions and becomes the second or third lowest-energy *trans*-conformer. Despite the change in free energy ordering, the additional d-aug basis functions will not have a quantitatively important effect on the conformationally-averaged dipole moments (*vida infra*).

Relative gas phase free energies of *trans*-4-MCHM conformers only really begin to converge with the inclusion of diffuse basis functions. Without augmented diffuse functions, MP2 and CCSD provide fluctuating state orderings and relative free energies. Though the free energy of **t2** is 0.05 kcal mol^{-1} lower than **t1** at the MP2 aug-cc-pVDZ level of theory, they become isoenergetic when d-aug basis functions are added. The relative free energies of conformers may be more susceptible to 1-electron basis set effects when implicit solvation is included. However, MP2 aug-cc-pVDZ still shows the best balance in reliability versus needed computational resources and time. When diffuse basis functions are added to the MP2 computations, the **c4** conformer is the lowest in free energy. Except for the high-energy **c5** conformer, the MP2 aug-cc-pVDZ relative free energies are typically within 0.07 kcal mol^{-1} of the MP2 d-aug-cc-pVDZ values. Diffuse basis functions are necessary to predict **c1** as the most stable conformer. The addition of d-aug basis functions stabilizes the relative free energies of **c5** by 0.11 kcal mol^{-1}, while slightly destabilizing relative free energies of **c3**, **c8**, and **c6**.

Influence of core correlation

There is no significant change between the valence MP2 aug-cc-pVDZ and aug-cc-pwCVDZ results. This is seen to be generally true over both isomers and in the gas and solvated phases. The effect of core correlation remains insignificant for these systems when the larger, augmented basis sets are employed.

Influence of conformational averaging on the resulting dipole moment

Gas-phase. Boltzmann-averaged gas phase dipole moments are reported in Figure 6. In the gas phase, the *trans* and *cis* isomers are computed to have a nearly identical dipole moment at almost all reliable levels of theory. Using MP2 aug-cc-pVDZ, the *trans* dipole moment is computed to be 1.6226 D, with the *cis* dipole moment only 0.0014 D lower (1.6202 D). Compare these equivalent dipole moments with the dipole moment of the lowest free energy conformers for *trans*-4-MCHM and *cis*-4-MCHM in Figure 6. At the MP2 aug-cc-pVDZ level of theory, the **t1** dipole moment is significantly lower (1.4737 D), and the **c1** dipole moment is lower still (1.3479 D). Using the dipole moment of the lowest energy conformers to make statements on relative isomeric solubility would be extremely dangerous when making the assumption that a C_s symmetric form of each isomer (**t3** or **t9**; **c5** and **c7**) would be the most stable conformer. One would have an equal probability of getting the right answer for the wrong reason, or the wrong answer for the wrong reason. The **t3** and **t9** dipole moments agree to within 0.0001 D, and are within 0.023 D of the Boltzmann-averaged value. However, the **c5** and **c7** dipole moments (1.4492 and 1.7222 D, respectively) bracket the Boltzmann-averaged *cis*-4-MCHM value of 1.6202 D by an uncomfortably wide margin.

Solvation phase. In terms of solubility trends, the solution phase Boltzmann-averaged dipole moments are better behaved (Figure 7). There tends to be a larger gap between the *trans* and *cis* dipole moments, as well as a less pronounced methodological dependence on the relative values. At the MP2 SMD aug-cc-pVDZ level of theory, the conformationally averaged *trans*-4-MCHM dipole moment is 2.4475 D, and the conformationally averaged *cis*-4-MCHM dipole moment is 2.4834 D. This absolute difference of 0.0358 D is a more comfortable separation than is observed in the gas phase computations, considering the overall level of accuracy of our electron correlation treatment.

Selection of the optimal method

Overall, the discussion in the previous sections drives the ultimate selection of the MP2 aug-cc-pVDZ method at the optimal method that balances accuracy and computational resources. Computation of MP2 aug-cc-pVDZ equilibrium geometry, harmonic vibrational frequencies, and dipole moment for a conformer of 4-MCHM required about 10 hours of wall time per conformer versus more than 500 hours of wall time per conformer at the MP2 aug-cc-pVTZ level of theory. This clearly makes application of larger basis set MP2 computations useless for rapidly determining the properties of emerging organic contaminants. Luckily, we observe little quantitative difference in results between the double zeta and triplet zeta basis sets. As mentioned, augmentation with diffuse functions greatly improves the results.

Other observations can be made that may hint at more improved *ab initio* methods providing improved accuracy, but these all would be more costly; for instance, it is observed that the CCSD SMD 6-31G(d') computations slightly increase the relative free-energy and dipole moment gaps between the isomers compared to MP2 theory. However, calibration of more sophisticated methodology beyond what has been tested here is currently intractable, and goes against the overarching goal of this investigation – to find an efficient *ab initio* methodology that can provide reliable dipole moments for organic contaminants.

At the optimal MP2 aug-cc-pVDZ level of theory with SMD solvation, we find the lowest-energy *trans*-4-MCHM conformer (**t4**) to be the overall global minimum, with the lowest-energy *cis*-4-MCHM conformer (**c4**) having a free-energy difference of 1.69 kcal mol^{-1}. Boltzmann-averaged solution phase dipole moments indicate the averaged *cis*-4-MCHM dipole moment (2.4834 D) to be slightly larger than the averaged *trans*-4-MCHM dipole moment (2.4475 D), which is consistent with experimental observations. While predictions at this level of theory in the gas phase indicate the same free-energy difference between the isomers, different conformers are identified as the minima, relative energetic ordering among the conformers is different, and the dipole moment trends are reversed. This reiterates the need for inclusion of solvation effects when using properties to predict aqueous phenomena.

Consideration of composite method alternatives

As an alternative to selecting a single "best" method, simple composite methods may allow a way to achieve the accuracy of more sophisticated methods at a decreased computational cost.[40] As such, we have explored the conformationally averaged dipole moments of 4-MCHM were explored with simple composite methods. First, energies and dipole moments were computed using MP2 aug-cc-pwCVDZ single points (with all electrons correlated) at the MP2 cc-pwCVDZ optimized geometries (with only valence electrons correlated), denoted in "Pople notation" as MP2 (all electron) aug-cc-pwCVDZ//MP2 (valence) cc-pwCVDZ and in the Tables/Figures as **CM1**. These results were compared to MP2 (all electron) aug-cc-pwCVDZ//MP2 (valence) aug-cc-pVDZ (**CM2**). Lastly, an electronic energy and dipole moment correction for each conformer from CCSD 6-31G(d') versus MP2 6-31G(d') was computed as an additive correction [Δ(CCSD)] to the MP2 aug-cc-pVDZ results,

$$\Delta E_e(CCSD) = E_e[\text{CCSD 6-31G(d')}] - E_e[\text{MP2 6-31G(d')}]$$

$$\Delta \mu(CCSD) = \mu[\text{CCSD 6-31G(d')}] - \mu[\text{MP2 6-31G(d')}] \,.$$

Tables S1 and 3 respectively show the relative energies and computed 4-MCHM dipole moments, along with Boltzmann-averaged *trans* and *cis* dipole moments using these composite methods. Quantitatively, there is very little change in dipole moments and relative free energies between MP2 aug-cc-pVDZ and MP2(all electron) aug-cc-pwCVDZ. Between the two methodologies with implicit solvation, the standard deviation in the free energy differences is only 0.014 kcal mol^{-1} and the standard deviation in dipole moments is only 0.0008 D. Conformer free energy differences between MP2(all electron) aug-cc-pwCVDZ and **CM2** are less than 0.001 kcal mol^{-1} in every case. However, dipole moments are overestimated slightly, with a mean signed deviation between the two methods of +0.0019 D. In terms of computing efficient and accurate dipole moments and relative energies, MP2 aug-cc-pVDZ seems to afford the best balance. Unfortunately, results obtained using **CM1** are in far worse agreement with both MP2 aug-cc-pVDZ and MP2 (all electron) aug-cc-pwCVDZ levels of theory. In the gas phase with **CM1** there is a systematic overstabilization of conformers **t6**, **t8**, **c8**, and **c9**. On the contrary in the solution phase, all relative energies compared to the lowest energy conformer (**t4**) are higher. The Boltzmann-averaged solution phase **CM1** dipole moments also disagree with MP2 aug-cc-pVDZ by predicting a larger *trans* dipole moment. Inclusion of diffuse functions in the geometry

optimization process is crucial to obtain accurate dipole moments. MP2 aug-cc-pVDZ computations with a small basis set CCSD additive correction provide consistent results to both MP2 aug-cc-pVDZ and MP2(all electronic) aug-cc-pwCVDZ methods. While the Δ(CCSD) additive correction likely provides improved results for relative energies and dipole moments, in the case of 4-MCHM there is no new insight gained by including Δ(CCSD). The prohibitive scaling of CCSD will render larger molecules intractable for most practicing computational chemists and is not recommended when computing dipole moments of conformationally flexible molecules containing 20+ atoms. In the end, the MP2 aug-cc-pVDZ level of theory seems to remain the superior method at balancing accuracy and time cost.

IV. Implications and Concluding Remarks

Molecular properties are frequently used in computational models of environmental fate and transport. However, often the values for molecular parameters are obtained via group additivity or semi-empirical methods. Rarely do these models account for 3D structure or explicit conformational averaging. Specifically in the case of the Elk River chemical spill, the dipole moments and molecular volumes of (4-methylcyclohexyl)methanol (4-MCHM) and other contaminants were used in engineering models to predict aqueous solubility.[13] The *trans* and *cis* isomers of the 4-MCHM mixture in that spill incident were not in thermal equilibrium at ambient temperatures along the pathways of environmental fate. Therefore, the conformational averaging of properties for each isomer must be determined separately.

The difference in computed *trans-/cis*-4-MCHM dipole moments is quite small, but the isomeric trend in GC retention time (*cis* > *trans*), carbon sorption behavior (*cis* < *trans*), and aqueous solubility (*cis* > *trans*) is experimentally observable.[9, 13, 14] In a previous study, our laboratories used the computed solution-phase dipole moments (*cis* > *trans*) to validate the experimental trend in the aqueous solubility of 4-MCHM via first principles electronic structure theory.[13] While factors beyond simple polar arguments contribute to the solubility, a species' dipole moment proves again to be a useful predictor of its aqueous behavior.[41] Stochastic conformational searching seems to offer a route to generate initial structures for a thorough evaluation of 3D molecular flexibility. However, users must beware of default software options typically employed in the context of small molecule drug discovery versus high-accuracy

computation of molecular properties. The Boltzmann-averaged dipole moment of *trans*-4-MCHM is only 0.039 D smaller than that of *cis*-4-MCHM at the MP2(all electron) aug-cc-pwCVDZ level of theory, when incorporating implicit solvation effects with SMD. In the current study, the conformationally averaged dipole moments of 4-MCHM have been further calibrated using a variety of electronic structure theory, including *ab initio* computations with larger basis sets than aug-cc-pwCVDZ and more robust treatments of electron correlation. Unfortunately, semi-empirical methods and DFT in isomeric dipole moments of 4-MCHM. Interestingly, it has been found that more efficient methodologies offer comparable results to the MP2(all electron) aug-cc-pwCVDZ level of theory, with valence MP2 aug-cc-pVDZ computations providing an appropriate balance between rigor and computational tractability. Leveraging supercomputer capabilities, reliable dipole moments for molecules with up to 30 atoms and reasonable conformational flexibility can be obtained on the time scale of a few hours to a few days.

This case study also highlights the need for careful and critical evaluation of computational predictions, even in the face of a developing disaster. Gas-phase estimates of the MCHM isomer dipole moments would predict reversed trends for the two isomers' environmental behavior (i.e. solubility, partitioning, etc.) Practitioners are cautioned against using gas phase property predictions when the computations will be used to inform solution phase transport properties. The subtleties of obtaining accurate molecular properties in a rapid, time-sensitive manner require further development before a true "turn-key" solution can be provided for first responders. In the meantime, these studies bolster the sentiment for increasing scientific support within, or parallel to, the incident command structure in responding to chemical spills. As quantum chemistry packages are developed to fully utilize the rapid growth of parallel computing and GPU technologies, the type of computational predictions described herein will become even more attractive and useful to provide rigorous first-principles predictions of molecular properties in time-sensitive situations.

ACKNOWLEDGMENTS

This project was funded by the National Science Foundation, Division of Chemical, Biological, and Environmental Transport, Award #1435289. The University of Memphis High Performance Computing Facility and Computational Research on Materials Institute (CROMIUM) are gratefully acknowledged for computing support.

SUPPORTING INFORMATION

Electronic Supporting Information (ESI). Insets of Figures 2 and 3, 3D structures for all conformers of 4-MCHM, relative free energies and Cartesian coordinates for all computed structures, are included as ESI. See DOI: 10.1039/x0xx00000x

AUTHOR INFORMATION

Corresponding Author

*Email: ndyonker@memphis.edu and w.alexander@memphis.edu

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

National Science Foundation, Division of Chemical, Biological, and Environmental Transport, CBET Awards #1435289

References

1. http://chemm.nlm.nih.gov/onsite.htm
2. EPA, TSCA Inventory. In Environmental Protection Agency, www.data.gov, accessed Sept. 2015.
3. Rosen, J.; Whelton, A. J.; McGuire, M. J.; Clancy, J. L.; Bartrand, T.; Eaton, A.; Patterson, J.; Dourson, M.; Nance, P.; Adams, C., WV TAP Final Report. West Virginia Testing

Assessment Project: Charleston, WV USA. July, Accessible at http://www.wvtapprogram.com. 2014.
4. Cooper, W. J., Responding to Crisis: The West Virginia Chemical Spill. *Environ. Sci. Technol.* **2014,** *48* (6), 3095-3095.
5. Manuel, J., Crisis and Emergency Risk Communication: Lessons from the Elk River Spill. *Environ Health Persp* **2014,** *122* (8), A214-A219.
6. http://yosemite.epa.gov/sab/SABPRODUCT.NSF/CCF982BA9F9CFCFA8525735200739805/$File/sab-07-011.pdf
7. R. D. Christie, A. E. G., R. J. Fortin Process for coal flotation using 4-methyl cyclohexane methanol frothers. April 10, 1994, 1990.
8. Perkin, W. H.; Pope, W. J., CIX.-Experiments on the synthesis of 1-methylcyclo-hexylidene-4-acetic acid, [graphic omitted]. Part I. *Journal of the Chemical Society, Transactions* **1908,** *93* (0), 1075-1085.
9. Foreman, W. T.; Rose, D. L.; Chambers, D. B.; Crain, A. S.; Murtagh, L. K.; Thakellapalli, H.; Wang, K. K., Determination of (4-methylcyclohexyl)methanol isomers by heated purge-and-trap GC/MS in water samples from the 2014 Elk River, West Virginia, chemical spill. *Chemosphere* **2015,** *131*, 217-224.
10. McGuire, M. J.; Rosen, J.; Whelton, A. J.; Suffet, I. H., An unwanted licorice odor in a West Virginia water supply. *J Am Water Works Ass* **2014,** *106* (6), 72-82.
11. Whelton, A. J.; McMillan, L.; Connell, M.; Kelley, K. M.; Gill, J. P.; White, K. D.; Gupta, R.; Dey, R.; Novy, C., Residential Tap Water Contamination Following the Freedom Industries Chemical Spill: Perceptions, Water Quality, and Health Impacts. *Environ. Sci. Technol.* **2015,** *49* (2), 813-823.
12. Bahadur, R.; Samuels, W. B., Modeling the Fate and Transport of a Chemical Spill in the Elk River, West Virginia. *J Environ Eng* **2015,** *141* (7).
13. Dietrich, A. M.; Thomas, A.; Zhao, Y.; Smiley, E.; Shanaiah, N.; Ahart, M.; Charbonnet, K. A.; DeYonker, N. J.; Alexander, W. A.; Gallagher, D. L., Partitioning, Aqueous Solubility, and Dipole Moment Data for cis- and trans-(4-Methylcyclohexyl)methanol, Principal Contaminants of the West Virginia Chemical Spill. *Environmental Science & Technology Letters* **2015,** *2* (4), 123-127.
14. Gallagher, D. L.; Phetxumphou, K.; Smiley, E.; Dietrich, A. M., Tale of Two Isomers: Complexities of Human Odor Perception for cis-and trans-4-Methylcyclohexane Methanol from the Chemical Spill in West Virginia. *Environ. Sci. Technol.* **2015,** *49* (3), 1319-1327.
15. Lan, J. Q.; Hu, M.; Gao, C.; Alshawabkeh, A.; Gu, A. Z., Toxicity Assessment of 4-Methyl-1-cyclohexanemethanol and Its Metabolites in Response to a Recent Chemical Spill in West Virginia, USA. *Environ. Sci. Technol.* **2015,** *49* (10), 6284-6293.
16. Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2015.
17. Ferguson, D. M.; Raber, D. J., A New Approach to Probing Conformational Space with Molecular Mechanics - Random Incremental Pulse Search. *J Am Chem Soc* **1989,** *111* (12), 4371-4378.
18. Chen, I. J.; Foloppe, N., Conformational sampling of druglike molecules with MOE and catalyst: Implications for pharmacophore modeling and virtual screening. *J Chem Inf Model* **2008,** *48* (9), 1773-1791.

19. Biarnes, X.; Ardevol, A.; Iglesias-Fernandez, J.; Planas, A.; Rovira, C., Catalytic Itinerary in 1,3-1,4-beta-Glucanase Unraveled by QM/MM Metadynamics. Charge Is Not Yet Fully Developed at the Oxocarbenium Ion-like Transition State. *J Am Chem Soc* **2011,** *133* (50), 20301-20309.
20. Iglesias-Fernandez, J.; Raich, L.; Ardevol, A.; Rovira, C., The complete conformational free energy landscape of beta-xylose reveals a two-fold catalytic itinerary for beta-xylanases. *Chemical Science* **2015,** *6* (2), 1167-1177.
21. Thompson, A. J.; Speciale, G.; Iglesias-Fernandez, J.; Hakki, Z.; Belz, T.; Cartmell, A.; Spears, R. J.; Chandler, E.; Temple, M. J.; Stepper, J.; Gilbert, H. J.; Rovira, C.; Williams, S. J.; Davies, G. J., Evidence for a Boat Conformation at the Transition State of GH76 alpha-1,6-Mannanases-Key Enzymes in Bacterial and Fungal Mannoprotein Metabolism. *Angew Chem Int Edit* **2015,** *54* (18), 5378-5382.
22. Williams, R. J.; Iglesias-Fernandez, J.; Stepper, J.; Jackson, A.; Thompson, A. J.; Lowe, E. C.; White, J. M.; Gilbert, H. J.; Rovira, C.; Davies, G. J.; Williams, S. J., Combined Inhibitor Free-Energy Landscape and Structural Analysis Reports on the Mannosidase Conformational Coordinate. *Angew Chem Int Edit* **2014,** *53* (4), 1087-1091.
23. M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., GAUSSIAN 09, Revision D. 01, Gaussian, Inc., Wallingford, CT, 2009.
24. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J Phys Chem B* **2009,** *113* (18), 6378-6396.
25. Becke, A. D., Density-Functional Thermochemistry .3. The Role of Exact Exchange. *J Chem Phys* **1993,** *98* (7), 5648-5652.
26. Lee, C. T.; Yang, W. T.; Parr, R. G., Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Physical Review B* **1988,** *37* (2), 785-789.
27. Harihara.Pc; Pople, J. A., Influence of Polarization Functions on Molecular-Orbital Hydrogenation Energies. *Theor Chim Acta* **1973,** *28* (3), 213-222.
28. The 6-31G(d') basis set has the d polarization functions for C, N, and O taken from the 6-311G(d) basis set, instead of the original arbitrarily assigned value of 0.8 used in the 6-31G(d) basis set. J. B. Foresman and Æ. Frisch, Exploring Chemistry with Electronic Structure Methods, 2nd Ed. Gaussian, Inc, Pittsburgh, PA, p. 110.
29. Csontos, J.; Kalman, M.; Tasi, G., Conformational analysis of melatonin at Hartree-Fock ab initio level. *J Mol Struc-Theochem* **2003,** *640*, 69-77.
30. Csontos, J.; Kalman, P.; Tasi, G.; Kalman, M.; Murphy, R. F.; Lovas, S., The effect of electron correlation on the conformational space of melatonin. *J Comput Chem* **2008,** *29* (9), 1466-1471.
31. Garden, A. L.; Paulot, F.; Crounse, J. D.; Maxwell-Cameron, I. J.; Wennberg, P. O.; Kjaergaard, H. G., Calculation of conformationally weighted dipole moments useful in ion-molecule collision rate estimates. *Chem Phys Lett* **2009,** *474* (1-3), 45-50.
32. Borowski, P., Conformational analysis of the chemical shifts for molecules containing diastereotopic methylene protons. *J Magn Reson* **2012,** *214*, 1-9.
33. Goerigk, L., How Do DFT-DCP, DFT-NL, and DFT-D3 Compare for the Description of London-Dispersion Effects in Conformers and General Thermochemistry? *J Chem Theory Comput* **2014,** *10* (3), 968-980.

34. Martin, J. M. L., What Can We Learn about Dispersion from the Conformer Surface of n-Pentane? *J Phys Chem A* **2013,** *117* (14), 3118-3132.
35. Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L., The Melatonin Conformer Space: Benchmark and Assessment of Wave Function and DFT Methods for a Paradigmatic Biological and Pharmacological Molecule. *J Phys Chem A* **2013,** *117* (10), 2269-2277.
36. Durig, J. R.; Zheng, C.; El Defrawy, A. M.; Ward, R. M.; Gounev, T. K.; Ravindranath, K.; Rao, N. R., On the relative intensities of the Raman active fundamentals, r(0) structural parameters, and pathway of chair-boat interconversion of cyclohexane and cyclohexane-d(12). *J Raman Spectrosc* **2009,** *40* (2), 197-204.
37. Weldon, A. J.; Tschumper, G. S., Intrinsic conformational preferences of and an anomeric-like effect in 1-substituted silacyclohexanes. *Int J Quantum Chem* **2007,** *107* (12), 2261-2265.
38. Weldon, A. J.; Vickrey, T. L.; Tschumper, G. S., Intrinsic conformational preferences of substituted cyclohexanes and tetrahydropyrans evaluated at the CCSD(T) complete basis set limit: Implications for the anomeric effect. *J Phys Chem A* **2005,** *109* (48), 11073-11079.
39. Kleinpeter, E.; Thielemann, J., Syntheses and conformational analyses of mono- and trans-1,4-dialkoxy substituted cyclohexanes - the steric substituent/skeleton interactions. *Tetrahedron* **2007,** *63* (37), 9071-9081.
40. DeYonker, N. J.; Cundari, T. R.; Wilson, A. K., "The correlation consistent Composite Approach (ccCA): Efficient and Pan-periodic Kinetics and Thermodynamics", in: "Progress in Theoretical Chemistry and Physics", J. Maruani, S. Wilson, W. N. Libscomp, ed., Springer, **19**, 197 (2009).
41. At the suggestion of a helpful anonymous reviewer, we determined the solvation energy differences between the *cis*- and *trans*-MCHM isomers to examine if the solvation energy would yield a more substantial isomeric difference to explain the observed experimental trends. The solvation energy, as computed by the difference in the SMD solution phase and gas-phase results, is $\Delta G_{solv} = G_{soln} - G_{gas}$. At the MP2 aug-cc-pVDZ level, the Boltzmann average of this quantity yields $\Delta G_{solv}(cis)$ = -4.548 kcal/mol and $\Delta G_{solv}(trans)$ = -4.533 kcal/mol; this agrees with the dipole moment and experimental solubility trends, but the difference between the isomers, $\Delta\Delta G_{solv}$, is only 0.015 kcal/mol, which is quite small. Interestingly, examination of only the lowest energy conformers of each isomer at the same level of theory, we observe the incorrect trend: ΔG_{solv}(**t4**) = -4.68 kcal/mol, ΔG_{solv}(**c4**) = -4.56 kcal/mol.

Figure 1. 3D structures of *trans*-4-MCHM conformers (**t1** on left, **t4** on the right). At the MP2 aug-cc-pVDZ level of theory with SMD, the two conformers differ in energy by only 0.12 kcal mol^{-1}, but their dipole moments differ by 0.23 Debye.

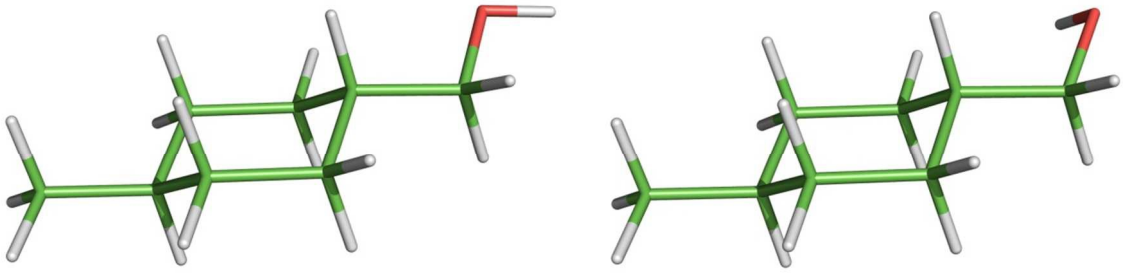


Figure 2. Gas phase free energy differences (in kcal mol^{-1}) for *trans*-4-MCHM conformers. Note that at almost every level of theory (with a few outliers observed with PM3 and PM6) the **t6, t7, t8,** and **t9** conformers have a relative free energy greater than 2.52 kcal mol^{-1} and are not shown. MMFF94x computations include implicit solvation within MOE and are shown for comparison purposes.

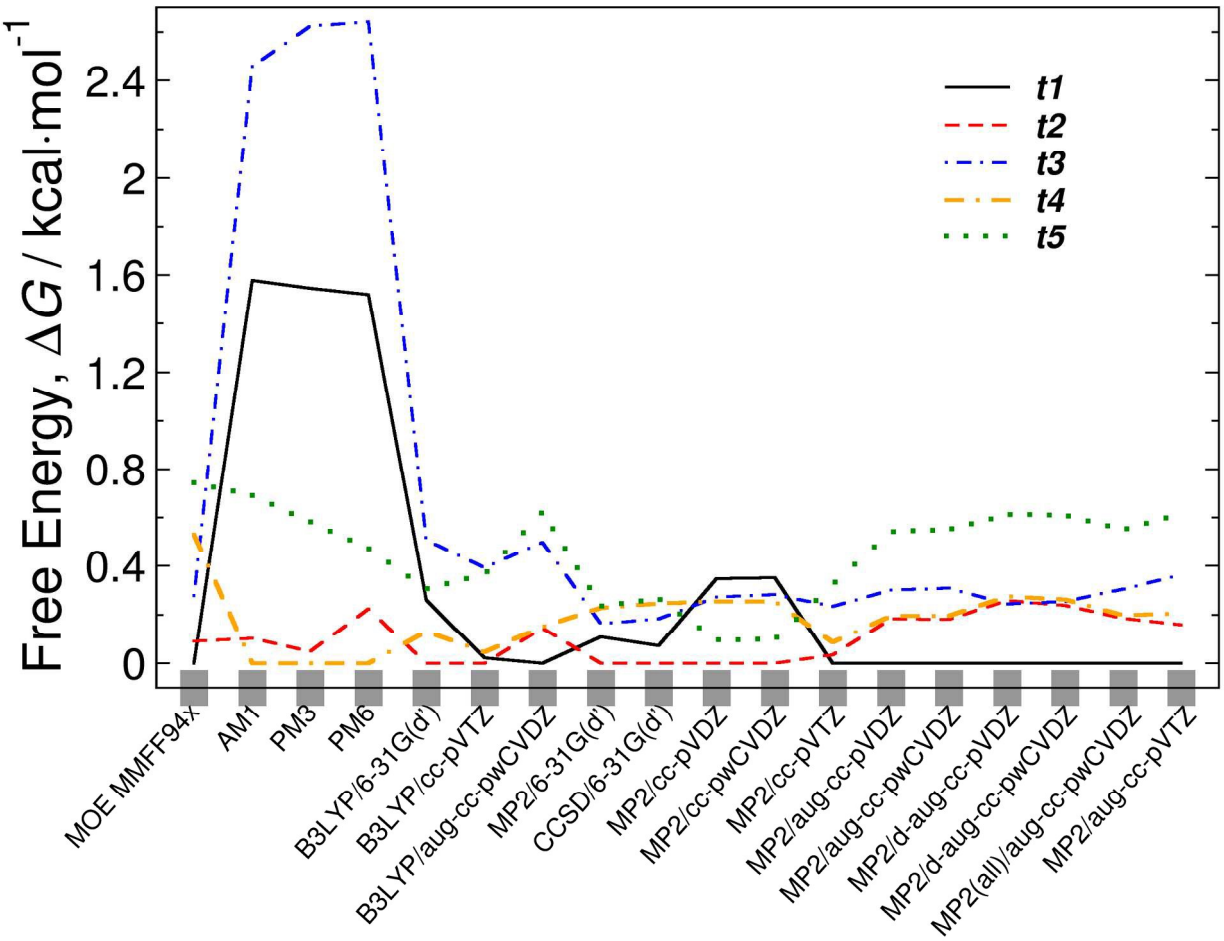


Figure 3. Gas phase free energy differences (in kcal mol^{-1}) for *cis*-4-MCHM conformers. MMFF94x computations include implicit solvation within MOE and are shown for comparison purposes.

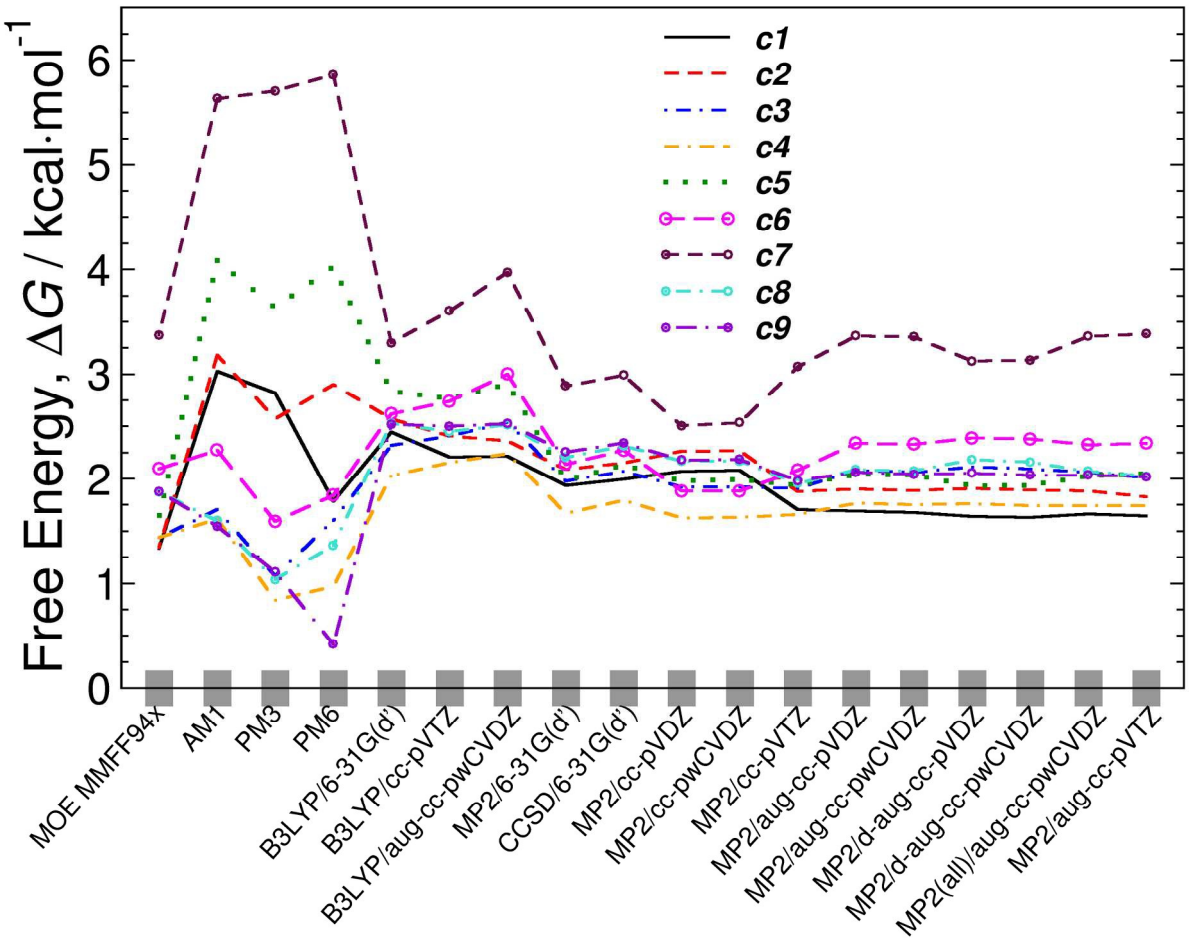


Figure 4. Solution phase (SMD solvation model) free energy differences (in kcal mol^{-1}) for *trans*-4-MCHM conformers. Note that at almost every level of theory (with one outlier observed with PM3) the **t6, t7, t8,** and **t9** conformers have a relative free energy greater than 2.73 kcal mol^{-1} and are not shown. MMFF94x computations include implicit solvation within MOE and are shown for comparison purposes.

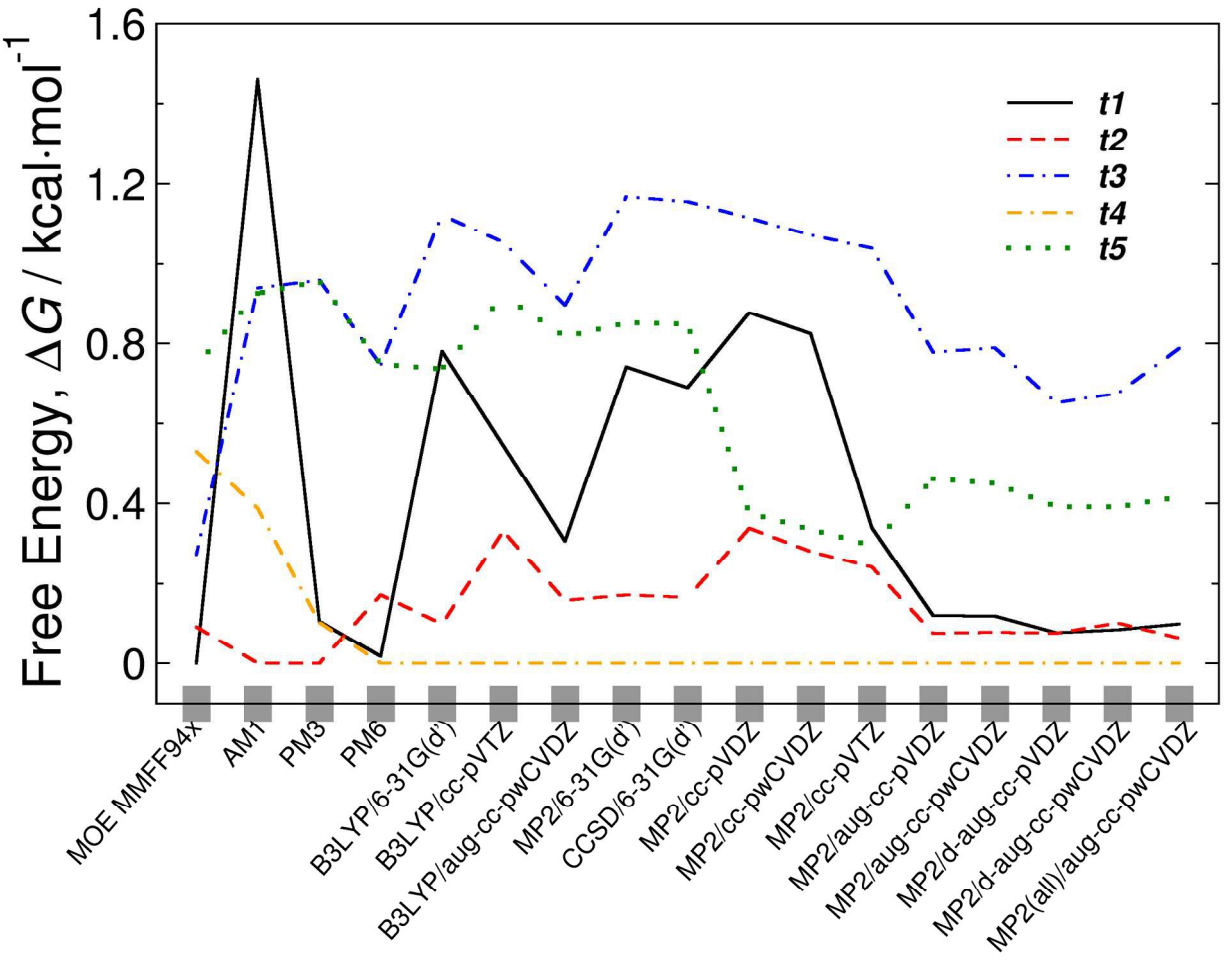


Figure 5. Solution phase (SMD solvation model) free energy differences (in kcal mol^{-1}) for *cis*-4-MCHM conformers. MMFF94x computations include implicit solvation within MOE and are shown for comparison purposes.

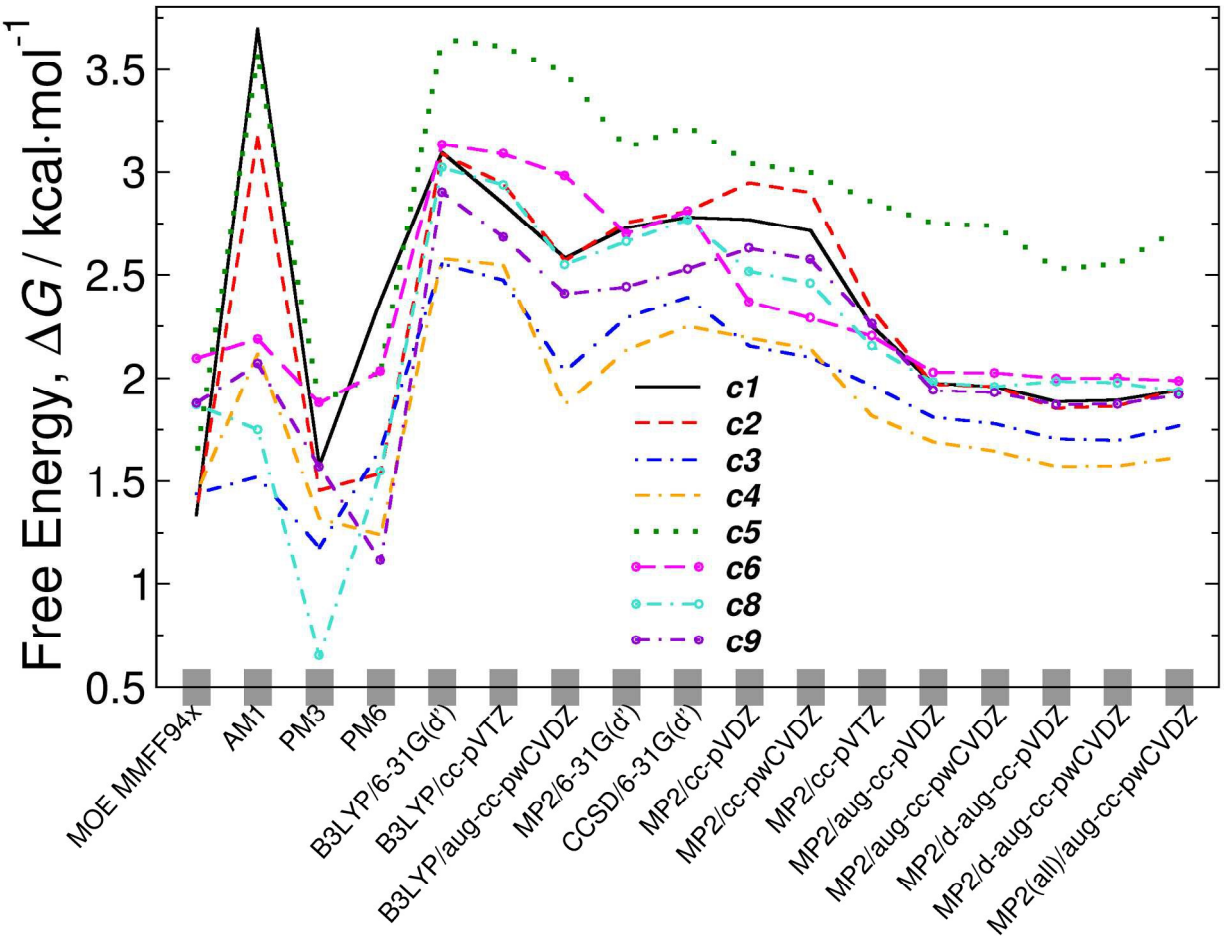


Figure 6. Gas phase dipole moments (in Debye) for *trans*-4-MCHM and *cis*-4-MCHM conformers. Boltzmann averaged dipole moments for the isomers (solid lines) are compared to the dipole moment belonging to the lowest-energy conformer (dashed line). The numbers represent the dipole moment of the individual low-energy conformers (see Table 1, *trans* conformers are given by black numbers for **t*n***, and *cis* conformers are given by red numbers for **c*n*).** Only lowest-energy conformers are shown.

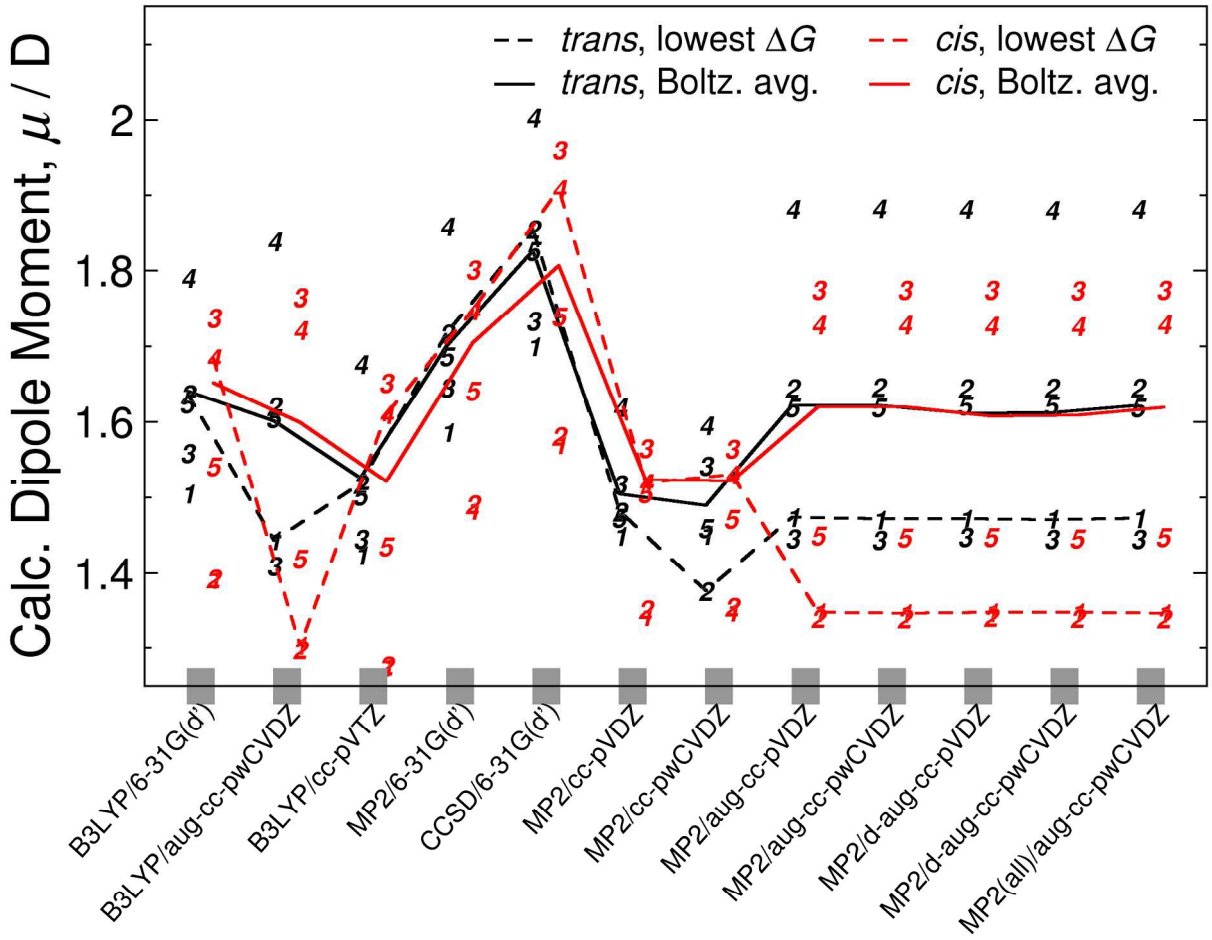


Figure 7. Solution phase (SMD solvation model) dipole moments (in Debye) for *trans*-4-MCHM and *cis*-4-MCHM conformers. Boltzmann averaged dipole moments for the isomers (solid lines) are compared to the dipole moment belonging to the lowest-energy conformer (dashed line). The numbers represent the dipole moment of the individual low-energy conformers (see Table 1, *trans* conformers are given by black numbers for **t*n***, and *cis* conformers are given by red numbers for **c*n*).** Only lowest-energy conformers are shown.

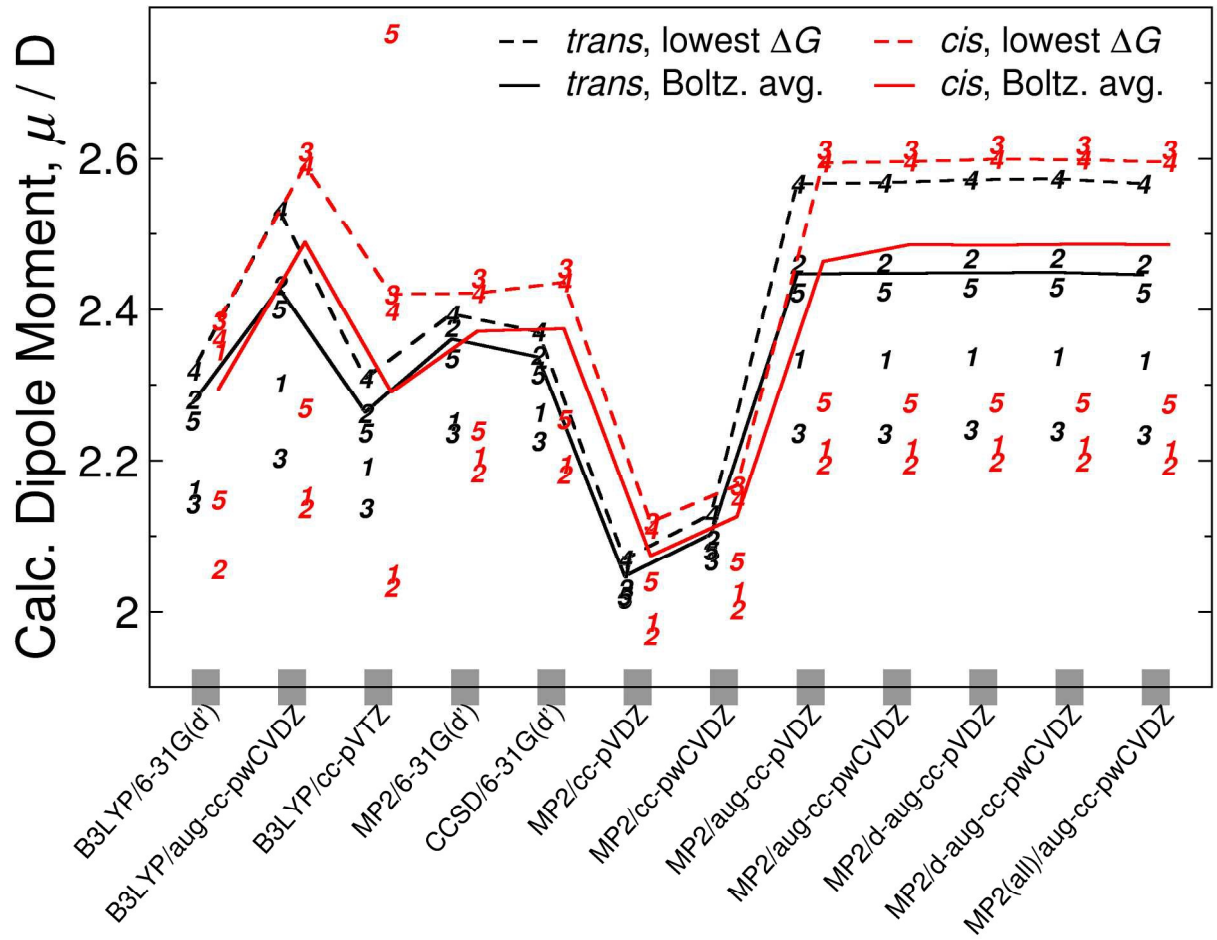


Table 1. Conformer label scheme (see corresponding 3D structures in Figures S3 & S4)

Conformer name	Label from Ref. 13	Isomer	substituent orientation CH_3OH / CH_3	Degeneracy of conformer
t1	1	*trans*	equatorial/equatorial	2
t2	3	*trans*	equatorial/equatorial	2
t3	5	*trans*	equatorial/equatorial	1
t4	6	*trans*	equatorial/equatorial	2
t5	7	*trans*	equatorial/equatorial	2
c1	9	*cis*	axial/equatorial	2
c2	11	*cis*	equatorial/axial	2
c3	13	*cis*	equatorial/axial	2
c4	15	*cis*	axial/equatorial	2
c5	16	*cis*	equatorial/axial	1
c6	17	*cis*	equatorial/axial	2
t6	19	*trans*	axial/axial	2
t7	21	*trans*	axial/axial	2
t8	22	*trans*	axial/axial	2
c7	23	*cis*	axial/equatorial	1
t9	24	*trans*	axial/axial	1
c8	n/a	*cis*	equatorial/axial	2
c9	n/a	*cis*	axial/equatorial	2

Table 2. Isomeric free energy differences of lowest-energy conformer computed at each level of theory.

	gas phase			solution phase		
	lowest trans conformer	lowest cis conformer	ΔG (kcal/mol)	lowest trans conformer	lowest cis conformer	ΔG (kcal/mol)
MMFF94x[a]				**t1**	**c1**	1.33
AM1	**t4**	**c9**	1.60	**t2**	**c3**	1.52
PM3	**t4**	**c4**	0.84	**t2**	**c8**	0.65
PM6	**t4**	**c9**	0.42	**t4**	**c9**	1.11
B3LYP 6-31G(d')	**t2**	**c4**	2.03	**t4**	**c3**	2.55
B3LYP cc-pVTZ	**t2**	**c4**	2.15	**t4**	**c3**	2.48
B3LYP aug-cc-pwCVDZ	**t1**	**c1**	2.21	**t4**	**c4**	1.87
MP2 6-31G(d')	**t2**	**c4**	1.67	**t4**	**c4**	2.14
CCSD 6-31G(d')	**t2**	**c4**	1.79[b]	**t4**	**c4**	2.25[b]
MP2 cc-pVDZ	**t2**	**c4**	1.63	**t4**	**c3**	2.15
MP2 cc-pwCVDZ	**t2**	**c4**	1.63	**t4**	**c3**	2.10
MP2 cc-pVTZ	**t1**	**c4**	1.66	**t4**	**c4**	1.82
MP2 aug-cc-pVDZ	**t1**	**c1**	1.69	**t4**	**c4**	1.69
MP2 aug-cc-pwCVDZ	**t1**	**c1**	1.67	**t4**	**c4**	1.65
MP2 d-aug-cc-pVDZ	**t1**	**c1**	1.64	**t4**	**c4**	1.57
MP2 d-aug-cc-pwCVDZ	**t1**	**c1**	1.63	**t4**	**c4**	1.58
MP2(all electron) aug-cc-pwCVDZ	**t1**	**c1**	1.66	**t4**	**c4**	1.62
MP2 aug-cc-pVTZ	**t1**	**c1**	1.64			

[a] MMFF94x computations include implicit solvation within MOE, are not corrected for thermal vibrational effects, and are shown for comparison purposes only.
[b] Vibrational thermal energy/entropy corrections are from MP2 6-31G(d') computations.

Table 3. Dipole moments (in Debye) for composite methods.

	MP2 aug-cc-pVDZ (gas)	MP2 (all electron) aug-cc-pwCVDZ (gas)	**CM1** (gas)	**CM2** (gas)	Δ(CCSD) (gas)	MP2 aug-cc-pVDZ (soln)	MP2 (all electron) aug-cc-pwCVDZ (soln)	**CM1** (soln)	**CM2** (soln)	Δ(CCSD) (soln)
t1	1.4737	1.4727	1.4881	1.4742	1.5869	2.3353	2.3336	2.2995	2.3378	2.3467
t2	1.6456	1.6459	1.6371	1.6487	1.7827	2.4608	2.4609	2.4152	2.4654	2.4232
t3	1.4460	1.4451	1.4300	1.4459	1.5341	2.2368	2.2348	2.2150	2.2392	2.2267
t4	1.8821	1.8821	1.8040	1.8855	2.0269	2.5667	2.5671	2.5003	2.5711	2.5434
t5	1.6195	1.6195	1.5975	1.6220	1.7580	2.4234	2.4232	2.3933	2.4274	2.4023
t6	1.4762	1.4751	1.4701	1.4770	1.5849	2.3668	2.3652	2.3305	2.3695	2.3765
t7	1.6063	1.6064	1.5942	1.6092	1.7462	2.4530	2.4537	2.4019	2.4573	2.4581
t8	1.9250	1.9260	1.8935	1.9285	2.0784	2.5750	2.5740	2.5254	2.5794	2.5756
t9	1.6457	1.6452	1.6110	1.6459	1.7174	2.3067	2.3064	2.2972	2.3097	2.2941
c1	1.3479	1.3469	1.3598	1.3500	1.4345	2.2173	2.2155	2.1748	2.2219	2.2066
c2	1.3407	1.3398	1.3513	1.3426	1.4319	2.1953	2.1940	2.1485	2.1999	2.1936
c3	1.7747	1.7748	1.7727	1.7765	1.9338	2.6119	2.6121	2.5745	2.6143	2.6247
c4	1.7288	1.7292	1.7304	1.7303	1.8897	2.5949	2.5951	2.5609	2.5970	2.6092
c5	1.4492	1.4478	1.4563	1.4495	1.5470	2.2774	2.2745	2.2376	2.2793	2.2921
c6	1.7298	1.7295	1.7149	1.7312	1.8940	2.5912	2.5908	2.5379	2.5933	2.6063
c7	1.7222	1.7237	1.6650	1.7224	1.8007	2.4267	2.4266	2.3928	2.4288	2.4189
c8	1.9139	1.9144	1.8609	1.9172	1.9109	2.5575	2.5577	2.4989	2.5622	2.5527
c9	1.8986	1.8991	1.8643	1.9023	1.9008	2.5414	2.5409	2.4811	2.5459	2.3926
trans	1.6226	1.6222	1.6070	1.6244	1.7451	2.4475	2.4461	2.4366	2.4503	2.4279
cis	1.6202	1.6196	1.7650	1.6218	1.7077	2.4834	2.4855	2.4287	2.4893	2.4619
Δμ(*trans-cis*)	0.0024	0.0026	-0.1580	0.0026	0.0373	-0.0358	-0.0393	0.0079	-0.0389	-0.0340