

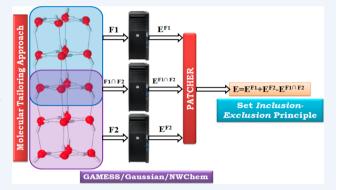
Molecular Tailoring Approach: A Route for ab Initio Treatment of **Large Clusters**

Nityananda Sahu and Shridhar R. Gadre*

Department of Chemistry, Indian Institute of Technology Kanpur, Kanpur 208 016, India

CONSPECTUS: Chemistry on the scale of molecular clusters may be dramatically different from that in the macroscopic bulk. Greater understanding of chemistry in this size regime could greatly influence fields such as materials science and atmospheric and environmental chemistry. Recent advances in experimental techniques and computational resources have led to accurate investigations of the energies and spectral properties of weakly bonded molecular clusters. These have enabled researchers to learn how the physicochemical properties evolve from individual molecules to bulk materials and to understand the growth patterns of clusters.

Experimental techniques such as infrared, microwave, and photoelectron spectroscopy are the most popular and powerful



tools for probing molecular clusters. In general, these experimental techniques do not directly reveal the atomistic details of the clusters but provide data from which the structural details need to be unearthed. Furthermore, the resolution of the spectral properties of energetically close cluster conformers can be prohibitively difficult. Thus, these investigations of molecular aggregates require a combination of experiments and theory. On the theoretical front, researchers have been actively engaged in quantum chemical ab initio calculations as well as simulation-based studies for the last few decades. To obtain reliable results, there is a need to use correlated methods such as Møller-Plesset second order method, coupled cluster theory, or dispersion corrected density functional theory. However, due to nonlinear scaling of these methods, optimizing the geometry of large clusters still remains a formidable quantum chemistry challenge.

Fragment-based methods, such as divide-and-conquer, molecular tailoring approach (MTA), fragment molecular orbitals, and generalized energy-based fragmentation approach, provide alternatives for overcoming the scaling problem for spatially extended molecular systems. Within MTA, a large system is broken down into two or more subsystems that can be readily treated computationally. Finally, the properties of the large system are obtained by patching the corresponding properties of all the subsystems. Due to these approximations, the resulting MTA-based energies carry some error in comparison with calculations based on the full system. An approach for correcting these errors has been attempted by grafting the error at a lower basis set onto a higher basis set. Furthermore, investigating the growth patterns and nucleation processes in clusters is necessary for understanding the structural transitions and the phenomena of magic numbers in cluster chemistry. Therefore, systematic building-up or the introduction of stochastics for generating molecular assemblies is the most crucial step for studying large clusters.

In this Account, we discuss the working principle of MTA for probing molecular clusters at ab initio level followed by a brief summary of an automated and electrostatics-guided algorithm for building molecular assemblies. The molecular aggregates presented here as test cases are generated based on either an electrostatic criterion or the basin hopping method. At MP2 level computation, the errors in MTA-based grafted energies are typically reduced to a submillihartree level, reflecting the potential of finding accurate energies of molecular clusters much more quickly. In summary, MTA provides a platform for effectively studying large molecular clusters at ab initio level of theory using minimal computer hardware.

1. INTRODUCTION

Weakly bonded molecular clusters, being a connecting link between isolated gas-phase molecules and condensed matter, represent a topic of fundamental interest in contemporary science.^{1,2} Understanding the nature of interactions in the clusters and their growth patterns, stability, and properties is an active area of research in physics, chemistry, and biology.^{3,4} Recent advances in high resolution spectroscopy have triggered a number of experimental measurements on molecular clusters. These experimental measurements in conjunction with ab initio

calculations are immensely useful for exploring the molecular geometries and energetics of clusters. However, the commonly used ab initio methods such as Hartree-Fock (HF), Møller-Plesset second order (MP2) theory and coupled cluster singles

Special Issue: Beyond QM/MM: Fragment Quantum Mechanical Methods

Received: February 24, 2014 Published: May 5, 2014



and doubles with perturbative triples (CCSD(T)) are known to formally scale as $O(N^4)$, $O(N^5)$, and $O(N^7)$, respectively, N being the number of basis functions. This requires large computational resources, namely, memory and CPU time. Figure 1 depicts the wall-clock time for energy evaluation of

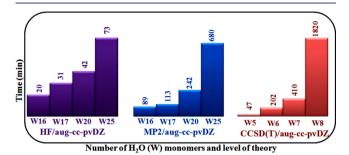


Figure 1. Single point energy calculation time on water (W) clusters of different sizes at HF, MP2, and CCSD(T) levels of theory.

water clusters, W_n, at these levels of theory using the aug-cc-pVDZ basis set on an Intel Core i7 (CPU @ 2.93 GHz) based machine with 32 GB RAM using the Gaussian package.⁵ Thus, the nonlinear scaling remains the main bottleneck that restricts the applicability of these methods to large systems.

The fragmentation-based methods offer an attractive alternative to probe large molecular systems⁶ as may be seen from the works of several research groups and references therein. Christoffersen and co-workers performed HF calculations, almost 40 years' ago, on biological molecules employing one Gaussian function per electron pair. Later, a divide-and-conquer (DC)-based algorithm was introduced by Yang⁸ for treating large molecules by computing the electron density of the subsystems within the Kohn-Sham scheme. Due to their applicability at either HF with subminimal basis sets or density functional theory (DFT), these two approaches were not adequate for large molecular systems. Molecular tailoring approach (MTA), a fragmentation method, was proposed by Gadre and co-workers during the early 1990s for estimating the molecular electrostatic potential (MESP) of zeolites. 9 MTA was later extended for the evaluation of molecular electronic energy, its gradients, and Hessian matrix and also for geometry optimization of large molecular systems. 10,11

Some more fragment-based approaches, aiming at computational efficiency, have been devised for treating large molecular systems. These methods include the DC method by Nakai and co-workers, 12 the fragment molecular orbital (FMO) method by Kitaura and co-workers, 13 the generalized energy-based fragmentation (GEBF) approach by Li and co-workers¹⁴ and many others. ^{15,16} All these approaches are justified by Kohn's near-sightedness principle, viz, the chemical environment of an atom is only minimally affected by the atoms lying beyond a certain distance from that atom. 17 A variety of molecular systems including polyene chains, polypeptides, proteins, aromatic oligoamides, water clusters, etc. have been investigated by employing these methods. However, owing to their approximate nature, all these approaches result in loss of accuracy during the calculations and hence require error minimization procedures. In the literature, these errors are minimized by applying many-body expansion procedures. ^{18–21} These include the many overlapping-body expansion due to Raghavachari, 19 electrostatically embedded many-body expansion by Truhlar and co-workers, 20 and the multicentered

integrated QM:QM method due to Tschumper and coworkers. ²¹

During the 1990s, two alternative approaches, IMOMO/ IMOMM and ONIOM, were developed by Morokuma and coworkers^{22–24} for treating large molecular systems, in which the active part of the system is treated at a higher level while the rest of the molecule is treated at a lower level of theory. 22 Thus, for obtaining the total energy of the system, three calculations, one at the higher level for the active system, one at a lower level for the whole system, and one at a lower level for the active system are required.²² On the other hand, ONIOM represents a multilayered approach, integrating with various levels of MO and MM methods for optimizing transition states and equilibrium geometries. ^{23,24} Another breakthrough in theoretical chemistry was achieved when Warshel and Levitt²⁵ introduced a hybrid QM/MM approach for molecular simulations of chemical processes and proteins. Being a combination of QM and MM, this approach showed great promise for describing chemical reactions and electronic processes.²⁶ However, this approach has some limitations. First, it is hard to optimize the QM part relative to the external charges since the QM nuclei collapse on the negatively charged external charges. Some of the atoms in the MM part bear no charge and hence are invisible to the QM part.

Therefore, it is preferable to investigate large molecular systems only at *ab initio* level of theory, assisted by fragmentation-based procedures. Due to the availability of high-end contemporary hardware, such calculations on large molecular clusters are indeed feasible. This Account reports the tools for treating such clusters followed by some case studies on weakly bonded molecular clusters.

2. MOLECULAR TAILORING APPROACH AND CLUSTER BUILDING ALGORITHM

This section briefly discusses the outline of MTA followed by a summary of the electrostatics-guided molecular cluster building algorithm.

2.1. Molecular Tailoring Approach (MTA)

In MTA, $^{9-11}$ a large molecular system under consideration is partitioned into a set of overlapping fragments on which *ab initio* calculations are carried out. The fragmentation may be done automatically or manually. The desired molecular property, P, of the parent system is obtained by patching those of the individual fragments via the set *inclusion–exclusion* principle, 10,11

$$P = \sum P^{F_i} - \sum P^{F_i \cap F_j} + \dots + (-1)^{k-1} \sum P^{F_i \cap F_j \cap \dots F_k} \dots$$
 (1)

Here, P^{F_i} denotes for the molecular property of *i*th fragment, $P^{F_i \cap F_j}$ stands for the molecular property of binary overlap of *i*th and *j*th fragments, and so on.

For assessing the quality of a fragmentation scheme, a distance-based parameter, R-goodness $(R_{\rm g})$ has been introduced in MTA. What does the $R_{\rm g}$ of an atom (A) in a given fragment denote? It is the minimum distance of an atom "A" present in the fragment from the atoms not belonging to that fragment (see Figure 2). If an atom "A" belongs to more than one fragment, then the maximum of the $R_{\rm g}$ values of "A" in different fragments is taken as the atomic $R_{\rm g}$ of "A". The minimum of all the atomic $R_{\rm g}$ values in the given molecular system is an indicator of the quality of a particular fragmentation scheme. Further details of this parameter can

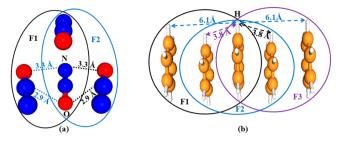


Figure 2. (a) Trimeric fragments for N_2O tetramer. R_g -values (in Å) of "O" and "N" atoms in F1 and F2 are 2.9 and 3.3, respectively. (b) Trimeric fragments for C_6H_6 pentamer. R_g -values (in Å) of the "H" atom in F1, F2, and F3 are 3.8, 6.1, and 3.8, respectively.

be found elsewhere. 10,27 The current version of the MTA enables the evaluation of energy, gradients, Hessian elements, molecular orbital energies, etc., as well as geometry optimization and vibrational frequency evaluation, of large molecular systems. 28–31 Due to the effective parallelization and the ability to work with Gaussian, GAMESS, and NWChem at the back-end, MTA has emerged as a powerful tool for *ab initio* treatment of large molecular systems. Recently, MTA has been employed for MP2 level geometry optimization and vibrational frequency calculation of clusters 18,29–31 of acetylene, benzene, carbon dioxide, and water.

With increasing size of the systems, MTA is expected to be more effective with respect to the corresponding full calculations (FCs). Especially during MP2 level geometry optimization, MTA shows a significant time advantage even with limited hardware. However, due to the approximate nature of MTA, the estimated property is always in error with respect to that obtained by the conventional FCs. Attempts have been made^{27,31} to minimize the error in energy through a grafting procedure as described in the algorithm (see Figure 3), keeping in mind that this error is generally independent of the basis set employed.^{27,31} Besides this, the major contribution to the error in energy comes from the HF level rather than from the second-order perturbative correction, that is, E(2). The method of correcting the MTA energies involves the grafting of the error at lower basis (LB) set onto a desired higher basis (HB) set.^{27,31}

It may be noted that even though the grafting involves a 4fold calculation, the time required for the whole procedure is still less than that for FCs. For instance, the full MP2/aug-ccpVDZ single point calculation of (H₂O)₂₀ (as a small test example) takes a total wall-clock time of 264 min, whereas the corresponding grafting procedure employing 6-31+G* as LB takes 161 min.²⁷ More importantly, this procedure can be extended to any correlated levels of theory within any basis sets. At MP2 level, it has been observed that the errors in MTAgrafted energies are reduced typically to less than ~0.5 mH. This can be illustrated through two test examples. At MP2/augcc-pVDZ (LB, 6-31+G*) level of theory, the error in the single point grafted energies for (C₂H₂)₁₀ and (CO₂)₁₃ clusters are reduced to 0.36 and 0.42 mH compared with the ungrafted values of 17 and 7 mH, respectively. 27 Similar results are also obtained when grafting is done by using 6-31+G* or aug-ccpVDZ as LB and aug-cc-pVTZ as HB. Thus, MTA in conjunction with grafting procedure has a potential to accurately mimic the FC values on minimal hardware.

2.2. Molecular Cluster Building Algorithm

Although dispersion interactions may be more dominant than the electrostatic ones in some systems, the latter is highly orientation dependent. Hence MESP is known to be effective for predicting patterns of binding in molecular clusters. 32,33 In recent work, Yeole and Gadre³⁴ have developed an automated code for systematically building up molecular clusters. The addition of a monomer to an *n*-mer cluster is performed based on the minimization of the energy employing the electrostatic potential for intermolecular complexation (EPIC) model.³³ Three minimum EPIC-energy structures of acetylene pentamer obtained by addition of monomer to the acetylene tetramer are shown as an illustration in Figure 4 and may be subjected to further ab initio level geometry optimization. This algorithm has been tested out and benchmarked^{29,30,34} on a variety molecular clusters of C₂H₂, C₆H₆, H₂O, ZnS, and CO₂. However, such a systematic method has a limitation since it cannot be used for systems that exhibit extra stability due to a sudden change in the structure with increasing size. Therefore, the combination of techniques such as basin hopping or genetic algorithm, with the quantum chemical methods is immensely useful for introducing stochastics in clusters. Some illustrative examples

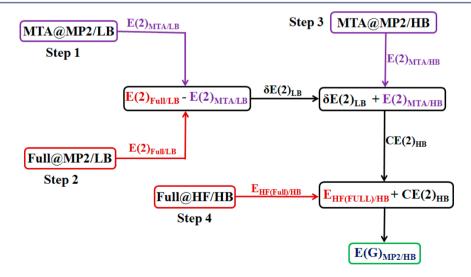


Figure 3. Grafting procedure at MP2 level of theory. $E_{\text{MP2}} = E_{\text{HF}} + E(2)$; $\delta E(2)_{\text{LB}}$, E(2) correction at LB; $\text{CE}(2)_{\text{HB}}$, corrected E(2) at HB; E(G), grafted energy.

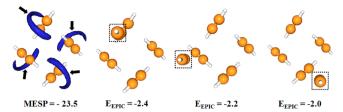


Figure 4. MESP isosurface (kcal/mol) of $(C_2H_2)_4$ at MP2/6-31+G* level of theory. Three minimum $(C_2H_2)_5$ structures with EPIC-energy (kcal/mol).

of application of MTA to weakly bound large molecular clusters are presented below, with starting geometries generated either by stochastics-based or systematic methods.

3. TEST STUDIES ON MOLECULAR CLUSTERS

Some test examples of MTA-based calculations on large clusters of water, carbon dioxide, hydrated sodium ion, and solvated ion pairs are given below.

3.1. Water Clusters

For understanding the nature of water clusters and identifying the hydrogen bond networks, several Monte Carlo (MC)- and MD-based studies, as well as *ab initio* calculations, have been performed. The clusters are described as a subject of theory is now feasible only with large parallel architectures. Solutions the accuracy and feasibility, the fragmentation-based approaches coupled with high-level correlated methods offer a cost-effective solution for such studies. The MC-based techniques are of paramount use for the introduction of stochastics in generating starting geometries whereas fragmentation methods are the important follow-up tools for *ab initio* studies of clusters.

In an extensive study,³¹ six low lying isomers each of $(H_2O)_{20}$ and $(H_2O)_{25}$ clusters are investigated employing a combination of MC and MTA. The geometries generated through the temperature basin paving (TBP) method³⁷ are used as starting structures for MTA-based geometry optimization at MP2/aug-cc-pVDZ level of theory. Three

low-lying isomers for each cluster³¹ are displayed in Figure 5. The edge-sharing pentagonal prism $(H_2O)_{20}$ -I is found to be the most stable isomer, in agreement with the previously reported global minimum,³⁵ whereas an internally solvated (19 + 1) type structure, $(H_2O)_{20}$ -II, not hitherto noticed in the literature is reported in ref 31. Similarly, a new global minimum, namely, $(H_2O)_{25}$ -I, different from the one reported earlier, is found.³¹ It comprises a $(H_2O)_{21}$ cluster with a laterally placed cyclic tetramer, whereas the other isomers, $(H_2O)_{25}$ -II and $(H_2O)_{25}$ -III, are characterized by a complete $(H_2O)_{24}$ cage made of cyclic tetramers, pentamers, and hexamers with a centrally placed water monomer (see Figure 5). The detailed energetics of all these isomers is reproduced at the MTA-optimized geometries from ref 31 by employing a better set of fragments and is displayed in Table 1.

Table 1. MTA Energies without Grafting, E(MTA), MTA-Grafted Energies, E(G), and Full-Calculated Energies E(FC), in au, for Three Isomers of $(H_2O)_{20}$ and $(H_2O)_{25}$ at MP2 Level Using 6-31+G* and aug-cc-pVDZ as LB and HB, respectively

system	$E(\text{MTA})_{\text{MP2/HB}}$	$E(G)_{MP2/HB}$	$E(FC)_{MP2/HB}$	ε^a (mH)		
$(H_2O)_{20}$ -I	-1525.594 16	-1525.574 96	-1525.575 26	0.30		
$(H_2O)_{20}$ -II	-1525.553 83	-1525.572 62	-1525.573 01	0.39		
$(H_2O)_{20}$ -III	$-1525.570\ 30$	-1525.57242	-1525.57279	0.36		
$(H_2O)_{25}$ -I	-1906.97672	-1906.98245	$-1906.982\ 80$	0.35		
$(H_2O)_{25}$ -II	-1906.97355	-1906.981 26	-1906.981 46	0.20		
$(H_2O)_{20}$ -III	-1906.97347	-1906.981 06	-1906.981 30	0.24		
$^{a}\varepsilon = E(G)_{MP2/HB} - E(FC)_{MP2/HB} .$						

As discussed earlier, MTA energies are expected to carry errors in relation to their FC counterparts. These errors are minimized employing the grafting procedure described earlier, using aug-cc-pVDZ and 6-31+G* as HB and LB, respectively. From Table 1, it can be noticed that the MTA-based estimated energies are in error by 2 to 20 mH with respect to their FC counterparts. On grafting, the error reduces to less than 0.4 mH for all the isomers in Table 1. Thus, MTA in conjunction with

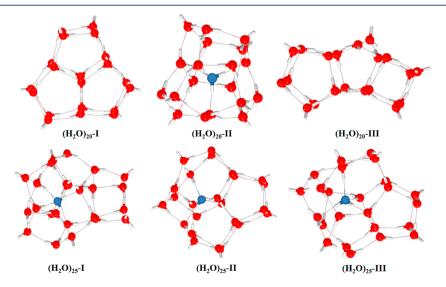


Figure 5. MTA-optimized $(H_2O)_{20}$ and $(H_2O)_{25}$ clusters at MP2/aug-cc-pVDZ level of theory. Internally solvated water molecule is marked in blue.³¹

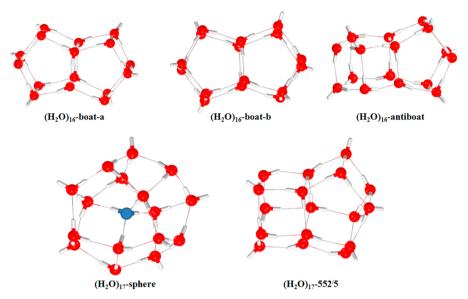


Figure 6. Three isomers of $(H_2O)_{16}$ and two isomers of $(H_2O)_{17}$ clusters at MP2/aug-cc-pVDZ level of theory. Internally solvated water molecule is marked in blue.³⁶

Table 2. MTA Energies without Grafting $(E(\text{MTA})_{\text{CCSD}(T)})$, Difference in MTA and Full-Calculated Energies (δ_{MP2}) at MP2/aug-cc-pVDZ level, and Grafted Energies $(E(G)_{\text{CCSD}(T)})$ and Full-Calculated Energies $(E(FC)_{\text{CCSD}(T)})$ at CCSD(T)/aug-cc-pVDZ for $(H_2O)_{16}$ and $(H_2O)_{17}$ Clusters^a

system	$E(MTA)_{CCSD(T)}$	$\delta_{ ext{MP2}}{}^{b}$	$E(G)_{CCSD(T)}$	$E(FC)_{CCSD(T)}^{c}$	
$(H_2O)_{16}$ -boat-a	-1220.655 90	-0.003 35	-1220.659 25	-1220.659 22	
$(H_2O)_{16}$ -boat-b	-1220.656 87	-0.002 11	-1220.658 98	-1220.658 99	
$(H_2O)_{16}$ -antiboat	-1220.653 47	-0.005 47	-1220.658 94	-1220.65893	
$(H_2O)_{17}$ -sphere	-1296.938 44	-0.015 11	-1296.953 55	-1296.953 62	
$(H_2O)_{17}-552^{/}5$	-1296.938 10	-0.013 47	-1296.951 57	-1296.951 53	
^a Energies in au. ${}^b\delta_{\text{MP2}} = E(\text{FC})_{\text{MP2}} - E(\text{MTA})_{\text{MP2}}$. ^c Reference36.					

grafting procedure is seen to be a powerful tool for studying large clusters employing correlated theory.

Following the success of the methodology at MP2 level, it is further extended to CCSD(T) level of theory. Unlike the above procedure, the extrapolation is done within one basis set but with two levels of theory, namely, MP2 and CCSD(T). Since the electronic correlation energy is found to be almost constant beyond the post-HF methods, MP2 (formally scales as $O(N^5)$) is expected to be a suitable starting point for approximating the highly accurate CCSD(T) energies as stated in eq 2.

$$E(G)_{CCSD(T)/DZ} = E(MTA)_{CCSD(T)/DZ} + E(full)_{MP2/DZ}$$
$$- E(MTA)_{MP2/DZ}...$$
(2)

For this purpose, the MP2/aug-cc-pVTZ level optimized geometries from the Supporting Information file of ref 36 are reoptimized at MP2/aug-cc-pVDZ level of theory, followed by an MTA-based single point energy calculation at CCSD(T)/aug-cc-pVDZ level of theory. The energetics of three isomers of $(H_2O)_{16}$ and two isomers of $(H_2O)_{17}$ clusters (see Figure 6) at CCSD(T)/aug-cc-pVDZ level of theory are reported in Table 2 along with the corresponding FC values. It is remarkable that the grafted MTA energies at CCSD(T)/aug-cc-pVDZ level match the corresponding FC values reported by Xantheas almost up to the fourth decimal. ³⁶ Since the choice of the basis set is no bar, it would be further possible to attain the "gold standard" of contemporary quantum chemistry, that is, CCSD(T)/CBS. Thus, the results of the present study pave a

way for taking up investigations on water clusters containing 100 monomers at the *ab initio* level of theory in the near future.

3.2. Carbon Dioxide Clusters

Carbon dioxide (CO_2) clusters rank second only to water clusters in terms of their interest and significance. They have been a subject of intense experimental and theoretical investigations, owing to the role of CO_2 in fundamental processes such as respiration and photosynthesis. In recent years, advances in high-resolution spectroscopic techniques, as well as powerful computers, in conjunction with dispersion-corrected density functionals have accelerated the investigations on these clusters. Some recent works employ an MD-based method for geometry optimization of $(CO_2)_n$ for $4 \le n \le 40$ by Takeuchi, high resolution spectroscopic studies of $(CO_2)_n$ for $6 \le n \le 13$ by Oliaee et al., 40,41 and MTA-based geometry optimization of the latter clusters at MP2/aug-cc-pVDZ level of theory.

It is indeed difficult to capture the large number of local minima on the PES, which necessitates a method for systematically growing these clusters. As discussed earlier, electrostatics is known to capture the directionality of the approach of an ensuing monomer, which is to be added to a cluster. $(CO_2)_n$ clusters in the range of n = 6-13 have been generated by adding CO_2 monomers or dimers to the existent lower-sized clusters through an automated cluster building algorithm, ³⁴ followed by the estimation of MP2/CBS binding energy and vibrational frequency calculation at MP2/aug-cc-pVDZ level of theory within MTA. ³⁰ The low-lying isomers of

clusters for n = 6-10 resemble the geometries derived from the MP2 level geometry optimization of the structures reported by Takeuchi.³⁹ An earlier study³⁰ reported the newly generated structures for higher-sized clusters (11, 12, and 13) to be lower in energy than those reported earlier.³⁹ However, upon energy correction through grafting procedure, the most-stable structures (see Figure 7) are in good qualitative agreement with those reported by Takeuchi.³⁹

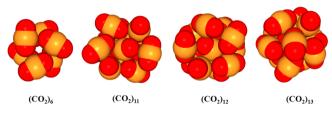


Figure 7. MTA-based lowest-energy isomers of $(CO_2)_n$ clusters, n = 6 and 11-13, at MP2/aug-cc-pVDZ level of theory.

The presence of long-range dispersion interactions and the C···O electrostatic interactions lead to the T- or slanted T-type orientations in the clusters, the number of the former ones increasing with cluster size. The structures in the range n=6-13 turn out to be asymmetric tops (A > B > C) except for N=6 and 13, which are symmetric tops (see Figure 7). Although the geometry optimization was unconstrained, the geometries of the lowest energy isomers of $(CO_2)_6$ and $(CO_2)_{13}$ did converge to symmetric tops, in agreement with the experiments. The symmetrical nature of these isomers may be responsible for the extra stability of the clusters. Figure 8

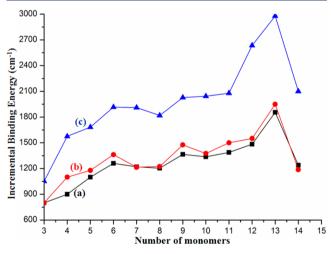


Figure 8. Incremental binding energy (cm^{-1}) of $(CO_2)_n$ clusters for n = 3-14, obtained using (a) M-O-M potential (black squares), ⁴² (b) SAPT-s (red circles), ⁴³ and (c) *ab initio* calculations at MP2/aug-cc-pVDZ level (blue triangles).

depicts the incremental binding energy, that is, the energy released when $(CO_2)_n$ is formed by addition of a CO_2 monomer to $(CO_2)_{n-1}$ clusters. These data exhibit local maxima for N=6 and 13 at MP2/aug-cc-pVDZ level of theory, in fair agreement with the estimated values employing M-O-M empirical intermolecular potentials⁴² and SAPT-s calculations. This energy is again found to decrease for $(CO_2)_{14}$ at all these levels of calculation. The minimal nature of the isomers is confirmed by the absence of imaginary frequencies obtained from the MTA-based vibrational calcu-

lations at MP2/aug-cc-pVDZ level of theory. The scaling factor is chosen to be 0.985 because the experimental C–O asymmetric stretch for the CO₂ molecule is observed at 2350.8 cm⁻¹ against the theoretically calculated value of 2386.6 cm⁻¹. These C–O asymmetric stretch frequencies generally reveal a blue shift with increasing cluster size, in agreement with the experimental results. He experimentally observed blue shift on going from (CO₂)₆ to (CO₂)₁₃ is approximately 14.2 cm⁻¹ compared with 17.7 cm⁻¹ (see Table 3) obtained from MTA-based MP2 level calculations. Besides this, the spectral pattern of the isomers of all clusters is found to resemble the respective experimental one.

Table 3. Comparison of MTA-Based Scaled Vibrational Frequency for Asymmetric C–O Stretch of $(CO_2)_n$ Clusters, n = 6-13, at MP2/aug-cc-pVDZ Level of Theory with the Experimental Values

cluster size	calculated frequency (cm ⁻¹)	experimental frequency a (cm $^{-1}$)		
6	2351.1	2353.5		
7	2353.5	2356.2		
8	2357.6			
9	2360.1	2358.7		
10	2360.3	2361.1		
11	2363.9	2361.8		
12	2366.3	2364.3		
13	2368.8	2367.7		
^a References 40 and 41.				

The $(CO_2)_{13}$ cluster, being a magic one,³⁹ is expected to retain its core upon increasing the size further by addition of a few monomers. In order to test out the cluster building algorithm,³⁴ a demonstrative calculation on $(CO_2)_{14}$ is carried out at MP2 level. Out of several possible geometries scanned at MP2/aug-cc-pvDZ level of theory, the most stable $(CO_2)_{14}$ isomer is indeed found to have the $(CO_2)_{13}$ core intact with a monomer situated at the periphery (see Figure 9). At MP2 level, the rotational constant values (in cm⁻¹), A = 0.0031, B = 0.0026, and C = 0.0026, of this isomer resemble those, A = 0.0026

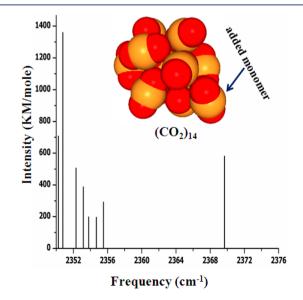


Figure 9. MTA-optimized geometry and vibrational spectrum for the low-lying $(CO_2)_{14}$ isomer at MP2/aug-cc-pVDZ level of theory.

0.0031, B = 0.0026, and C = 0.0025, of the isomer obtained from the geometry optimization of the structure provided by Takeuchi.³⁹ The geometry and the MTA-based vibrational spectrum for this isomer is displayed in Figure 9. Thus, the combined use of the cluster builder algorithm and MTA enables the putative minima search as well as calculation of vibrational spectra clusters. The results for larger CO_2 clusters would be beneficial for benchmarking the future experimental investigations on these clusters.

3.3. Sodium Cation in Water Clusters

Small microsolvated clusters of ions have attracted considerable interest because of their importance in physicochemical and biological domains. 44,45 Despite the advancements in experimental techniques, the structure of solvated ions is not yet well-understood due to the variation of the number of water molecules (varying from 4 to 8) in the first solvation shell of a common ion such as sodium. 46 For understanding relative stability of Na+ ion in different coordinated states and for a comparison of the results with experimental data, several theoretical calculations have been reported. We present a test example below using the recent B3LYP-level geometry optimization results of Sastry and co-workers.⁴⁷ In order to assess the feasibility and to benchmark the grafting procedure for charged systems, MTA-based geometry optimization of three low-lying isomers of Na⁺(H₂O)₁₈ cluster is performed at MP2/aug-cc-pVDZ level of theory (see Figure 10). For

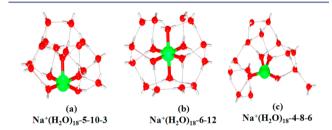


Figure 10. Three isomers of $\mathrm{Na^+(H_2O)_{18}}$ clusters at MP2/aug-cc-pVDZ level of theory. Sodium atoms are marked in green.

maintaining the charge and spin multiplicity throughout, the fragmentation on these clusters is done by including the cation in every fragment. MTA-based grafted energies of these isomers (Figure 10a-c) are -1534.783 39, -1534.765 90, and -1534.753 29 au as against their FC values of -1534.783 29, -1534.765 74, and -1534.753 49 au, respectively. This shows that the grafted energies fall within 0.2 mH of their FC counterparts, reflecting the excellent performance of grafting procedure for accurate energetics, with the stability of isomers in agreement with that at B3LYP level. Thus, looking at the success of the grafting procedure for estimating the energetics of charged species inside water cages, it is evident that high-level calculations are feasible for addressing solvation problems, even with a minimal hardware.

3.4. Sodium Chloride Ion Pair in Acetonitrile

The stability and structural properties of contact ion pairs (CIPs) and solvent-separated ion pairs (SSIPs) in nonaqueous medium, for example, acetonitrile (CH₃CN), is also a topic of current interest. ⁴⁸ For a further demonstration of the capability of MTA to handle large-sized clusters, a test calculation of sodium and chloride SSIP encapsulated in (CH₃CN)₉₀ clusters is attempted at M06-2X/6-31+G*//M06-2X/3-21G level of theory. The resulting geometry of the SSIP isomer (may not be

a minimum) is displayed in Figure 11. The initial structure was generated by the group of Professor Tembe at I.I.T. Bombay,

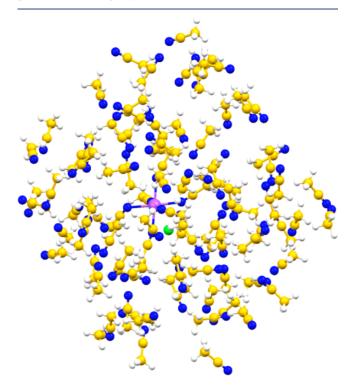


Figure 11. Na $^+$ Cl $^-$ (CH $_3$ CN) $_{90}$ SSIP (Na $^+$ Cl = 7.6 Å) optimized at M06-2X/3-21G level of theory. Sodium and chlorine atoms are marked in green and magenta, respectively.

using the GROMACS 4.5.4 software package⁴⁹ and the united atom optimized potential for liquid simulations (OPLS) force field. So Ab initio geometry optimization of Na⁺Cl⁻(CH₃CN)₉₀ is difficult even at a low level of theory. A computationally feasible fragmentation scheme of a total of eight main fragments with an average size of 160 atoms is employed for the single point energy (E = -12566.07553 au) calculation at M06-2X/6-31+G* (total number of basis functions for the whole system is 5716) level of theory. The test example is included only to bring out the feasibility of use of MTA for carrying out such a computationally expensive calculation. No attempt has been made to explore energy minima on the PES.

4. CONCLUSIONS

High-level correlated methods are necessary for reliable theoretical investigations of weakly bonded molecular clusters. Due to the nonlinear scaling and the requirement large memory and disk space coupled with the presence of several local energy minima, such calculations are not always feasible using the standard packages. On the other hand, such calculations are rendered feasible though the QM/MM techniques. However, these still face some difficulties in accuracy, owing to the parametrization involved therein. The fragmentation-based approaches offer an attractive solution for investigating large molecular clusters even with limited computational power offered by off-the-shelf hardware. As demonstrated in the current study, molecular tailoring approach (MTA) enables such calculations on rather modest hardware. In continuation with the parallel software developed in our group, 10,51,52 MTA turns out to be a powerful tool for geometry optimization and calculation of vibrational spectra of a variety of molecular

clusters. This Account has illustrated some such studies on water, carbon dioxide clusters, and hydrated ions. The errors resulting due to the approximate nature of MTA can be reduced through the grafting procedure or many-body corrections. Furthermore, for understanding the structural transitions and the unusual stability of clusters, either systematic building or the introduction of stochastics is necessary in cluster chemistry. Thus, the combination of these methods with MTA can open up a possibility of investigating large-sized clusters even at the present gold standard of quantum chemistry, that is, CCSD(T)/CBS, in the years to come.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gadre@iitk.ac.in.

Notes

The authors declare no competing financial interest.

Biographies

Nityananda Sahu received his B.Sc. (2008) and M.Sc. (2010) degrees from Utkal University, Odisha, India. He is currently pursuing a Ph.D. at IIT Kanpur on quantum chemical investigations of large water clusters and explicit hydration of molecules.

Shridhar R. Gadre received his Ph.D. degree from IIT Kanpur in 1979. After postdoctoral research with Professors Robert G. Parr and Robert L. Matcha in USA, he served the University of Pune as a Lecturer (1980–1988) and Professor (1988–2010). He joined IIT Kanpur as a Professor in 2010. His research areas include topography of molecular scalar fields, information theory in quantum chemistry, and development of parallel *ab initio* codes for large molecules.

ACKNOWLEDGMENTS

N.S. is thankful to the UGC, New Delhi, for a Fellowship. S.R.G. acknowledges financial support from DST, New Delhi [projects (SR/S2-/JCB/41/2006) and (SR/S1/PC-37/2011)] and C-DAC, Pune. Thanks are due to Dr. G. N. Sastry and Professor B. L. Tembe for providing the starting geometries of some clusters studied in this work and Dr. Sachin D. Yeole for cluster building code.

■ REFERENCES

- (1) Braga, D.; Grepioni, F. From Molecule to Molecular Aggregation: Clusters and Crystals of Clusters. *Acc. Chem. Res.* **1994**, *27*, 51–56.
- (2) Fehliner, T. P.; Halet, J.-F.; Saillard, J.-Y. *Molecular Clusters: A Bridge to Solid-State Chemistry*; Cambridge University Press: Cambridge, U.K., 2007, pp 378.
- (3) Chalasinski, G.; Szczesniak, M. M. Origins of Structure and Energetics of van der Waals Clusters from *ab initio* Calculations. *Chem. Rev.* **1994**, *94*, 1723–1765.
- (4) Desfrançois, C.; Carles, S.; Schermann, J. P. Weakly Bound Clusters of Biological Interest. *Chem. Rev.* **2000**, *100*, 3943–3962.
- (5) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.;

- Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
- (6) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.* **2012**, *112*, 632–672.
- (7) Spangler, D.; Christoffersen, R. E. Ab Initio Calculations on Large Molecules Using Molecular Fragments. Generalization of the Molecular Fragment Basis at the Minimum Basis Set Level. Int. J. Ouantum Chem. 1980, 17, 1075–1097.
- (8) Lee, C.; Yang, W. The Divide-and-Conquer Density-Functional Approach: Molecular Internal Rotation and Density of States. *J. Chem. Phys.* **1992**, *96*, 2408–2411.
- (9) Gadre, S. R.; Shirsat, R. N.; Limaye, A. C. Molecular Tailoring Approach for Simulation of Electrostatic Properties. *J. Phys. Chem.* **1994**, 98, 9165–9169.
- (10) Ganesh, V.; Dongare, R. K.; Balanarayan, P.; Gadre, S. R. Molecular Tailoring Approach for Geometry Optimization of Large Molecules: Energy Evaluation and Parallelization Strategies. *J. Chem. Phys.* **2006**, *125*, No. 104109.
- (11) Gadre, S. R.; Ganesh, V. Molecular Tailoring Approach: Towards PC-based *ab Initio* Treatment of Large Molecules. *J. Theor. Comput. Chem.* **2006**, *5*, 835–855.
- (12) Kobayashi, M.; Nakai, H. Dual-Level Hierarchical Scheme for Linear-Scaling Divide-and-Conquer Correlation Theory. *Int. J. Quantum Chem.* **2009**, 109, 2227–2237.
- (13) Fedorov, D. G.; Kitaura, K., Eds. The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems; CRC Press: Boca Raton, FL, 2009.
- (14) Hua, W.; Fang, T.; Li, W.; Yu, J.-G.; Li, S. Geometry Optimizations and Vibrational Spectra of Large Molecules from a Generalized Energy-Based Fragmentation Approach. *J. Phys. Chem. A* **2008**, *112*, 10864–10872.
- (15) Bettens, R. P. A.; Lee, A. M. A New Algorithm for Molecular Fragmentation in Quantum Chemical Calculations. *J. Phys. Chem. A* **2006**, *110*, 8777–8785.
- (16) Collins, M. A. Molecular Potential Energy Surfaces Constructed from Interpolation of Systematic Fragment Surfaces. *J. Chem. Phys.* **2007**, *127*, No. 024104.
- (17) Kohn, W. Density Functional and Density Matrix Method Scaling Linearly with the Number of Atoms. *Phys. Rev. Lett.* **1996**, *76*, 3168–3171.
- (18) Mahadevi, A. S.; Rahalkar, A. P.; Gadre, S. R.; Sastry, G. N. *Ab Initio* Investigation of Benzene Clusters: Molecular Tailoring Approach. *J. Chem. Phys.* **2010**, *133*, No. 164308.
- (19) Mayhall, N. J.; Raghavachari, K. Many-Overlapping-Body (MOB) Expansion: A Generalized Many-Body Expansion for Nondisjoint Monomers in Molecular Fragmentation Calculations of Covalent Molecules. J. Chem. Theory Comput. 2012, 8, 2669–2675.
- (20) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Correlation Energy, with Applications to the Calculation of Accurate Second-Order Møller—Plesset Perturbation Theory Energies for Large Water Clusters. J. Chem. Theory Comput. 2007, 3, 1342—1348
- (21) Bates, D. M.; Smith, J. R.; Janowski, T.; Tschumper, G. S. Development of a 3-Body:Many-Body Integrated Fragmentation Method for Weakly Bound Clusters and Application to Water Clusters $(H_2O)_{n=3-10,16,17}$. J. Chem. Phys. **2011**, 135, No. 044123.
- (22) Maseras, F.; Morokuma, K. IMOMM: A New Integrated *Ab Initio* + Molecular Mechanics Geometry Optimization Scheme of Equilibrium Structures and Transition States. *J. Comput. Chem.* **1995**, *16*, 1170–1179.
- (23) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels-Alder Reactions and Pt(P(t-Bu)₃)₂ + H₂ Oxidative Addition. *J. Phys. Chem.* **1996**, *100*, 19357–19363.

- (24) Vreven, T.; Mennucci, B.; da Silva, C. O.; Morokuma, K.; Tomasi, J. The ONIOM-PCM Method: Combining the Hybrid Molecular Orbital Method and the Polarizable Continuum Model for Solvation. Application to the Geometry and Properties of a Merocyanine in Solution. *J. Chem. Phys.* **2001**, *115*, 62–72.
- (25) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (26) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.
- (27) Sahu, N.; Yeole, S. D.; Gadre, S. R. Appraisal of Molecular Tailoring Approach for Large Clusters. *J. Chem. Phys.* **2013**, *138*, No. 104101
- (28) Rahalkar, A. P.; Gadre, S. R. Tailoring Approach for Obtaining Molecular Orbitals for Large Systems. *J. Chem. Sci.* **2012**, *124*, 149–158.
- (29) Rahalkar, A. P.; Yeole, S. D.; Gadre, S. R. Acetylene Aggregates via Cluster Building Algorithm and Molecular Tailoring Approach. *Theor. Chem. Acc.* **2012**, *131*, 1095–1101.
- (30) Yeole, S. D.; Sahu, N.; Gadre, S. R. Structures, Energetics and Vibrational Spectra of CO₂ Clusters through Molecular Tailoring and Cluster Building Algorithm. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7718–7723.
- (31) Furtado, J. P.; Rahalkar, A. P.; Shanker, S.; Bandyopadhyay, P.; Gadre, S. R. Facilitating Minima Search for Large Water Clusters at MP2 Level via Molecular Tailoring. *J. Phys. Chem. Lett.* **2012**, *3*, 2253–2258
- (32) Gadre, S. R.; Bhadane, P. K. Patterns in Hydrogen Bonding via Electrostatic Potential Topography. *J. Chem. Phys.* **1997**, *107*, 5625–5626.
- (33) Pundlik, S. S.; Gadre, S. R. Structure and Stability of DNA Base Trimers: An Electrostatic Approach. *J. Phys. Chem. B* **1997**, *101*, 9657–9662.
- (34) Yeole, S. D.; Gadre, S. R. Molecular Cluster Building Algorithm: Electrostatic Guidelines and Molecular Tailoring Approach. *J. Chem. Phys.* **2011**, *134*, No. 084111.
- (35) Fanourgakis, G. S.; Aprà, E.; de Jong, W. A.; Xantheas, S. S. High-Level *Ab Initio* Calculations for the Four Low-Lying Families of Minima of (H₂O)₂₀: II. Spectroscopic Signatures of the Dodecahedron, Fused Cubes, Face-Sharing Pentagonal Prisms and Edge-Sharing Pentagonal Prisms Hydrogen Bonding Networks. *J. Chem. Phys.* **2005**, *122*, No. 134304.
- (36) Yoo, S.; Aprà, E.; Zeng, X. C.; Xantheas, S. S. High-Level *Ab Initio* Electronic Structure Calculations of Water Clusters $(H_2O)_{16}$ and $(H_2O)_{17}$: A New Global Minimum for $(H_2O)_{16}$. *J. Phys. Chem. Lett.* **2010**, *1*, 3122–3127.
- (37) Shanker, S.; Bandyopadhyay, P. Monte Carlo Temperature Basin Paving with Effective Fragment Potential: An Efficient and Fast Method for Finding Low-Energy Structures of Water Clusters (H₂O)₂₀ and (H₂O)₂₅. *J. Phys. Chem. A* **2011**, *115*, 11866–11875.
- (38) Ramabhadran, R. O.; Raghavachari, K. Extrapolation to the Gold-Standard in Quantum Chemistry: Computationally Efficient and Accurate CCSD(T) Energies for Large Molecules Using an Automated Thermochemical Hierarchy. *J. Chem. Theory Comput.* **2013**, *9*, 3986–3994.
- (39) Takeuchi, H. Geometry Optimization of Carbon Dioxide Clusters $(CO_2)_n$ for $4 \le n \le 40$. *J. Phys. Chem. A* **2008**, *112*, 7492–7497.
- (40) Oliaee, J. N.; Dehghany, M.; Moazzen-Ahmadi, N.; McKellar, A. R. W. Spectroscopic Identification of Carbon Dioxide Clusters: (CO₂)₆ to (CO₂)₁₃. *Phys. Chem. Chem. Phys.* **2011**, *13*, 1297–1300.
- (41) Oliaee, J. N.; Dehghany, M.; McKellar, A. R. W.; Moazzen-Ahmadi, N. High Resolution Infrared Spectroscopy of Carbon Dioxide Clusters up to $(CO_2)_{13}$. J. Chem. Phys. **2011**, 135, No. 044315.
- (42) Murthy, C. S.; Singer, K.; McDonald, I. R. Interaction Site Models for Carbon Dioxide. *Mol. Phys.* **1981**, *44*, 135–143.

- (43) Bukowski, R.; Sadlej, J.; Jeziorski, B.; Jankowski, P.; Szalewicz, K.; Kucharski, S. A.; Williams, H. L.; Rice, B. M. Intermolecular Potential of Carbon Dioxide Dimer From Symmetry-Adapted Perturbation Theory. J. Chem. Phys. 1999, 110, 3785–3803.
- (44) Tielrooij, K. J.; Garcia-Araez, N.; Bonn, M.; Bakker, H. J. Cooperativity in Ion Hydration. *Science* **2010**, 328, 1006–1009.
- (45) Asmis, K. R.; Neumark, D. M. Vibrational Spectroscopy of Microhydrated Conjugate Base Anions. *Acc. Chem. Res.* **2012**, *45*, 43–52
- (46) Ohtaki, H.; Radnai, T. Structure and Dynamics of Hydrated Ions. Chem. Rev. 1993, 93, 1157-1204.
- (47) Neela, Y. I.; Mahadevi, A. S.; Sastry, G. N. First Principles Study and Database Analyses of Structural Preferences for Sodium Ion (Na⁺) Solvation and Coordination. *Struct. Chem.* **2013**, *24*, 67–79.
- (48) Nguyen, T. V.; Hughes, S. R.; Peslherbe, G. H. Microsolvation of the Sodium and Iodide Ions and Their Ion Pair in Acetonitrile Clusters: A Theoretical Study. *J. Phys. Chem. B* **2008**, *112*, 621–635.
- (49) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
- (50) Ulmschneider, J. P.; Ulmschneider, M. B. United Atom Lipid Parameters for Combination with the Optimized Potentials for Liquid Simulations All-Atom Force Field. *J. Chem. Theory Comput.* **2009**, *5*, 1803–1813.
- (51) Limaye, A. C.; Gadre, S. R. A General Parallel Solution to the Integral Transformation and Second-Order Møller-Plesset Energy Evaluation on Distributed Memory Parallel Machines. *J. Chem. Phys.* **1994**, *100*, 1303–1307.
- (52) Rahalkar, A. P.; Katuoda, M.; Gadre, S. R.; Nagase, S. Molecular Tailoring Approach in Conjunction with MP2 and RI-MP2 Codes: A Comparison with Fragment Molecular Orbital Method. *J. Comput. Chem.* **2010**, *31*, 2405–2418.