Accurate dipole polarizabilities for water clusters $n=$**2-12** at the coupled-cluster level of theory and benchmarking of various density functionals

Jeff R. Hammond[*]

Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439 and

The James Franck Institute, Department of

Chemistry and Department of Computer Science,

The University of Chicago, Chicago, IL 60637

Niranjan Govind and Karol Kowalski

William R. Wiley Environmental Molecular Sciences Laboratory,

Pacific Northwest National Laboratory,

K8-91, P. O. Box 999, Richland, WA 99352

Jochen Autschbach

Department of Chemistry, University of Buffalo,

State University of New York, Buffalo, NY 14260

Sotiris S. Xantheas[†]

Chemical and Materials Sciences Division,

Pacific Northwest National Laboratory, 902 Battelle Boulevard,

P.O. Box 999, MS K1-83, Richland, WA 99352

(Dated: Revised October 21, 2009)

[*] Electronic mail: jhammond@mcs.anl.gov

[†] Electronic mail: sotiris.xantheas@pnl.gov

Abstract

The static dipole polarizabilities of water clusters ($2 \leq N \leq 12$) are determined at the Coupled-Cluster level of theory (CCSD). For the dipole polarizability of the water monomer it was determined that the role of the basis set is more important than that of electron correlation and that the basis set augmentation converges with two sets of diffuse functions. The CCSD results are used to benchmark a variety of density functionals while the performance of several families of basis sets (Dunning, Pople and Sadlej) in producing accurate values for the polarizabilities was also examined. The Sadlej family of basis sets was found to produce accurate results when compared to the ones obtained with the much larger Dunning basis sets. It was furthermore determined that the PBE0 density functional with the aug-cc-pVDZ basis set produces overall remarkably accurate polarizabilities at a moderate computational cost.

I. INTRODUCTION

Liquid water is a ubiquitous solvent, which is important in sustaining life on earth through its participation in many chemical and biological functions. Simulations of liquid water are presently performed primarily in two varieties: first-principles molecular dynamics simulations (FPMD) wherein the energies and forces are usually computed within the density functional theory (DFT) framework and those where a classical potential (force field) is employed. The complexity of models associated with both types of approaches has increased significantly over the years. Whereas in the past FPMD with simple gradient-corrected functionals (BP86) [1] and the classical TIP3P water model [2] constituted the standard approaches, the pursuit of greater accuracy has led to models of ever increasing complexity. To this end, several Generalized Gradient Approximation (GGA) and GGA-hybrid functionals have been applied to water (i.e. Ref.[3]), whereas classical force fields have evolved from simple, pairwise-additive, rigid-body models like TIP3P to more sophisticated, many-body polarizable, flexible models which incorporate various levels of the relevant physics (see Refs. [4–12] for a few examples).

The fundamental interactions between water molecules have been probed during high-resolution experiments on water clusters which have yielded the minimum-energy structures for small clusters [13, 14]. These experiments affirmed early benchmark quantum chemical studies for water clusters up to the hexamer [15]. Over the past 20 years, the synergy between theory and experiment has led to a detailed understanding of the competing physical effects that determine the average structure, such as the competition between hydrogen bond count and conformational strain as well as the vital role of zero-point vibrational energy in determining the most stable isomer among competing structures (cf. water hexamer [16]).

There have been numerous investigations [17, 18] of the performance of DFT for polarizabilities focusing on either small molecules or the shortcomings of approximate exchange-correlation functionals for extended systems [19]. At least three groups have previously considered the electric properties of water clusters [20–23], while Santra and coworkers [24] benchmarked DFT for other properties of water clusters. The present study extends the previous work by considering the polarizability of water clusters at the CCSD linear response (LR) level of theory [25], which combines the advantages of both yielding highly accurate polarizabilities [26] and being able to be applied to large molecular systems [27, 28]. The

role of the Gaussian basis set is evaluated by comparing results using basis sets from the Dunning [29], Pople [30] and Sadlej [31] families. Finally, we consider a variety of density functionals (B3LYP [32–34], PBE0 [35, 36], X3LYP [37], HCTH120 [38, 39], Becke98 [40], CAMB3LYP [41] and CAMPBE0 [42]), not all of which have been previously evaluated for large molecular systems. For typical levels of theory like Hartree-Fock (HF) and MP2, the shortcomings of the former for polarizabilities are well-documented [15], whereas the accuracy of the latter for polarizabilities is not sufficiently greater than DFT to justify the increase in computational cost, except in extended systems where the DFT electric-field response is non-physical [19] and thus conventional DFT cannot be used (see Ref. [43] for one solution).

The main goal of the present study is to identify which density functionals most accurately reproduce the polarization in assemblies of water molecules, when compared to the most accurate CC results, and are therefore more appropriate to be used for the simulation of the properties of large assemblies of water molecules and liquid water. The paper is organized as follows: In Section II we outline the computational details. In Section III we present accurate results for the dipole polarizabilities of water clusters and evaluate the effects of the basis set and the level of electron correlation. In that section we also compare the results of various density functionals with the high-level wavefunction-theory results. Final conclusions are drawn in Section IV.

II.   COMPUTATIONAL DETAILS

In order to obtain accurate values for the polarizabilities of water clusters, a hierarchy of calculations has been performed. First, the polarizabilities of the water monomer and dimer were investigated with high level correlation methods — up to CCSDTQ for the monomer and CCSDT for the dimer — to understand correlation effects. In addition, the basis set dependence of the monomer was investigated at the SCF, CCSD and CCSDT level for up to sextuple-zeta and triply-augmented basis sets. Second, the polarizabilities of small water clusters (n=1-4) were calibrated at the CCSD level of theory with the aug-cc-pVXZ (X=D,T,Q) basis sets to quantify the basis set dependence of CCSD as intermolecular interactions become important. Third, the results with basis sets from the Dunning [29], Pople [30] and Sadlej [31] families were compared with each other at the DFT level of theory

for large clusters. Finally, the DFT polarizabilities of clusters were compared to the CCSD results for the aug-cc-pVDZ basis set, which was found to be the best compromise between accuracy and computational cost due to the relatively small differences between aug-cc-pVDZ and aug-cc-pVTZ for larger clusters. For example, the isotropic polarizability of the trimer with the aug-cc-VDZ, aug-cc-VTZ and aug-cc-VQZ basis sets are 28.94, 29.02 and 28.94, respectively. Similarly small differences are observed for the tetramer.

The labeling conventions for the various isomers of the various water clusters used in the figures are defined in Table I.

The coupled-cluster (CC) dipole polarizabilities were computed with the NWChem [44] Tensor Contraction Engine (TCE) response property module as described in Refs. [27] and [45]. The largest CCSD-LR calculation — $(H_2O)_{12}$ with the aug-cc-pVDZ basis set — was run on 256 nodes (4 cores per node) of the Chinook supercomputer with wall times per iteration of ~ 600, ~ 1300 and ~ 900 seconds for the T, Λ and $T^{(1)}$ equations, respectively, for a total wall time of ~ 33 hours.

The polarizability calculations at the DFT level, including asymptotically-corrected functionals [46], were also performed with NWChem using the implementation by Autschbach [47] which is based on techniques reported in Refs.[48–51] and references cited therein. For comparison, the $(H_2O)_{12}$ calculation with the CAMPBE0 functional and the aug-cc-pVDZ basis set took less than an hour on 8 nodes (4 cores per node). Asymptotically-corrected functionals require approximately twice the wall time as conventional GGA hybrids under the current implementation.

The water monomer and dimer polarizability calculations at the CCS and CC2 [52] levels of theory were performed using Dalton 2.0 [53], while MP2, CCSD(T) [54] and CC3 [55] results were obtained using Aces II (MAB) [56]. Results for CCSDT and CCSDTQ (first implemented in Ref. [57]) were obtained with NWChem as described in Ref. [26]. To ensure compatibility between codes, spherical angular functions were employed for calculations of the monomer and dimer (presented in Tables II and III and Figures 1 and 2). Larger cluster calculations (including those reported in Table IV) used Cartesian angular functions [58].

Where there was a choice between orbital-relaxed and orbital-unrelaxed (response) property evaluation in the iterative coupled-cluster methods, the response formalism was used. It is well-known that the orbital-relaxed formalism cannot be used for dynamic properties due to the introduction of spurious poles [52], and since many experiments are performed in this

5

context, it is more useful to understand the accuracy of response methods. Additionally, the NWChem implementation of CC properties supports only the response formalism at this time and codes which implement orbital-relaxed properties are not capable of treating the largest systems considered here.

We used the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ optimal cluster geometries from earlier studies for the water dimer [59], trimer through hexamers [5], octamers (D_{2d} and S_4) [60], $n = 11, 13$ [63], and $n = 7, 9, 10, 12, 14, 15$ and 16 [64].

All correlated calculations use the frozen core approximation.

III.   RESULTS

A.   Electron correlation and basis set effects for the polarizability of the water monomer and dimer

High-level calculations which approach the limit for the treatment of both electron correlation and the completeness of the orbital basis set were performed in order to arrive at a converged value for the monomer as well as investigate the possible limitations for the larger clusters. In Table II, we compare wavefunction-based methods ranging from SCF to CCSDT for Dunning's aug-cc-pVDZ and aug-cc-pVTZ basis sets as well as CCSDTQ with the smaller aug-cc-pVDZ basis. The convergence in the correlation with the aug-cc-pVDZ basis for the isotropic polarizability component of the monomer and dimer is displayed in Figure 1.

As noted in earlier studies [65–70], the polarizability of the water monomer has a strong dependence on the treatment of both electron correlation and the orbital basis set. The MP2/aug-cc-pVDZ results are quite accurate for the isotropic component for both the monomor and dimer but noticable less accurate for the anisotropy of the monomer. Other low-order methods, such as SCF, CCS and CC2, are much less accurate for α_{iso}. Although the MP2 and CC2 results are close to the CCSDT values for the dimer anisotropy with the aug-cc-pVDZ basis, the monomer results suggest this is an artifact for this basis, and that for aug-cc-pVTZ and larger sets, CCSD will be closer to CCSDT.

The approximate triples method CC3 reproduces the CCSDT result to within 0.003 a.u. and appears to be a very useful method for including triple excitations without the storage

requirements of CCSDT (see also Ref. [70]). The effect of quadruples (quantified by the difference between the CCSDTQ and CCSDT results) is almost one-third of that of the triples (quantified by the difference between the CCSDT and CCSD results): -0.019 a.u. versus -0.065 a.u. for the isotropic polarizability. While this is interesting in the context of theoretical benchmarking for the polarizability of the water monomer, higher-order correlation is not likely to play a significant role for the larger clusters, as these corrections would most likely be dwarfed by lower-order correlation effects introduced by double excitations. In addition, the calculation of the larger clusters, even the dimer, is not currently feasible at the CCSDTQ level of theory even with the smaller aug-cc-pVDZ basis set. While CCSDT may be possible for small clusters using the aug-cc-pVDZ basis set, this will not produce meaningful results due to the imbalance of using high-level correlation treatment with a double-zeta basis set, as is clear from the monomer data. The best way to quantify correlation beyond CCSD in clusters using an appropriate basis set would be with a parallel implementation of CC3-LR, which has not yet been implemented in any suite of electronic structure software.

The results with the aug-cc-pVDZ and aug-cc-pVTZ basis sets indicate that the numbers listed in Table II are not close to being converged with the basis set. Table III lists the results at the SCF, CCSD and CCSDT levels of theory for very large basis sets, up to sextuple-zeta quality and several levels of augmentation within the Dunning construction scheme [29]. In agreement with previous findings [15] it is obvious that SCF is inadequate for both the isotropic and anisotropic components of the polarizability, demonstrating the large role of electron correlation in the dipole polarizability, even for a simply hydrogen-bonded system near its equilibrium geometry. The difference in α_{iso} between CCSD and CCSDT appears large with the aug-cc-pVDZ basis (0.065 au) but it drops to 0.011 a.u. for the d-aug-cc-pVQZ basis, a trend observed previously for other small molecules [26]. The opposite trend is observed for the anisotropic component, with the effect of triples increasing from -0.008 a.u. with aug-cc-pVDZ to 0.046 a.u. for d-aug-cc-pVQZ. The comparison between the CCSD and CCSDT results for the isotropic polarizability of the water monomer indicates that the role of triples is less than 1% and that calculations at the CCSD level of electron correlation can yield quite accurate results. Further evidence for the usefulness of CCSD is that the effect of triples on the polarizability is less noticeable for larger molecules, such as benzene [27, 71].

For the dipole polarizability of the water monomer it was furthermore determined that the role of the basis set is even more important than that of electron correlation. For the (singly) augmented basis sets of the Dunning series, aug-cc-pVXZ, the change in the polarizability, and in particular the anisotropic component, is still noticeable between the quadruple- and quintuple-zeta quality sets at the CCSDT level. However, as demonstrated in Figure 2, the convergence in the cardinal number with both the singly- and doubly-augmented series is systematic. The basis set augmentation converges with two sets of diffuse functions (d-aug) level since the addition of a third set of diffuse functions (t-aug) has a negligible effect.

While the monomer polarizability is not the primary target of the present study, the accurate electronic polarizabilities computed here can be compared to experiment upon inclusion of the vibrational contribution of Bishop and Cheung (0.25 a.u.) [72]. If we add this to the estimated CCSDT/CBS value ($\sim$9.58 a.u.) the total — 9.83 a.u. — is between the experimental result of Zeiss and Meath (9.63 a.u.) [73] and that of Murphy (9.92 a.u.) [74].

B. Accurate polarizabilities of water clusters n=1-4

Because of the observed strong dependence of the monomer's polarizability on both the level of electron correlation and basis set, small water clusters were first studied using high-level methods in order to obtain reasonable results for their polarizabilities. In Table IV we consider the CCSD and CCSD(T) total energies and CCSD-LR isotropic and anisotropic polarizabilities for the aug-cc-pVXZ basis sets (X=D,T,Q) for small water clusters. Although, both the basis set and electron correlation dependence of the total energies are appreciable, the basis set dependence of the polarizabilities is, in contrast, relatively small, and our results for the monomer and dimer are consistent with those of Maroulis [65, 68, 69]. There is a variance with the basis set size for the anisotropic component, which is larger than the one for the isotropic component. While the change from the aug-cc-pVTZ to aug-cc-pVQZ basis for α_{ani} is larger than that from the aug-cc-pVDZ to aug-cc-pVTZ, α_{iso} changes negligibly in both cases for the water dimer and trimer clusters. These results indicate that the use of CCSD/aug-cc-pVDZ polarizabilities to calibrate the polarizabilities of the larger clusters is justified.

C. Benchmarking of various families of basis sets for computing water cluster polarizabilities

The basis set dependence of the water cluster polarizabilities was evaluated by comparing the results with the aug-cc-pVDZ and aug-cc-pVTZ basis sets. For the CCSD level of theory and all six DFT functionals considered in this study, the absolute difference in the isotropic and anisotropic components of the polarizability was less than 0.5 a.u. and generally decreased with cluster size. For the isotropic component, the polarizability change from the aug-cc-pVDZ to the aug-cc-pVTZ basis set was always positive. The difference between the results obtained with the aug-cc-pVDZ and the aug-cc-pVTZ basis sets for the decamer was 0.21, 0.14 and 0.21 au for the B3LYP, PBE0 and B98 functionals, respectively. The effect of the basis set on the polarizability per water molecule is thus very small for the larger clusters with the family of Dunning basis sets.

In addition to verifying that the aug-cc-pVDZ basis was sufficient for the purposes of comparing methods, numerous other basis sets were evaluated with the B3LYP and PBE0 functionals. In Figure 3, cluster polarizabilities with PBE0 are reported for five commonly used Pople family basis sets: 6-31+G*, 6-31++G**, 6-311+G*, 6-311++G** and 6-311++G($3df$,$3pd$). While the Pople basis sets systematically approach the apparent CBS limit for the isotropic polarizability, they do so very slowly. Clearly, a critical component is missing from the basis set which results in the large gap between the 6-311++G** and 6-311++G($3df$,$3pd$) and between 6-311++G($3df$,$3pd$) and aug-cc-pVQZ, the largest basis set considered. The comparison suggests that the less-complete Pople basis sets are insufficient for producing accurate polarizabilities due to the lack of additional diffuse functions, which are present in both the Dunning and Sadlej basis set families and, to some extent, in the 6-311++G($3df$,$3pd$) set. Four of the Pople family of basis sets do not reproduce either the magnitude of the polarizability per molecule — which is 9.9 for the Dunning basis sets but less than 9 for all but 6-311++G($3df$,$3pd$) — or the qualitative trends in the relative magnitude of the polarizability for different isomers of clusters. For example, for the water hexamer the ordering of the polarizabilities of the book and cyclic isomers changes between the Pople and Dunning families of basis sets. Only the 6-311++G($3df$,$3pd$) basis, which is twice as large as aug-cc-pVDZ, produces the qualitative features of the hexamer seen with all of the Dunning and Sadlej basis sets. This is a problem if the principle of minimum

9

polarizability is to be used within the conceptual DFT framework to predict the stability of various isomers. The relative ordering of the polarizabilities is also reversed for the two isomers of W_{17} between the Pople and Dunning basis sets. Similarly, the Pople basis sets of double-zeta quality (6-31+G* and 6-31++G**) also improperly order the polarizabilities of W_{11} while the larger triple-zeta quality Pople sets (6-311+G* and 6-311++G**) agree with the ordering predicted by the smaller aug-cc-pVDZ basis set.

The basis set convergence evaluation ofr the Dunning and Sadlej basis sets (see Figures 4 and 5) suggests that all four of the Sadlej basis sets considered here (POL, HYPOL, Z2POL, Z3POL) closely reproduced the polarizabilities obtained with the family of the Dunning basis sets and that the Dunning basis sets are themselves a rapidly converging set of approximations for the polarizabilities. The Sadlej basis sets are specifically designed for polarizability calculations and while they are of approximately double-zeta quality for energies, they produce accurate polarizabilities due to the fact that they include several diffuse functions designed for precisely this purpose. The results with the POL and HYPOL basis sets were closer to the ones obtained with the aug-cc-pVTZ basis set, whereas the results with the Z2POL and Z3POL basis sets were closer to the ones obtained with the aug-cc-pVDZ basis set. The excellent agreement between the results obtained with the POL and aug-cc-pVQZ basis sets is consistent with previous findings for polyacenes [27]. This finding is of practical significance since the POL basis set has approximately the same rank as the aug-cc-pVDZ. The only downside of using the computationally less expensive POL basis set is the absence of validation of convergence via a systematic hierarchy of basis sets which can be extended in both cardinality and augmentation, as is the case for the Dunning series. As seen in Figure 5, the aug-cc-pVTZ and aug-cc-pVQZ basis sets are nearly indistinguishable for $n \geq 4$ and the difference between these two sets and the smaller aug-cc-pVDZ decreases smoothly with increasing n. This provides substantial evidence that the use of the aug-cc-pVDZ basis set for benchmarking the larger clusters is a reasonable method.

D. Benchmarking of various density functionals for computing water cluster polarizabilities

The isotropic polarizability component per water molecule for water clusters up to W_{12} for CCSD with the aug-cc-pVDZ basis set is shown in Figure 6. The errors with respect

to CCSD for the various density functionals are shown in Figure 7. While all functionals display relatively similar trends, the rate at which each functional reaches the large cluster value of its percent error with respect to the CCSD result varies. For both the HCTH120 and B98 functionals, the error decreases rapidly from $n=1$-4, where for the PBE0, B3LYP and X3LYP functionals the error remains almost constant for all clusters. While the asymptotic (large cluster size) error is approximately already reached for the hexamer, estimating the polarizability itself requires substantially larger clusters due to the fluctuations with respect to the different networks present in the different isomers. Starting at the heptamer, the polarizability per molecule is essentially constant as the cluster size increases, suggesting large cluster values for α_{iso}/n of ca. 9.2, 9.6, 9.7, 9.8 and 9.8 for B98, HCTH120, PBE0, B3LYP and X3LYP, respectively. The B3LYP and X3LYP functionals produce essentially indistinguishable results for the polarizability. Finally, the PBE0 and CAMB3LYP functionals are clearly most able to reproduce the CCSD isotropic polarizabilities, with errors ranging from 0.17% to 0.46% and -0.53% to 0.22%, respectively, for $2 \leq n \leq 21$.

The polarizability anisotropy, whose magnitude is highly dependent on the cluster hydrogen bonding topology, is shown in Figure 8. Results are reported only up to $n=15$ because large clusters have a substantially larger anisotropy, a fact that tends to obscure the trends; also the CCSD results are available only up to $n \leq 12$. The CCSD anisotropy is usually lower than the DFT results, with most of the exceptions occurring either for the octamer or the CAM functionals. Except for HCTH120, all the DFT functionals considered here are able to reproduce the CCSD anisotropy to within 15% for $n < 10$. This can be attributed to the importance of exact exchange in obtaining an accurate anisotropy, since HCTH120 is a pure GGA functional. Due to the stronger basis set dependence of the anisotropic polarizability, the aug-cc-pVDZ set most likely produces an upper-bound for both DFT and CC. For CC, the evidence for this can be seen in Table IV. Nonetheless, the observation that DFT exagerrates α_{ani} with respect to CCSD is unlikely to change with basis set saturation.

IV. CONCLUSIONS

Accurate results for the electric properties of water clusters have been obtained using CC methods. These results have beed subsequently used to evaluate the accuracy of seven DFT functionals. The functionals considered in this study were chosen from a much larger set

based upon their ability to accurately predict the polarizability of a single water molecule and/or because they are common functionals used to simulate both water clusters and bulk water. In particular, the HCTH class of functionals has been suggested within the CPMD community as being capable of reproducing important macroscopic properties of liquid water [75, 76]. However, the results of this study demonstrate that PBE0 is much superior to the other five functionals for yielding accurate polarizabilities, and by inference, superior to many other functionals not considered here that are known to produce even less accurate polarizabilities. We found that functionals not containing some portion of the exact (Hartree-Fock) exchange were significantly less accurate for electric properties than their hybrid counterparts. In this respect, PBE0 is more accurate than PBE, B3LYP is better than BLYP, and so forth. The ramifications, specifically the computational cost, of employing exact exchange during CPMD simulations are greater than in molecular simulations with local basis functions [77]. However, exact exchange cannot be neglected especially for simulations intending to directly quantify polarization. It is quite clear that the HCTH functional should not be immediately adopted as a substitute for BLYP and that more research targeted towards the development of new density functionals is necessary.

V.   ACKNOWLEDGMENTS

[1] K. Laasonen, M. Sprik, M. Parrinello and R. Car, *J. Chem. Phys.* **99**, 9080 (1993).

[2] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys* **79**, 926 (1983).

[3] M. Guidon, F. Schiffmann, J. Hutter and J. VandeVondele, *J. Chem. Phys.* **128**, 214104 (2008).

[4] C. J. Burnham and S. S. Xantheas, *J. Chem. Phys.* **116**, 1479 (2002).

[5] S. S. Xantheas, C. J. Burnham and R. J. Harrison, *J. Chem. Phys.* **116**, 1493 (2002).

[6] C. J. Burnham and S. S. Xantheas, *J. Chem. Phys.* **116**, 1500 (2002)

[7] C. J. Burnham and S. S. Xantheas, *J. Chem. Phys.* **116**, 5115 (2002).

[8] G. S. Fanourgakis and S. S. Xantheas, *J. Phys. Chem. A* **110**, 4100 (2006)

[9] G. S. Fanourgakis and S. S. Xantheas, *J. Chem. Phys.* **128**, 074506 (2008)

[10] G. Lamoureux, A. D. MacKerell Jr. and B. Roux, *J. Chem. Phys.* **119**, 5185 (2003).

[11] J. Chen and T. Martínez, *Chem. Phys. Lett.* **438**, 315 (2007).

[12] H. A. Stern, F. Rittner, B. J. Berne and R. A. Friesner, *J. Chem. Phys.* **115**, 2237 (2001).

[13] N. Pugliano and R. J. Saykally, *Science* **257**, 1937 (1992); J. D. Cruzan, L. B. Braly, K. Liu, M. G. Brown, J. G. Loeser and R. J. Saykally, *Science* **271**, 59 (1996); K. Liu, M. G. Brown, J. D. Cruzan and R. J. Saykally, *Science* **271**, 62 (1996); K. Liu, J. D. Cruzan and R. J. Saykally, *Science* **271**, 929 (1996); J. K. Gregory, D. C. Clary, K. Liu, M. G. Brown and R. J. Saykally, *Science* **275**, 814 (1997); J. B. Paul, R. A. Provencal, C. Chapo, K. Roth, R. Casaes and R. J. Saykally, *J. Phys. Chem. A* 103, 2972 (1999); J. B. Paul, C. P. Collier and R. J. Saykally, *J. Phys. Chem. A* **101**, 5211 (1997); M. G. Brown, F. N. Keutsch and R. J. Saykally, *J. Chem. Phys.* **109**, 9645 (1998); Frank N. Keutsch and R. J. Saykally, *Proc. Nat. Acad. Sci.* **98**, 10533 (2001); F. N. Keutsch, M. G. Brown, P. B. Petersen, R. J. Saykally, M. Geleijns and A. van der Avoird, *J. Chem. Phys.* **114**, 3994 (2001); F. N. Keutsch, R. S. Fellers, M. R. Viant and R. J. Saykally, *J. Chem. Phys.* **114**, 4005 (2001).

[14] K. Nauta and R. E. Miller, *Science* **287**, 293 (2000).

[15] S. S. Xantheas and T. H. Dunning Jr., *J. Chem. Phys.* **99**, 8774 (1993); S. S. Xantheas, *J. Chem. Phys.* **100**, 7523 (1994); S. S. Xantheas, *J. Chem. Phys.* **102**, 4505 (1995).

[16] K. Liu, M. G. Brown, C. Carter, R. J. Saykally, J. K. Gregory and D. C. Clary, *Nature* **381**, 501 (1996).

[17] M. J. Stott and E. Zaremba, *Phys. Rev. A* **21**, 12 (1980); S. J. A. van Gisbergen, J. G. Snijders, and E. J. Baerends, *J. Chem. Phys.* **103**, 9347 (1995); C. Jamorski, M. E. Casida, and D. R. Salahub, *J. Chem. Phys.* **104**, 5134 (1996); S. J. A. van Gisbergen, V. P. Osinga, O. V. Gritsenko, R. van Leeuwen, J. G. Snijders, and E. J. Baerends, *J. Chem. Phys.* **105**, 3142 (1996); P. Calaminici, K. Jug and A. M. Koster, *J. Chem. Phys.* **109**, 7756 (1998); A. J. Cohen and Y. Tantirungrotechai, *Chem. Phys. Lett.* **299**, 465 (1999); C. Adamo, M. Cossi, G. Scalmani and V. Barone, *Chem. Phys. Lett.* **307**, 265 (1999); S. J. A. van Gisbergen, J. M. Pacheco, and E. J. Baerends, *Phys. Rev. A* **63**, 063201 (2001); P. Mori-Sanchez, Q. Wu, and W. Yang, *J. Chem. Phys.* **119**, 11001 (2003); H. Sekino, Y. Maeda, M. Kamiya, and K. Hirao, *J. Chem. Phys.* **126**, 014107 (2007).

[18] P. Sałek, T. Helgaker, O. Vahtras, H. Ågren, D. Jonssen and J. Gauss, *Mol. Phys.*, **103**, 439 (2005).

[19] B. Champagne, E. A. Perpete, S. J. A. van Gisbergen, E.-J. Baerends, J. G. Snijders, C. Soubra-Ghaoui, K. A. Robins and B. Kirtman, *J. Chem. Phys.* **109**, 10489 (1998); S. J. A. van Gisbergen, P. R. T. Schipper, O. V. Gritsenko, E. J. Baerends, J. G. Snijders, B. Champagne, and B. Kirtman, *Phys. Rev. Lett.* **83**, 694 (1999); B. Champagne, E. A. Perpète, D. Jacquemin, S. J. A. van Gisbergen, E. J. Baerends, C. Soubra-Ghaoui, K. Robins, and B. Kirtman, *J. Phys. Chem. A* **104**, 4755 (2000); B. Champagne, F. A. Bulat, W. Yang, S. Bonness and B. Kirtman, *J. Chem. Phys* **125**, 194114 (2006).

[20] T. K. Ghanty and S. K. Ghosh, *J. Chem. Phys.* **118**, 8547 (2003).

[21] M. Yang, P. Senet and C. Van Alsenoy, *Int. J. of Quantum Chem.* **101**, 535 (2004).

[22] A. Krishtal, P. Senet, M. Yang and C. Van Alsenoy, *J. Chem. Phys.* **125**, 034312 (2006).

[23] A. Morita, *J. Comp. Chem. 23*, 1466 (2002).

[24] B. Santra, A. Michaelides, and M. Scheffler *J. Chem. Phys.* **127**, 184104 (2007). B. Santra, A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi and M. Scheffler *J. Chem. Phys.* **129**, 194111 (2008).

[25] H. J. Monkhorst, *Int. J. Quantum Chem.* **S11**, 421 (1977); H. Koch and P. Jørgensen, *J. Chem. Phys.* **93**, 3333 (1990); O. Christiansen, P. Jørgensen and C. Hättig, *Int. J. Quantum Chem.* **68**, 1 (1998).

[26] J. R. Hammond and K. Kowalski, *J. Chem. Phys.*, **128**, 224102 (2008); the CCSDTQ implementation, while not published, uses an identical strategy to CCSDT.

[27] J. R. Hammond, K. Kowalski and W. A. de Jong, *J. Chem. Phys.*, **127**, 144105 (2007).

[28] K. Kowalski, J. R. Hammond, W. A. de Jong and A. J. Sadlej, *J. Chem. Phys.* **129**, 226101 (2008).

[29] T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989); R. A. Kendall, T.H. Dunning Jr. and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992); D. E. Woon and T.H. Dunning Jr., *J. Chem. Phys.* **98**, 1358 (1993); D. E. Woon and T.H. Dunning Jr., *J. Chem. Phys.* **100**, 2975 (1994).

[30] P. C. Hariharan and J. A. Pople, *Theor. Chimica Acta* **28**, 213 (1973); R. Krishnan, J.S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.* **72**, 650 (1980); M. M. Francl, W. J. Petro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees and J. A. Pople, *J. Chem. Phys.* **77**, 3654 (1982); T. Clark, J. Chandrasekhar and P. v. R. Schleyer, *J. Comp. Chem.* **4**, 294 (1983); P. M. W. Gill, B. G. Johnson, J. A. Pople and M. J. Frisch, *Chem. Phys. Lett.* **197**, 499 (1992).

[31] A. J. Sadlej, *Coll. Czech. Chem. Comm.* **53**, 1995 (1988); A. J. Sadlej and M. Urban, *J. Mol. Struct. (THEOCHEM)* **234**, 147 (1991); A. J. Sadlej, *Theor. Chim. Acta* **79**, 123 (1992); A. J. Sadlej, *Theor. Chim. Acta* **81**, 45 (1992); A. J. Sadlej, *Theor. Chim. Acta* **81**, 339 (1992).

[32] A. D. Becke, *Phys. Rev. A* 88, 3098 (1988);

[33] C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988);

[34] A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).

[35] J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996); **78**, 1396 (1997).

[36] C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1998).

[37] X. Xu and W. A. Goddard III, *Proc. Nat. Acad. Sci.* **101**, 2673 (2004).

[38] F. A. Hamprecht, A. J. Cohen, D. J. Tozer and N. C. Handy, *J. Chem. Phys.* **109**, 6264 (1998).

[39] A. D. Boese, N. L. Doltsinis, N. C. Handy and M. Sprik. *J. Chem. Phys.* **112**, 1670 (2000).

[40] H. L. Schmider and A. D. Becke, *J. Chem. Phys.* **108**, 9624 (1998).

[41] T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.* **393**, 51 (2004).

[42] M. A. Rohrdanz and J. M. Herbert, *J. Chem. Phys.* **129**, 034107 (2008); A. W. Lange, M. A. Rohrdanz and J. M. Herbert, *J. Phys. Chem. B* **112**, 6304 (2008).

[43] M. van Faassen, P. L. de Boeij, R. van Leeuwen, J. A. Berger, and J. G. Snijders, Phys. Rev.

Lett. **88**, 186401 (2002); M. van Faassen, P. L. de Boeij, R. van Leeuwen, J. A. Berger, and J. G. Snijders, J. Chem. Phys. **118**, 1044 (2003).

[44] E. J. Bylaska, W. A. de Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, D. Wang, E. Aprá, T. L. Windus, J. Hammond, J. Autschbach, P. Nichols, S. Hirata, M. T. Hackler, Y. Zhao, P.-D. Fan, R. J. Harrison, M. Dupuis, D. M. A. Smith, J. Nieplocha, V. Tipparaju, M. Krishnan, A. Vazquez-Mayagoitia, Q. Wu, T. Van Voorhis, A. A. Auer, M. Nooijen, L. D. Crosby, E. Brown, G. Cisneros, G. I. Fann, H. Fruchtl, J. Garza, K. Hirao, R. Kendall, J. A. Nichols, K. Tsemekhman, K. Wolinski, J. Anchell, D. Bernholdt, P. Borowski, T. Clark, D. Clerc, H. Dachsel, M. Deegan, K. Dyall, D. Elwood, E. Glendening, M. Gutowski, A. Hess, J. Jaffe, B. Johnson, J. Ju, R. Kobayashi, R. Kutteh, Z. Lin, R. Littlefield, X. Long, B. Meng, T. Nakajima, S. Niu, L. Pollack, M. Rosing, G. Sandrone, M. Stave, H. Taylor, G. Thomas, J. van Lenthe, A. Wong and Z. Zhang. "NWChem, A Computational Chemistry Package for Parallel Computers, Version 5.1.1" (2009), Pacific Northwest National Laboratory, Richland, Washington 99352-0999, USA. A modified version.

[45] J. R. Hammond, M. Valiev, W. A. de Jong and K. Kowalski, *J. Phys. Chem. A*, **111**, 5492 (2007).

[46] N. Govind, M. Valiev, L. Jensen and K. Kowalski, *J. Phys. Chem. A*, Article ASAP (2009).

[47] J.Autschbach, Comp. Lett. **3**, 131 (2007).

[48] M. Krykunov and J. Autschbach, J. Chem. Phys. **123**, 114103 (2005).

[49] M. Krykunov and J. Autschbach, J. Chem. Phys. **125**, 034102 (2006).

[50] M. Krykunov and J. Autschbach, J. Chem. Phys. **126**, 024101 (2007).

[51] L. Jensen, J. Autschbach, M. Krykunov and G. C. Schatz, J. Chem. Phys. **127**, 134101 (2007).

[52] O. Christiansen, H. Koch, P. Jørgensen, *Chem. Phys. Lett.* **243**, 409 (1995).

[53] *DALTON, a molecular electronic structure program, Release 2.0* (2005), see `http://www.kjemi.uio.no/software/dalton/dalton.html`.

[54] M. Urban, J. Noga, S. J. Cole and R. J. Bartlett, *J. Chem. Phys.* **83**, 4041 (1985). K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989). K. Raghavachari, J. A. Pople, E. S. Replogle and M. Head-Gordon, *J. Phys. Chem.* **94**, 5579 (1990).

[55] O. Christiansen, H. Koch and P. Jørgensen, *J. Chem. Phys.* **103**, 7429 (1995); H. Koch, O. Christiansen, P. Jørgensen, A. M. Sanchez de Meras and T. Helgaker, *J. Chem. Phys.* **106**,

1808 (1997); K. Hald, F. Pawlowski, P. Jørgensen and C. Hättig, *J. Chem. Phys.* **118**, 1292 (2003).

[56] *Aces II*, a quantum chemical program package written by J. F. Stanton, J. Gauss, J .D. Watts, P. G. Szalay, R. J. Bartlett with contribution from A. A. Auer, D. B. Bernholdt, O. Christiansen, M. E. Harding, M. Heckert, O. Heun, C. Huber, D. Jonsson, J. Juselius, W. J. Lauderdale, T. Metzroth, K. Ruud and the integral packages *MOLECULE* (J. Almlöf and P. R. Taylor), Props (P. R. Taylor), and ABACUS (T. Helgaker, H. Å. Jensen, P. Jørgensen and J. Olsen). See also J. F. Stanton, J. Gauss, J. D. Watts, W. J. Lauderdale, R.J. Bartlett, *Int. J. Quantum Chem. Symp.* **26**, 879 (1992) as well as : `http://www.aces2.de` for the current version.

[57] M. Kállay and J. Gauss, J. Mol. Struct. (THEOCHEM) **768**, 71 (2006).

[58] While the difference between spherical and Cartesian angular functions is noticable for the monomer and dimer, the effect for larger systems is nominal. For the octamer, the effect on the isotropic (anisotropic) polarizability is 0.02 a.u. (0.002 a.u.) per molecule, which does not affect the interpretation of the results.

[59] S. S. Xantheas, *J. Chem. Phys.* **104**, 8821 (1996).

[60] S. S. Xantheas and E Aprá, *J. Chem. Phys.* **120**, 823 (2004).

[61] G. S. Fanourgakis, E. Aprá and S. S. Xantheas, *J. Chem. Phys.* **121**, 2655 (2004).

[62] A. Lagutschenkov, G. S. Fanourgakis, G. Niedner-Schatteburg and S. S. Xantheas, *J. Chem. Phys.* **122**, 194310 (2005).

[63] S. Bulusu, S. Yoo, E Aprá, S. Xantheas and X. C. Zeng, *J. Phys. Chem. A* **110**, 11781 (2006)

[64] E. Aprá, J. R. Hammond and S. S. Xantheas, to be published.

[65] G. Maroulis, *J. Chem. Phys.*, **94**, 1182 (1991).

[66] B. Datta, P. Sen, D. Mukherjee, *J. Phys. Chem.* 99, 6441 (1995).

[67] D. Spelsberg and W. Meyer, *J. Chem. Phys.* **108**, 1532 (1998).

[68] G. Maroulis, *Chem. Phys. Lett.* **289**, 403 (1998).

[69] G. Maroulis, *J. Chem. Phys.*, **113**, 1813 (2000).

[70] G. Avila, *J. Chem. Phys.*, **122**, 144310 (2005).

[71] J. R. Hammond, W. A. de Jong and K. Kowalski, unpublished results.

[72] D. M. Bishop and L. M. Cheung, *J. Phys. Chem. Ref. Data* **11**, 1 (1982).

[73] G. D. Zeiss and W. J. Meath, *Mol. Phys.* **30**, 161 (1975).

[74] W. F. Murphy, *J. Chem. Phys.* **67**, 5877 (1977).

[75] M. Sprik, J. Hutter and M. Parrinello, *J. Chem. Phys.* **105**, 1142 (1996).

[76] A. D. Boese, N. L. Doltsinis, N. C. Handy and M. Sprik, *J. Chem. Phys.* **112**, 1670 (2000).

[77] T. Todorova, A. Sietsonen, J. Hutter, I.-F. W. Kuo and C. J. Mundy, *J. Phys. Chem. B* **110**, 3685 (2006).

FIG. 1: Isotropic polarizabilities with the aug-cc-pVDZ basis set for various correlation methods for the water monomer and dimer. All correlated calculations use the frozen core approximation.
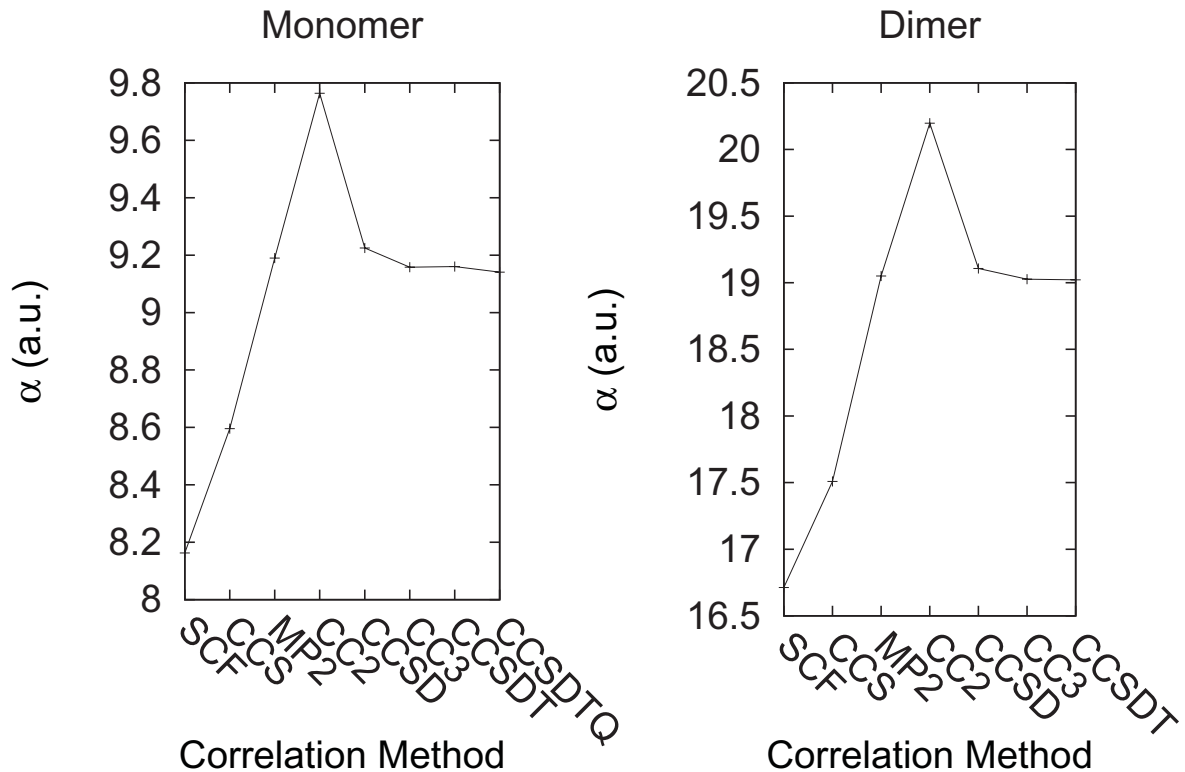

TABLE I: Labeling scheme for clusters with multiple isomers. $W_n=(H_2O)_n$.

W_6		W_{11}		W_{20}	
	b=Book		a=434		d=Dodecahedron
	c=Cage		b=4412		e=Edge-sharing
	y=cYclic		c=443		f=Face-sharing
	p=Prism		d=515		c=fused Cubes
			e=551		
W_8	d=D_{2d}	W_{17}	i=Interior		
	s=S_4		s=Surface		

FIG. 2: Basis set convergence at the SCF, CCSD and CCSDT (frozen core) levels of theory for the monomer. The triply augmented series are indistinguishable from the doubly augmented series.
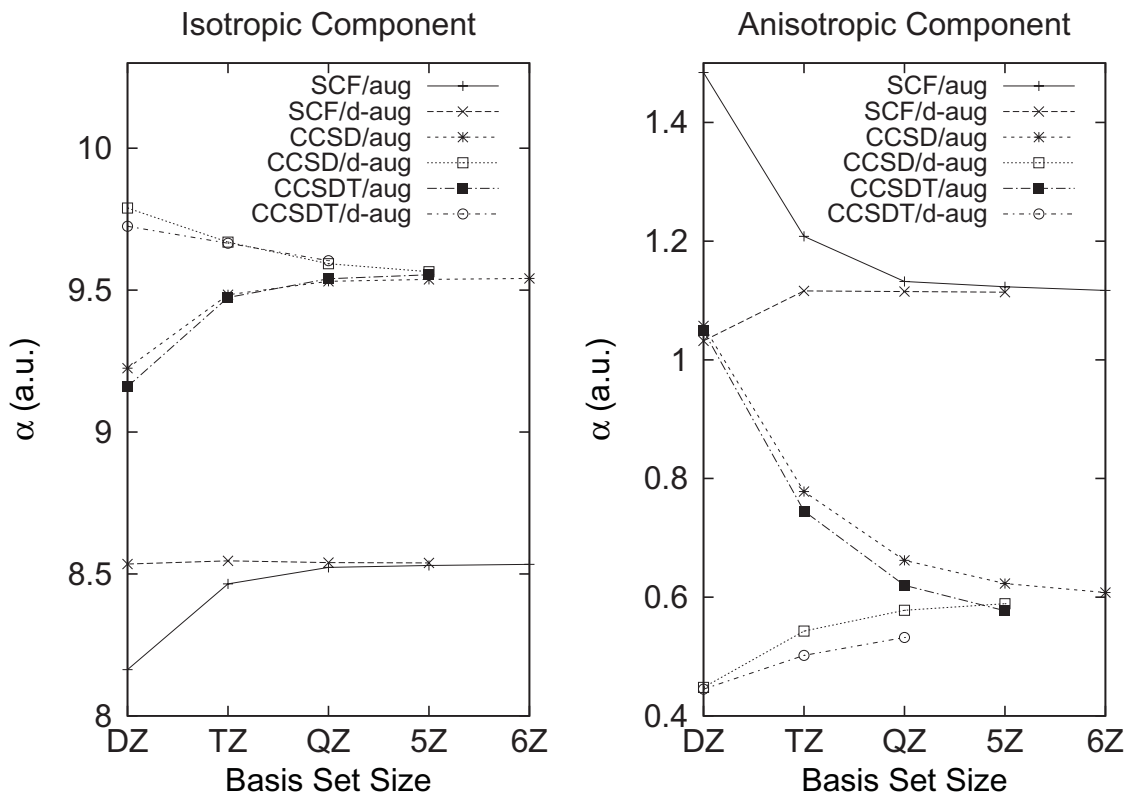
FIG. 3: Evaluation of the Pople family of basis sets for the polarizability of water clusters with the PBE0 functional.
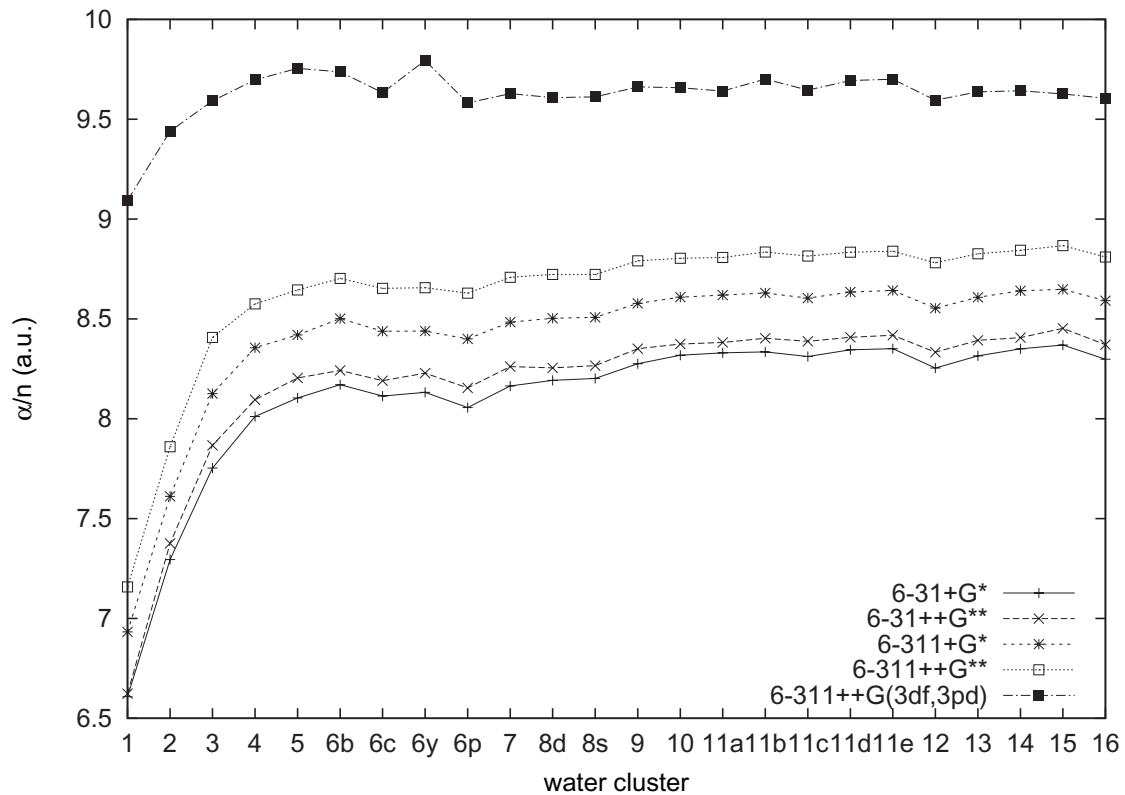
FIG. 4: Evaluation of the Sadlej family of basis sets for the polarizability of water clusters with the PBE0 function.
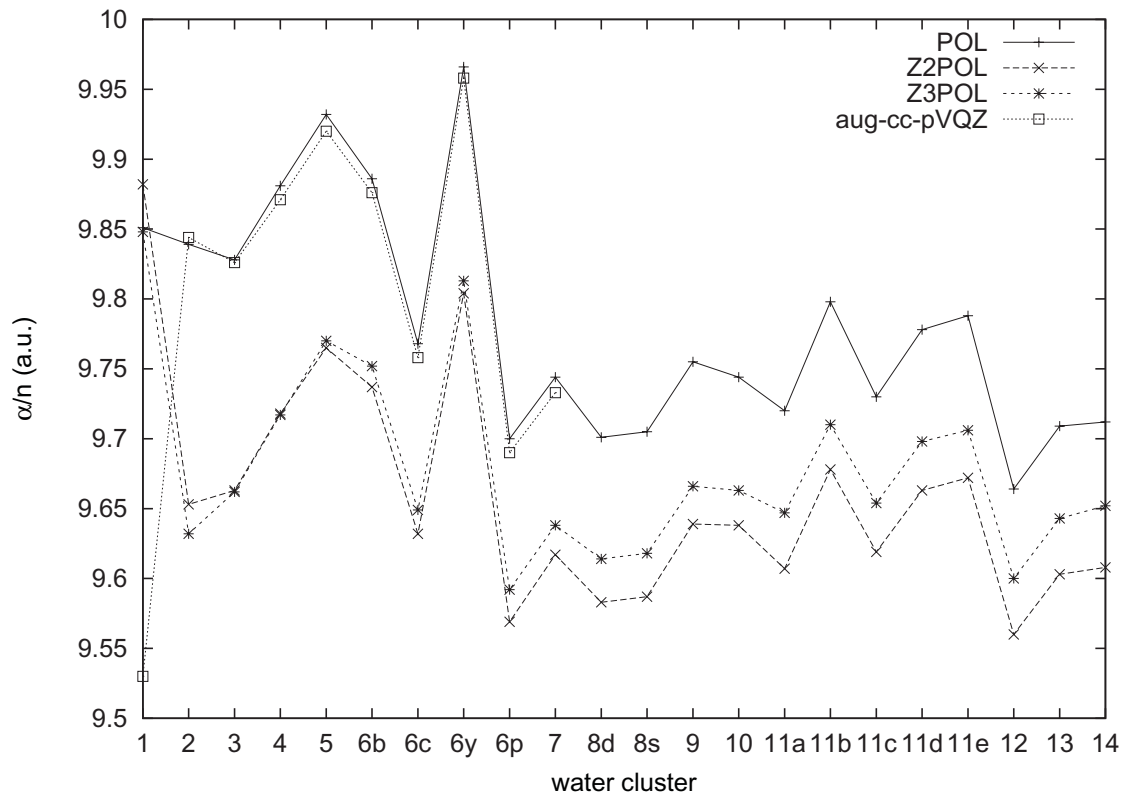
FIG. 5: Evaluation of the Dunning series of basis sets for the polarizability of water clusters with the PBE0 function.
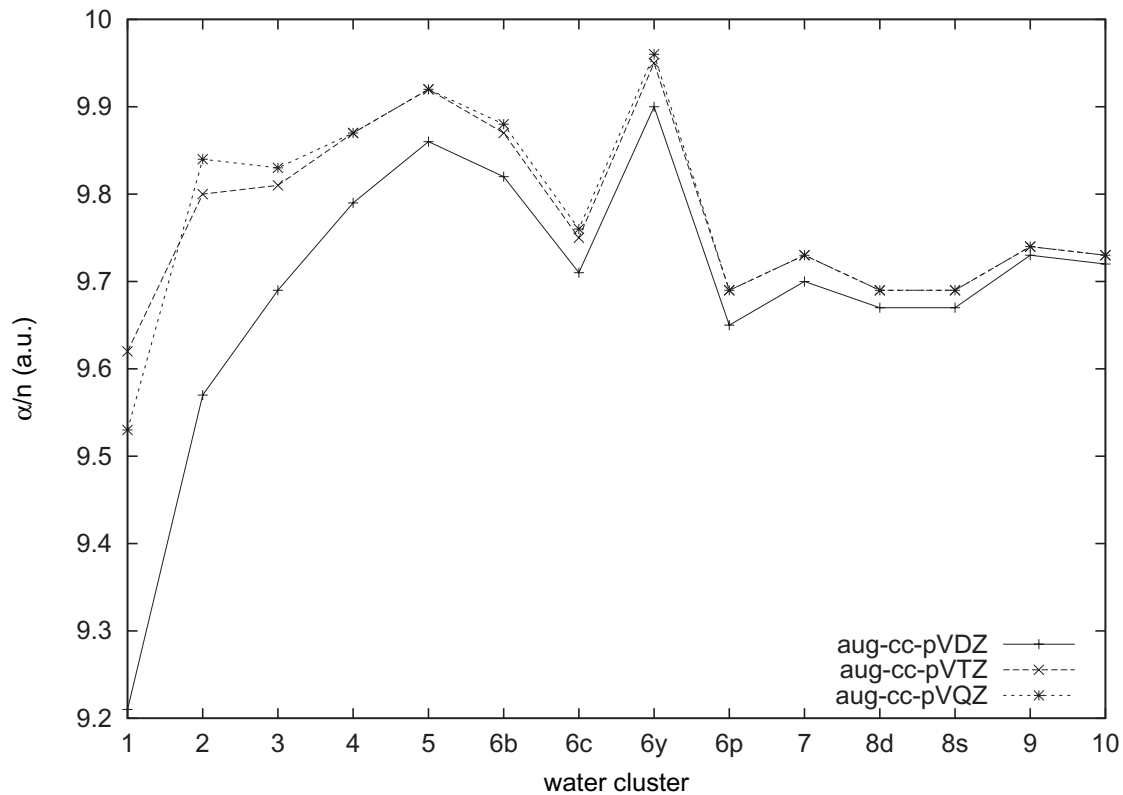
FIG. 6: Isotropic polarizabilities of various isomers of water clusters at the CCSD/aug-cc-pVDZ (frozen core) level of theory.
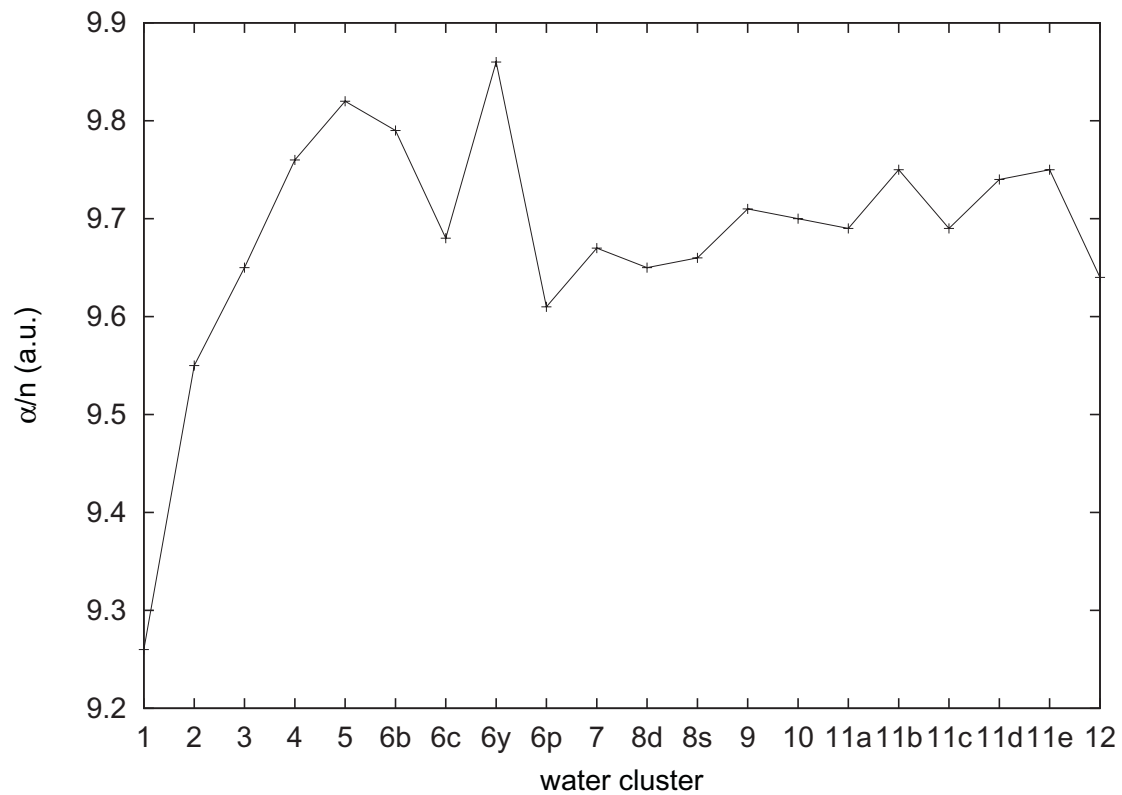
FIG. 7: Errors (with respect to CCSD) in the isotropic polarizability per molecule for a number of density functionals with the aug-cc-pVDZ basis set.
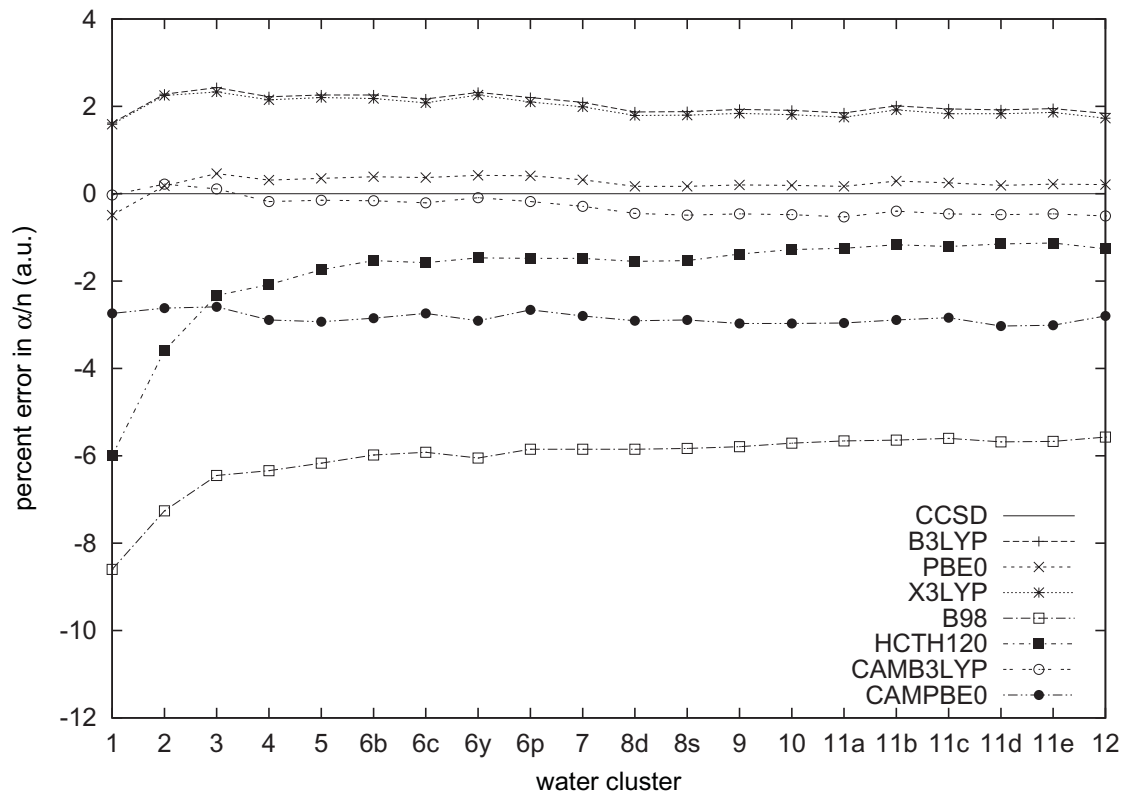
FIG. 8: Errors (with respect to CCSD) in the anisotropic polarizability per molecule for a number of density functionals with the aug-cc-pVDZ basis set.
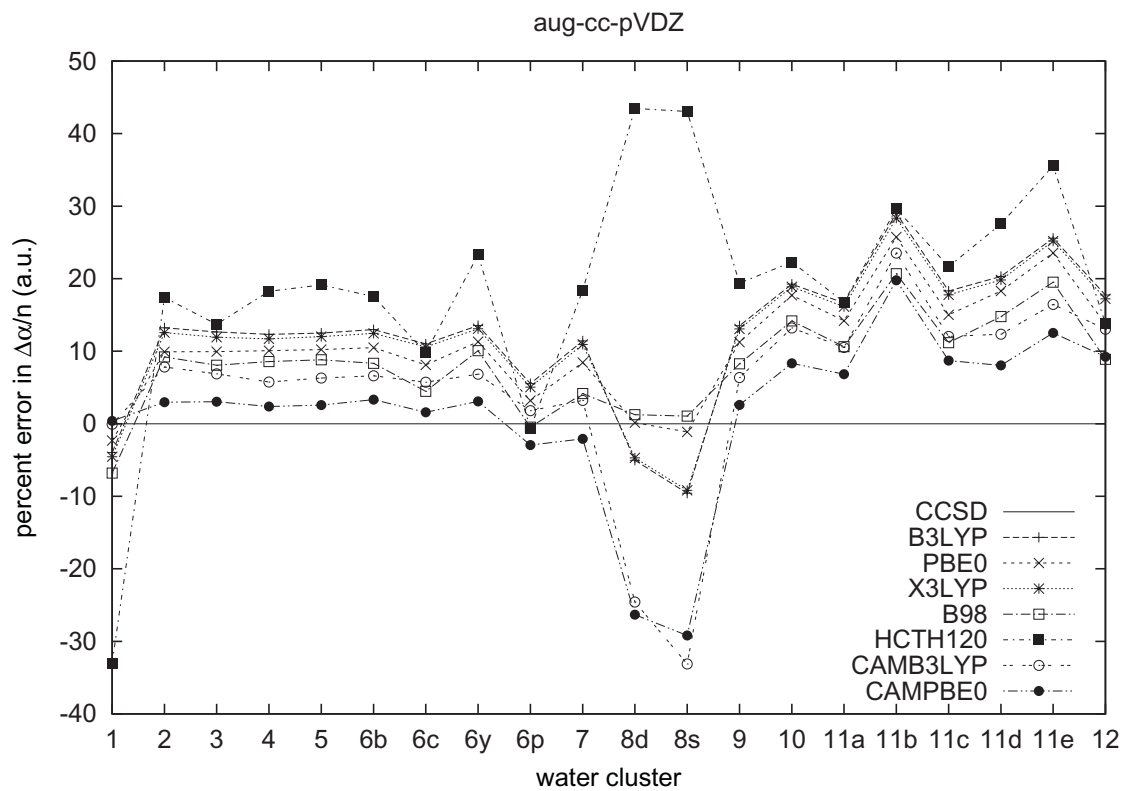
TABLE II: Static dipole polarizabilities of the water monomer and the dimer. The isotropic (α_{iso}) and anisotropic (α_{ani}) components are reported for aug-cc-pVDZ (DZ) and aug-cc-pVTZ (TZ) basis sets using a variety of methods. The MP2 and CCSD(T) results are, by necessity, orbital-relaxed; all other results use the response formalism.

| | Method | α_{iso} | | α_{ani} | |
		DZ	TZ	DZ	TZ
Monomer	SCF	8.163	8.465	1.484	1.208
	CCS	8.596	8.905	1.666	1.375
	MP2	9.190	9.556	0.971	0.638
	CC2	9.764	10.099	0.970	0.634
	CCSD	9.225	9.484	1.057	0.778
	CC3	9.158	9.477	1.048	0.744
	CCSDT	9.160	9.474	1.049	0.745
	CCSDTQ	9.141	-	1.051	-
Dimer	SCF	16.715	17.082	3.089	2.930
	CCS	17.509	17.899	2.984	2.851
	MP2	19.051	-	3.524	-
	CC2	20.197	20.584	3.578	3.390
	CCSD	19.107	19.357	3.471	3.238
	CCSD(T)	18.940	-	3.613	-
	CC3	19.026	-	3.638	-
	CCSDT	19.022	-	3.609	-

TABLE III: Comparison of the water monomer polarizabilities at the SCF, CCSD and CCSDT levels of theory with the Dunning family of basis sets.

Basis	α_{iso}			α_{ani}		
	SCF	CCSD	CCSDT	SCF	CCSD	CCSDT
aug-cc-pVDZ	8.163	9.225	9.160	1.484	1.057	1.049
aug-cc-pVTZ	8.465	9.484	9.474	1.208	0.778	0.745
aug-cc-pVQZ	8.523	9.531	9.540	1.132	0.662	0.620
aug-cc-pV5Z	8.530	9.538	9.554	1.123	0.623	0.577
aug-cc-pV6Z	8.534	9.541	-	1.117	0.608	-
d-aug-cc-pVDZ	8.535	9.789	9.725	1.032	0.448	0.445
d-aug-cc-pVTZ	8.546	9.669	9.665	1.116	0.543	0.502
d-aug-cc-pVQZ	8.540	9.593	9.604	1.115	0.578	0.532
d-aug-cc-pV5Z	8.539	9.565	-	1.114	0.589	-
t-aug-cc-pVDZ	8.533	9.785	9.722	1.033	0.436	0.432
t-aug-cc-pVTZ	8.545	9.672	9.668	1.112	0.541	0.500
t-aug-cc-pVQZ	8.540	9.593	9.605	1.114	0.578	0.531
t-aug-cc-pV5Z	-	9.565	-	-	0.589	-
Experiment [73]	-	9.630	-	-	-	-
Experiment [74]	-	9.92	-	-	-	-

TABLE IV: Basis set dependence of the polarizabilities of small water clusters at the CCSD level of theory.

α (bohr^3)	CCSD/DZ		CCSD/TZ		CCSD/QZ	
n	Iso	Ani	Iso	Ani	Iso	Ani
1	9.26	1.03	9.54	0.71	9.56	0.61
2	19.11	3.47	19.36	3.24	19.31	2.83
3	28.94	5.28	29.02	5.21	28.94	4.87
4	39.06	7.46	38.96	7.51	-	-