Low energy isomers of $(H_2O)_{25}$ from a hierarchical method based on Monte Carlo temperature basin paving and molecular tailoring approaches benchmarked by MP2 calculations

Cite as: J. Chem. Phys. **141**, 164304 (2014); https://doi.org/10.1063/1.4897535
Submitted: 05 June 2014 . Accepted: 28 September 2014 . Published Online: 23 October 2014

Nityananda Sahu, Shridhar R. Gadre, Avijit Rakshit, Pradipta Bandyopadhyay, Evangelos Miliordos, and Sotiris S. Xantheas

Low energy isomers of $(H_2O)_{25}$ from a hierarchical method based on Monte Carlo temperature basin paving and molecular tailoring approaches benchmarked by MP2 calculations

Nityananda Sahu,[1] Shridhar R. Gadre,[1,a)] Avijit Rakshit,[2] Pradipta Bandyopadhyay,[2] Evangelos Miliordos,[3] and Sotiris S. Xantheas[3,a)]
[1]*Department of Chemistry, Indian Institute of Technology Kanpur, Kanpur 208016, India*
[2]*School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi 110067, India*
[3]*Physical Sciences Division, Pacific Northwest National Laboratory, 902 Battelle Boulevard, P.O. Box 999, MS K1-83, Richland, Washington 99352, USA*

(Received 5 June 2014; accepted 28 September 2014; published online 23 October 2014)

We report new global minimum candidate structures for the $(H_2O)_{25}$ cluster that are lower in energy than the ones reported previously and correspond to hydrogen bonded networks with 42 hydrogen bonds and an interior, fully coordinated water molecule. These were obtained as a result of a hierarchical approach based on initial Monte Carlo Temperature Basin Paving sampling of the cluster's Potential Energy Surface with the Effective Fragment Potential, subsequent geometry optimization using the Molecular Tailoring Approach with the fragments treated at the second order Møller-Plesset (MP2) perturbation (MTA-MP2) and final refinement of the entire cluster at the MP2 level of theory. The MTA-MP2 optimized cluster geometries, constructed from the fragments, were found to be within <0.5 kcal/mol from the minimum geometries obtained from the MP2 optimization of the entire $(H_2O)_{25}$ cluster. In addition, the grafting of the MTA-MP2 energies yields electronic energies that are within <0.3 kcal/mol from the MP2 energies of the entire cluster while preserving their energy rank order. Finally, the MTA-MP2 approach was found to reproduce the MP2 harmonic vibrational frequencies, constructed from the fragments, quite accurately when compared to the MP2 ones of the entire cluster in both the HOH bending and the OH stretching regions of the spectra.
 [http://dx.doi.org/10.1063/1.4897535]

Water clusters have received considerable attention over the past two decades since they are considered as effective conduits for understanding the role of water in several chemical and biological processes.[1–4] A large number of experimental studies have been carried out[5–13] aimed at probing the molecular structures and properties of water clusters of increasing size. Overall, the field of water clusters is of fundamental interest in the contemporary science. Since the number of local minima on the cluster's potential energy surface (PES) increases almost exponentially with its size,[11] a plethora of optimization techniques and potential functions have been developed for investigating water clusters with varying sizes.[14–26] In general, because of the *ad hoc* parameterization of empirical interaction potentials for water resulting in the incomplete treatment of many-body interactions, such investigations may lead to an inconsistent prediction of the energetics and the relative rank ordering of the various cluster isomers. Therefore, obtaining the structures of these low-energy candidates using reliable *ab initio* methods is indispensible towards assessing the accuracy of empirical force fields for water as well as benchmarking new theories (such as density functionals). Complementary to this effort, more efficient and reliable algorithms to effectively sample the underlying multidimensional cluster PESs are needed.

It is nowadays well established[27–30] that an accurate description of both the hydrogen bond energy and cooperativity requires methods that include electron correlation, such as the second order Møller-Plesset perturbation theory (MP2) or Coupled Cluster with Singles, Doubles, and perturbative Triple replacements [CCSD(T)]. In recent years, the availability of efficiently parallelizable modules of these highly correlated methods on leadership class computer architectures has rendered such calculations feasible for medium size water clusters ($15 < n < 25$).[31–38] Some notable results reported as early as 2004 range from the MP2 single point energies with an augmented correlation-consistent quadruple zeta quality basis set (aug-cc-pVQZ) for various isomers[31,32] of $(H_2O)_{20}$ to the CCSD(T)/aug-cc-pVTZ energies of $(H_2O)_8$.[39] In 2009 the most time consuming (T) part of a CCSD(T)/aug-cc-pVTZ single point calculation[35] for the cage $(H_2O)_{24}$ required ∼3 h on all 223 200 processors of Oak Ridge National Laboratory's "jaguar" supercomputer (ranked #1 on the top 500 list at the time), required a total of 150 TeraBytes (TB) of memory and demonstrated a sustained performance of 1.39 PetaFLOP/s (1.39×10^{15} floating point operations/s) in double precision. Since then, high level calculations on medium size water clusters have become commonplace.[36–38] Since these computational resources are not widely available, the introduction of more cost-effective approximations, such as fragmentation-based approaches, offers a useful and efficient alternative for the study of water clusters with >20 molecules.[39–42] It

a)Authors to whom correspondence should be addressed. Electronic addresses: gadre@iitk.ac.in and sotiris.xantheas@pnnl.gov.

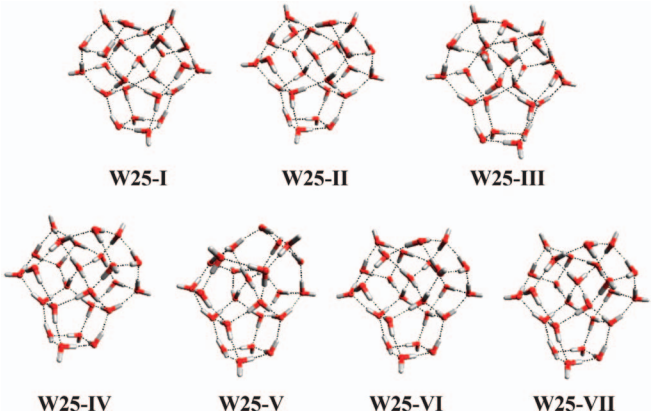

FIG. 1. Optimized geometries of the seven low-lying $(H_2O)_{25}$ isomers at the MP2/aug-cc-pVDZ level of theory.

is conceivable that the two approaches can be combined in an effective strategy, which consists of using the more efficient fragmentation-based approach to identify potential candidates for low-lying structures and the full, highly correlated method [MP2, CCSD(T)] for their subsequent refinement. In this context, molecular dynamics or Monte Carlo based sampling methods such as Basin Hopping (BH)[43,44] or Temperature Basin Paving (TBP)[25] are crucial for identifying candidate structures of enhanced stability for medium size clusters ($n > 10$).

The present work reports 7 new low-lying isomers of $(H_2O)_{25}$. These were obtained from a hierarchical approach based on a Monte Carlo Temperature Basin Paving (MCTBP)[26] method combined with the fragment-based Molecular Tailoring Approach (MTA)[45–48] and subsequently refined by performing MP2 calculations for the entire cluster. The new structures obtained with this approach (*cf.* Figure 1) are lower in energy than the ones reported earlier[41] for this cluster, a fact that further demonstrates the efficacy of the used approach.

Monte Carlo (MC) sampling techniques are a practical tool for locating low energy structures of molecular clusters. The BH algorithm is, in particular, one of the prominent and widely used methods combining MC sampling with local optimization.[43,44] Shanker and Bandyopadhyay[25] developed a variant of the BH method, termed TBP, where the MC sampling is biased by defining an effective temperature, the value of which is dynamically changed based on the sampling history. In the present work this effective temperature is defined as

$$T(E,t) = T_{initial} + \exp[C' H(E,t)], \qquad (1)$$

where $T_{initial}$ is the initial temperature, $T(E, t)$ the temperature for the energy bin with value E at step t, $H(E, t)$ the energy histogram, and C' a normalization constant. The MC acceptance/rejection is carried out as follows: if the energy of the "new" state is lower than the "old" one, the "new" state is always accepted. Otherwise, the "old" state is accepted with probability $\min(1, \exp(-\frac{E_{new}-E_{old}}{kT_{old}}))$, where E_{new} and E_{old} are the energies of the "new" and "old" states, respectively, and T_{old} is the temperature of the "old" state. If the structure is trapped in a local minimum, its effective temperature increases with repeated visits to that minimum. This will eventually increase the effective temperature so it can overcome an energy barrier. Assigning a lower weight to the already visited structures further reduces the sampling of similar structures during the MC process.

The TBP method has been implemented in a local version of GAMESS program package[49] and the interaction between water molecules was described with the Effective Fragment Potential (EFP) model.[50] During the MCTBP simulations the (binding) energy range sampled was from −246 to −175 kcal/mol. The energetically most competitive minima were chosen using a cut-off value of −244 kcal/mol (energy of the most structure reported by Bandow and Hartke[51]). The energetically most competitive minima were chosen using a cut-off value of −244 kcal/mol (energy of the most favorable structure reported by Bandow and Hartke (51) and C′ was taken as 0.015.

The seven lowest energy isomers (**W25-I** through **W25-VII**) are shown in Figure 1. Their energies are lower than the previously reported[41] local minima, also sampled with the EFP potential albeit with a different sampling approach. The new lower energy structures were obtained as a consequence of the new implementation of the TBP sampling algorithm described above. Since the visit of similar structures during the course of optimization is reduced, a larger part of the PES can be sampled. The low energy TBP/EFP geometries were further subjected to MTA-based MP2 geometry optimizations with the aug-cc-pVDZ basis set[52,53] (MTA-MP2/aug-cc-pVDZ) using the Gaussian 09 package.[54] Within MTA, the system is divided into N smaller overlapping fragments, following several criteria like the R_g (R-goodness) parameter,[46,47] which are computed with the full electronic structure calculations (presently MP2). A molecular property, P, for the entire cluster is obtained from the values of the fragments via the set inclusion-exclusion principle[47]

$$P = \sum_{i=1}^{N} P^{f_i} - \sum_{i>j}^{N} P^{f_i \cap f_j} + \cdots + (-1)^{k-1} \sum_{i>j>\cdots>k}^{N} P^{f_i \cap f_j \cdots \cap f_k}. \qquad (2)$$

P^{f_i} denotes the value of P for the ith fragment, $P^{f_i \cap f_j}$ stands for the value of the binary overlap between the ith and jth fragments and so on. The sign $(-1)^{k-1}$ is determined by the number k of the fragments included in each summation. Details of the algorithm, fragmentation, and parameters used in the MTA are given in Refs. 45–48, 55–57. The optimization process is converged when G_{max} and G_{rms} are <0.001 and <0.0005, respectively, and the energy difference $<10^{-5}$ a.u. between four successive optimization steps. The fragmentation scheme is somewhat different for each of the seven $(H_2O)_{25}$ isomers consisting of 14–18 main fragments with maximum sizes between 11 and 14 molecules and an average size of 10 molecules. The total number of fragments ranges from 78 (**W25-III**) to 132 (**W25-V**); see the supplementary material for details on the specific fragments and their size for each isomer.[65]

The MTA-based geometry optimization has been shown to provide good geometries for large molecular clusters.

TABLE I. Energies (a.u.) of the seven $(H_2O)_{25}$ isomers at the MP2, MTA-MP2, and GMTA-MP2 levels with the aug-cc-pVDZ basis set. Numbers in parentheses indicate the relative energy rank of the isomers.

	MP2[a]	MTA-MP2[a]		GMTA-MP2[b]	
Isomer	E (a.u.)	E (a.u.)	δE[c] (mE_h)	E (a.u.)	δE[d] (mE_h)
W25-I	−1906.98472 (1)	−1906.98312 (3)	1.60	−1906.98441 (1)	0.31
W25-II	−1906.98459 (2)	−1906.98526 (1)	−0.67	−1906.98424 (2)	0.35
W25-III	−1906.98425 (3)	−1906.97746 (4)	6.79	−1906.98396 (3)	0.29
W25-IV	−1906.98397 (4)	−1906.98424 (2)	−0.27	−1906.98377 (4)	0.20
W25-V	−1906.98380 (5)	−1906.97374 (5)	10.06	−1906.98353 (5)	0.27
W25-VI	−1906.98358 (6)	−1906.96731 (7)	16.27	−1906.98329 (6)	0.29
W25-VII	−1906.98309 (7)	−1906.97353 (6)	9.56	−1906.98290 (7)	0.19

[a]With the aug-cc-pVDZ basis set.
[b]Obtained from Eq. (3) with (sb) = 6-31+G* and (lb) = aug-cc-pVDZ.
[c]$\delta E = E$(MTA-MP2) − E(MP2).
[d]$\delta E = E$(GMTA-MP2) − E(MP2).

In general, the MTA-optimized geometries differ from the ones optimized during the entire cluster calculation. To further quantify this difference, the MTA-optimal (MTA-MP2/aug-cc-pVDZ) geometries were used as starting points for the MP2/aug-cc-pVDZ optimization of the entire (25 molecule) cluster, performed with the NWChem suite of codes.[58] The MP2/aug-cc-pVDZ optimized geometries for the seven isomers of Figure 1 and their Cartesian coordinates are provided in the supplementary material.[65] In the following, MTA-MP2 denotes the results from the MTA method that combines several fragments of reduced sizes (see the supplementary material[65]) for each isomer, whereas MP2 is used to denote the results for the entire cluster of 25 molecules.

The maximum, average and RMS differences between the MP2 and MTA-MP2 optimized geometries amount to 0.06 Å, 0.02 Å, and 0.02 Å for the hydrogen bond H···O lengths and to 0.006 Å, 0.001 Å, and 0.002 Å for the covalent O–H bond lengths, respectively. These small geometry changes are consistent with the fact that the entire cluster MP2 geometry optimization starting from the MTA-MP2 optimized geometries lowers the MP2 energy by just 5×10^{-4} a.u. (0.31 kcal/mol). All seven low-energy isomers correspond to networks that are structurally similar to one another, each having 8 free O–H groups and 42 hydrogen bonds (see Fig. 1). In addition, they all have 8 DAA-, 8 DDA-, and 9 DDAA-type molecules (D = proton Donor, A = proton Acceptor). This results in 16 monomers connected via three hydrogen bonds and 9 monomers with four hydrogen bonds to nearest neighbors. These isomers can be thought of being assembled by laterally placing a water tetramer to the bottom of a $(H_2O)_{21}$ cluster containing one internally solvated water monomer.[59]

The energy difference between the MP2 and MTA-MP2 energies appears to be in general independent of the basis set used. This observation is further exploited to graft the MTA-MP2 absolute energies by adding a "correction" estimated with a smaller basis (sb) to the MTA-MP2 energies computed with a larger basis (lb). This error seems to originate mainly from the HF rather than the MP2 treatment.[41,60] The grafting (GMTA-MP2) procedure is presently applied by adding a correction obtained with the 6-31+G* basis set as (sb) to the energies obtained with the aug-cc-pVDZ basis set (lb) as[60]

$$E(\text{GMTA-MP2}/lb) = E(\text{HF}/lb) + [E_C(\text{MP2}/sb) - E_C(\text{MTA-MP2}/sb)] + E_C(\text{MTA-MP2}/lb), \quad (3)$$

where E_C is the correlation energy. Table I lists the MP2, MTA-MP2, and GMTA-MP2 energies for all seven isomers. The MP2 and MTA-MP2 energies differ (δE in Table I) by a few mE_h (10^{-3} a.u.), while the two methods predict a quite different energy ordering among the isomers (*cf.* Table I). This may be due to the different fragmentation schemes used to obtain the MTA-MP2 optimized geometries for each isomer (see the supplementary material[65] for the fragmentation scheme and size of fragments used for each isomer). Upon grafting, however, the difference (δE) between the MP2 and GMTA-MP2 energies is reduced by an order of magnitude to just <0.4 mE_h and the relative energy ordering of the isomers is restored to the one found at MP2. All 7 isomers reported here lie *below* the most stable one reported earlier[41] at the GMTA-MP2/aug-cc-pVDZ level, 2.0 mE_h for the most stable (**W25-I**), and 0.5 mE_h for the less stable (**W25-VII**).

In a similar manner, the grafting of the energies at the MTA-MP2/aug-cc-pVTZ level can be performed using either the 6-31+G* or the aug-cc-pVDZ basis set as (sb) as shown in Table II. Again, the MTA-MP2 energies differ from the MP2 ones in a non-systematic way by several mE_h yielding a different ordering, whereas the GMTA-MP2 results display, once again, sub-mE_h accuracy and an identical energy rank compared to the MP2 results. The different network topologies of the seven low-lying isomers span a range of just 1.2 kcal/mol at the MP2/aug-cc-pVTZ level.

The use of Eq. (3) requires a HF/(lb) calculation, which could be cumbersome for a large cluster. This can be avoided by invoking the nearly-basis-set-independent MTA-MP2 energy error, by introducing the alternative equation to generate grafted MTA energies

$$E(\text{GMTA-MP2}/lb) = [E(\text{MP2}/sb) - E(\text{MTA-MP2}/sb)] + E(\text{MTA-MP2}/lb). \quad (4)$$

TABLE II. Energies (a.u.) of the seven $(H_2O)_{25}$ isomers at the MP2, MTA-MP2, and GMTA-MP2 levels with the aug-cc-pVTZ basis set. Numbers in parentheses indicate the relative energy rank of the isomers.

	MP2[a]	MTA-MP2[a]		GMTA-MP2[b]		GMTA-MP2[c]	
Isomer	E (a.u.)	E (a.u.)	δE[d] (mE_h)	E (a.u.)	δE[e] (mE_h)	E (a.u.)	δE[e] (mE_h)
W25-I	−1908.67617 (2)	−1908.67470 (3)	1.47	−1908.67589 (2)	0.28	−1908.67630 (1)	−0.13
W25-II	−1908.67654 (1)	−1908.67695 (1)	−0.41	−1908.67606 (1)	0.48	−1908.67628 (2)	0.26
W25-III	−1908.67565 (4)	−1908.66883 (4)	6.82	−1908.67569 (3)	−0.04	−1908.67562 (4)	0.03
W25-IV	−1908.67570 (3)	−1908.67600 (2)	−0.30	−1908.67565 (4)	0.05	−1908.67573 (3)	−0.03
W25-V	−1908.67503 (5)	−1908.66504 (5)	9.99	−1908.67508 (5)	−0.05	−1908.67510 (5)	−0.07
W25-VI	−1908.67493 (6)	−1908.65871 (7)	16.22	−1908.67491 (6)	0.02	−1908.67498 (6)	−0.05
W25-VII	−1908.67463 (7)	−1908.66489 (6)	9.74	−1908.67465 (7)	−0.02	−1908.67445 (7)	0.18

[a]With the aug-cc-pVTZ basis set.
[b]Obtained from Eq. (3) with (*sb*) = aug-cc-pVDZ and (*lb*) = aug-cc-pVTZ.
[c]Obtained from Eq. (4) with (*sb*) = aug-cc-pVDZ and (*lb*) = aug-cc-pVTZ. The energies with the aug-cc-pVDZ basis are taken from Table I.
[d]$\delta E = E$(MTA-MP2) − E(MP2).
[e]$\delta E = E$(GMTA-MP2) − E(MP2).

This requires MP2 calculations for the entire cluster only with the (*sb*) basis set. Table II lists the GMTA-MP2/aug-cc-pVTZ results obtained via Eq. (4). The very good agreement between the two sets of grafted (GMTA-MP2) energies obtained via Eqs. (3) and (4) (*cf.* Table II) clearly demonstrates the independence of the grafting process on the basis set. The grafted energies obtained via eq. (4) show a smaller error with respect to MP2 when compared to the corresponding ones obtained from Eq. (3), but they interchange the energy ordering between isomers (I) and (II). This energy difference is, however, just 0.08 (aug-cc-pVDZ) and 0.23 (aug-cc-pVTZ) kcal/mol at the MP2 level, with their relative order being switched between the two basis sets. In this respect, both MP2 and GMTA-MP2 predict these two isomers to be practically isoenergetic.

The cluster binding energies are computed as $D_e = E[(H_2O)_{25}] - 25 \times E[H_2O]$. Their values with the aug-cc-pVDZ and aug-cc-pVTZ basis sets are extrapolated to the Complete Basis Set limit, CBS = $D_e(\infty)$, by the two-point formula $D_e(\text{aug-cc-pV}n\text{Z}) = D_e(\infty) + A\ n^{-3}$, which for $n = 2, 3$ (after eliminating A) yields[61]

$$D_e(\infty) = [27D_e(\text{aug} - \text{cc} - \text{pVTZ}) - 8D_e(\text{aug} - \text{cc} - \text{pVDZ})]/19. \quad (5)$$

This expression has been shown to generally overestimate the CBS limit,[62] however, more accurate estimates require calculations with the larger aug-cc-pVQZ basis set and estimates of the Basis Set Superposition Error (BSSE) in conjunction with three-point expressions. The later has been shown[63] – *albeit* for smaller clusters – not to alter the relative order of isomers. The cluster binding energies D_e at the MP2 and GMTA-MP2, listed in Table III, differ by <0.5 kcal/mol from each other with the two methodologies yielding the same relative order at the CBS limit.

All reported conformers are very close in energy. Their relative stability depends on the overall hydrogen bonded network and specifically on the H-bond strengths and, in turn, on the H-bond lengths. From **W25-I** to **W25-VII** the number of H-bonds with shorter lengths (1.6–1.7 Å) is decreased from 10 (**W25-I**) to 8 (**W25-II**) to 6 (**W25-VII**). Furthermore, there is an increase in the average H-bond lengths on going from W25-I to W25-VII. This may be a factor for describing the stability of the isomers that will be explored in future studies that consider many cluster families and isomers.

It is of further interest to assess the accuracy of MTA in reproducing the vibrational spectra via the fragmentation scheme[55] when compared to the ones obtained by MP2 for the entire cluster. In Figure 2 we compare the harmonic spectra of the lowest two isomers (**W25-I** and **W25-II**) obtained with the two methods MP2 (blue line) and MTA-MP2 (red line) with the aug-cc-pVDZ basis (see the supplementary material[65] for the harmonic vibrational frequencies and infra-red (IR) intensities). Panels (a) and (b) display the comparison in the HOH bending, whereas panels (c) and (d) in

TABLE III. GMTA-MP2 and MP2 binding energies (kcal/mol) with the aug-cc-pVDZ and aug-cc-pVTZ basis sets and estimates of the complete basis set (CBS) limit via Eq. (5). Numbers in parentheses denote the relative energy rank ordering of isomers.

Isomer	aug-cc-pVDZ[a]	aug-cc-pVTZ[b]	CBS
	GMTA-MP2		
W25-I	−289.70 (1)	−283.72 (2)	−281.20 (2)
W25-II	−289.59 (2)	−283.83 (1)	−281.40 (1)
W25-III	−289.41 (3)	−283.60 (4)	−281.15 (4)
W25-IV	−289.29 (4)	−283.57 (3)	−281.16 (3)
W25-V	−289.14 (5)	−283.21 (5)	−280.71 (5)
W25-VI	−288.99 (6)	−283.11 (6)	−280.63 (6)
W25-VII	−288.75 (7)	−282.94 (7)	−280.49 (7)
	MP2		
W25-I	−289.89 (1)	−283.90 (2)	−281.38 (2)
W25-II	−289.81 (2)	−284.13 (1)	−281.74 (1)
W25-III	−289.60 (3)	−283.57 (4)	−281.03 (4)
W25-IV	−289.42 (4)	−283.60 (3)	−281.15 (3)
W25-V	−289.31 (5)	−283.18 (5)	−280.60 (5)
W25-VI	−289.18 (6)	−283.12 (6)	−280.57 (6)
W25-VII	−288.87 (7)	−282.93 (7)	−280.43 (7)

[a]Using monomer energies E(MP2/aug-cc-pVDZ//MP2/aug-cc-pVDZ) = −76.26091 a.u.
[b]Using monomer energies E(MP2/aug-cc-pVTZ//MP2/aug-cc-pVDZ) = −76.32895 a.u.

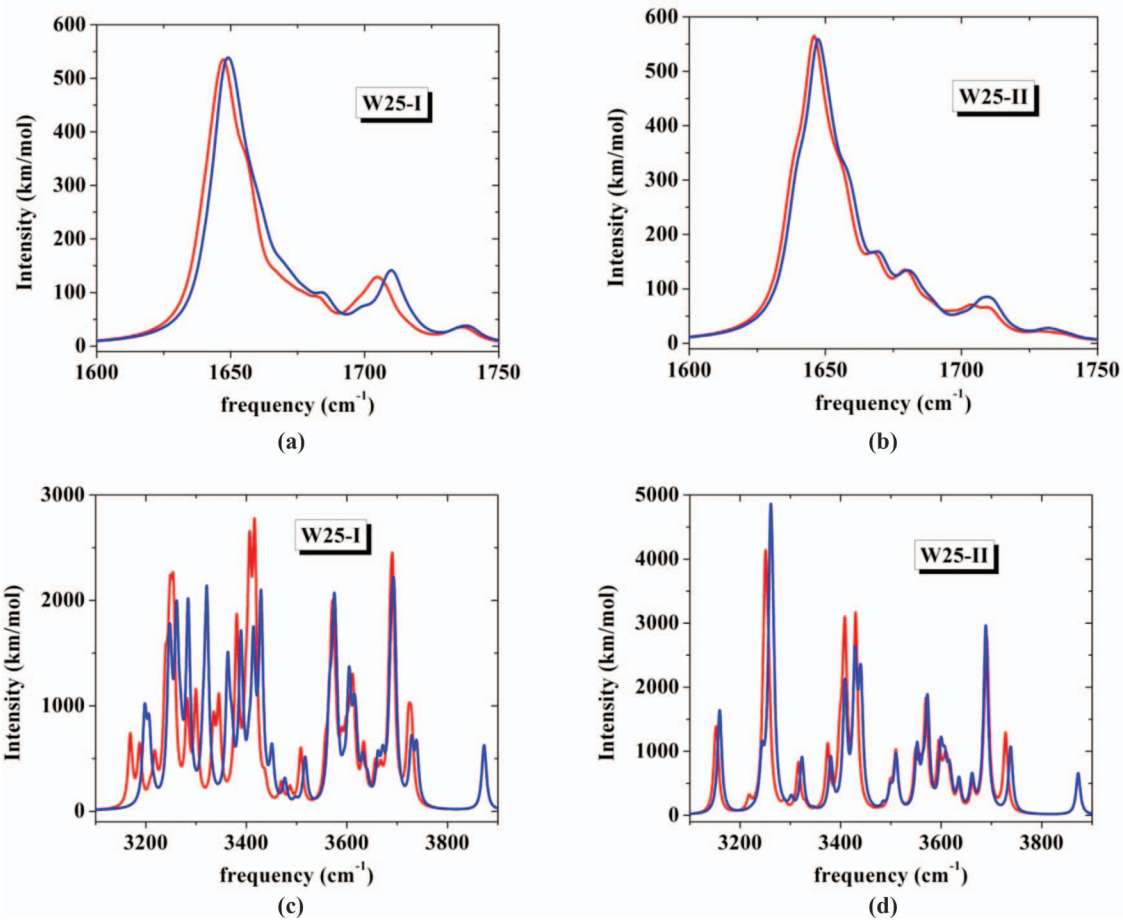

FIG. 2. Comparison between the vibrational spectra of the two lowest isomers of the $(H_2O)_{25}$ cluster obtained at the MP2/aug-cc-pVDZ (red lines) and MTA-MP2/aug-cc-pVDZ (blue lines). (a)–(b): HOH bending (1600−1750 cm^{-1}) region, (c)–(d): OH stretching (3000−4000 cm^{-1}) region.

the OH stretching region of the spectra (constructed from the stick diagrams using Lorentzian functions with full-width at half-maximum of 5 cm^{-1}). The agreement between the MP2 and MTA-MP2 spectra is excellent in the 1600−1750 cm^{-1} region for both isomers and in the 3000−4000 cm^{-1} region for the **W25-II** isomer. As regards the **W25-II** isomer, the peaks in the 3100−3400 cm^{-1} region are shifted by ∼10 cm^{-1}. The spectra for the remaining five isomers at the MTA-MP2/aug-cc-pVDZ level are reported in the supplementary material.[65] The spectra of **W25-I** in the OH stretching region (Fig. 2(c)) exhibit the major peaks recently reported in the experimental spectra. Small deviations of 4.5 and 3.6 cm^{-1} for the 1712 cm^{-1} band between **W25-I** and **W25-II** as well as larger differences in the intensities of two narrow peaks around 3400 cm^{-1} can be attributed to the approximate character of the MTA method. Following the success of the grafting procedure for energies, in a recent development, a similar procedure has been implemented for the accurate estimation of MTA-based Hessian and dipole derivatives. This leads to remarkably accurate vibrational spectra of molecules. The details of this development will be reported in a future publication.

In conclusion, we report seven new low-lying isomers of $(H_2O)_{25}$ obtained via a hierarchical procedure based on sampling the PES with the EFP potential with subsequent optimization using the MTA-MP2 fragmentation approach and final refinement at the MP2 level for the entire cluster. The new implementation of the MC-based TBP method enhances the sampling of the PES by reducing the time that similar structures are visited during the MC steps. This resulted in the identification of new isomers that are lower in energy than the ones reported previously with less efficient sampling approaches. The MTA-MP2 fragmentation approach splits the 25 molecule cluster into smaller fragments with no more than 14 water molecules and yields optimized geometries that are quite close to the fully optimized MP2 ones for the entire cluster. The success of the method hinges upon the fact that the error in MTA energies is almost independent of the basis set used. Excellent agreement between the MTA-MP2 and MP2 frequencies in both the HOH-bending and the OH-stretching regions of the spectra was found. Furthermore, the application of the grafted MTA-MP2 methodology yields electronic energies with sub-mE_h accuracy and restores the energy rank order of the isomers with respect to the MP2 results for the whole cluster. The appreciable savings in computational time and resources required by the MTA-MP2 method (estimated cost of MTA-MP2 for all fragments is ∼20% of that for the MP2 calculation of the entire cluster) render this approach an extremely useful tool for calculating the properties of large systems. We finally note that a similar approach with the one used here has previously predicted[41] a new isomer as a candidate for

the global minimum of $(H_2O)_{20}$. This new isomer has been identified in recent spectroscopic Infra-Red (IR) experimental studies,[64] further supporting the usefulness of the adopted approach in accurately sampling the underlying cluster PES and truthfully ranking the candidates for the low-lying cluster isomers.

ACKNOWLEDGMENTS

N.S. thanks the University Grants Commission, New Delhi for a fellowship. S.R.G. acknowledges the Department of Science and Technology (DST), New Delhi [Project Nos. (SR/S2/JCB/41/2006) and (SR/S1/PC-37/2011)] and the Centre for Development of Advanced Computing (C-DAC) Pune, for financial support. P.B. thanks D. S. T. for Grant No. SR/S1/PC-45/2009. We acknowledge the help of Mr. Sudhanshu Shanker for coding the Monte Carlo Algorithm used. E.M. and S.S.X. acknowledge support from the US Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences. Pacific Northwest National Laboratory (PNNL) is a multiprogram national laboratory operated for the US DOE by Battelle. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

[1] K. Liu, J. D. Cruzan, and R. J. Saykally, Science **271**, 929 (1996).
[2] M. M. Teeter, Proc. Natl. Acad. Sci. U.S.A. **81**, 6014 (1984).
[3] J. Israelachvili and H. Wennerström, Nature (London) **379**, 219 (1996).
[4] R. Ludwig, Angew. Chem., Int. Ed. **40**, 1808 (2001).
[5] U. Buck and F. Huisken, Chem. Rev. **100**, 3863 (2000).
[6] K. Nauta and R. E. Miller, Science **287**, 293 (2000).
[7] F. N. Keutsch and R. J. Saykally, Proc. Natl. Acad. Sci. U.S.A. **98**, 10533 (2001).
[8] J. P. Devlin, J. Sadlej, and V. Buch, J. Phys. Chem. A **105**, 974 (2001).
[9] P. Andersson, C. Steinbach, and U. Buck, Eur. Phys. J. D **24**, 53 (2003).
[10] C. Steinbach, P. Andersson, J. K. Kazimirski, U. Buck, V. Buch, and T. A. Beu, J. Phys. Chem. A **108**, 6165 (2004).
[11] C. J. Burnham, S. S. Xantheas, M. A. Miller, B. E. Applegate, and R. E. Miller, J. Chem. Phys. **117**, 1109 (2002).
[12] J. S. Prell and E. R. Williams, J. Am. Chem. Soc. **131**, 4110 (2009).
[13] T. Hamashima, K. Mizuse, and A. Fujii, J. Phys. Chem. A **115**, 620 (2011).
[14] F. H. Stillinger and A. Rahman, J. Chem. Phys. **60**, 1545 (1974).
[15] O. Matsuoka, E. Clementi, and M. Yoshimine, J. Chem. Phys. **64**, 1351 (1976).
[16] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[17] Y. Wang and J. M. Bowman, J. Chem. Phys. **134**, 154510 (2011).
[18] P. N. Day, R. Pachter, M. S. Gordon, and G. N. Merrill, J. Chem. Phys. **112**, 2063 (2000).
[19] C. J. Burnham and S. S. Xantheas, J. Chem. Phys. **116**, 5115 (2002).
[20] G. S. Fanourgakis and S. S. Xantheas, J. Chem. Phys. **128**, 074506 (2008).
[21] J. K. Kazimirski and V. Buch, J. Phys. Chem. A **107**, 9762 (2003).
[22] H. Kabrede and R. Hentschke, J. Phys. Chem. B **107**, 3914 (2003).
[23] H. Takeuchi, J. Chem. Inf. Model. **48**, 2226 (2008).
[24] S. Kazachenko and A. J. Thakkar, Chem. Phys. Lett. **476**, 120 (2009).
[25] S. Shanker and P. Bandyopadhyay, J. Phys. Chem. A **115**, 11866 (2011).
[26] A. Rakshit and P. Bandyopadhyay, Comput. Theor. Chem. **1021**, 206 (2013).
[27] S. S. Xantheas, T. H. Dunning, Jr., J. Chem. Phys. **99**, 8774 (1993).
[28] K. Kim and K. D. Jordan, J. Phys. Chem. **98**, 10089 (1994).
[29] P. N. Krishnan, J. O. Jensen, and L. A. Burke, Chem. Phys. Lett. **217**, 311 (1994).
[30] E. Miliordos and S. S. Xantheas, J. Chem. Phys. **139**, 114302, (2013).
[31] G. S. Fanourgakis, E. Aprà, and S. S. Xantheas, J. Chem. Phys. **121**, 2655 (2004).
[32] G. S. Fanourgakis, E. Aprà, W. A. de Jong, and S. S. Xantheas, J. Chem. Phys. **122**, 134304 (2005).
[33] S. S. Xantheas and E. Aprà, J. Chem. Phys. **120**, 823 (2004).
[34] S. Bulusu, S. Yoo, E. Aprà, S. S. Xantheas, and X. C. Zeng, J. Phys. Chem. A **110**, 11781 (2006).
[35] E. Aprà, R. J. Harrison, W. A. de Jong, A. P. Rendell, V. Tipparaju, and S. S. Xantheas, in *SC'09: Proceedings of the Conference on High Performance Computing, Networking, Storage and Analysis, SESSION: Gordon Bell finalists, article No. 66* (ACM, New York, NY USA, 2009).
[36] S. Yoo, M. V. Kirov, and S. S. Xantheas, J. Am. Chem. Soc. **131**, 7564 (2009).
[37] S. Yoo, E. Aprà, X. C. Zeng, and S. S. Xantheas, J. Phys. Chem. Lett. **1**, 3122 (2010).
[38] R. M. Shields, B. Temelso, K. A. Archer, T. E. Morrell, and G. C. Shields, J. Phys. Chem. A **114**, 11725 (2010).
[39] D. G. Fedorov and K. Kitaura, J. Chem. Phys. **123**, 134103 (2005).
[40] Z. Yang, S. Hua, W. Hua, and S. Li, J. Phys. Chem. A **114**, 9253 (2010).
[41] J. P. Furtado, A. P. Rahalkar, S. Shanker, P. Bandyopadhyay, and S. R. Gadre, J. Phys. Chem. Lett. **3**, 2253 (2012).
[42] A. Saha and K. Raghavachari, J. Chem. Theory Comput. **10**, 58 (2014).
[43] D. J. Wales and M. P. Hodges, Chem. Phys. Lett. **286**, 65 (1998).
[44] J. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 6611 (1987).
[45] S. R. Gadre, R. N. Shirsat, and A. C. Limaye, J. Phys. Chem. **98**, 9165 (1994).
[46] K. Babu and S. R. Gadre, J. Comput. Chem. **24**, 484 (2003).
[47] V. Ganesh, R. K. Dongare, P. Balanarayan, and S. R. Gadre, J. Chem. Phys. **125**, 104109 (2006).
[48] S. R. Gadre and V. Ganesh, J. Theor. Comput. Chem. **5**, 835 (2006).
[49] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, Jr., J. Comput. Chem. **14**, 1347 (1993).
[50] M. S. Gordon, M. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, and W. J. Stevens, J. Phys. Chem. A **105**, 293 (2001).
[51] B. Bandow and B. Hartke, J. Phys. Chem. A **110**, 5809 (2006).
[52] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[53] R. A. Kendall, T. H. Dunning, Jr. and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[54] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, Gaussian 09, Revision A.01, Gaussian, Inc., Wallingford, CT, 2009.
[55] A. P. Rahalkar, V. Ganesh, and S. R. Gadre, J. Chem. Phys. **129**, 234101 (2008).
[56] A. P. Rahalkar, M. Katuoda, S. R. Gadre, and S. Nagase, J. Comput. Chem. **31**, 2405 (2010).
[57] A. P. Rahalkar, S. D. Yeole, V. Ganesh, and S. R. Gadre, "Molecular tailoring: An art of the possible for *ab initio* treatment of large molecules and molecular clusters linear scaling techniques," in *Computational Chemistry and Physics* (Springer, Dordrecht/Heidelberg/London/New York, 2011), Vol. 13, pp. 199–225.
[58] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. van Dam, D. Wang, J. Nieplocha, E. Aprà, T. L. Windus, and W. A. de Jong, Comp. Phys. Comm. **181**, 1477 (2010).
[59] J. Cui, H. Liu, and K. D. Jordan, J. Phys. Chem. B **110**, 18872 (2006).
[60] N. Sahu, S. D. Yeole, and S. R. Gadre, J. Chem. Phys. **138**, 104101 (2013).
[61] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).
[62] K. A. Peterson, J. R. Lyons, and J. S. Francisco, J. Chem. Phys. **125**, 084314 (2006).
[63] S. S. Xantheas, C. J. Burnham, and R. J. Harrison, J. Chem. Phys. **116**, 1493 (2002).
[64] C. C. Pradzynski, C. W. Dierking, F. Zurheide, R. M. Forck, T. Zeuch, U. Buck, and S. S. Xantheas, "Infrared detection of $(H_2O)_{20}$ isomers of exceptional stability: a drop-like and a face-sharing pentagonal prism cluster," Phys. Chem. Chem. Phys. (published online).
[65] See supplementary material at http://dx.doi.org/10.1063/1.4897535 for cartesian coordinates, details of fragmentation scheme, and MTA-based vibrational spectra for the MP2/aug-cc-pVDZ optimized geometries.