Eur. Phys. J. D **24**, 57–60 (2003)
DOI: 10.1140/epjd/e2003-00182-9

**The European
Physical Journal D**

Efficient global geometry optimization of clusters

Method, and application to water clusters

B. Hartke[a]

Institut für Physikalische Chemie, Christian-Albrechts-Universität, Olshausenstraße 40, 24098 Kiel, Germany

Received 10 September 2002

Abstract. The method of global geometry optimization of atomic and molecular clusters by evolutionary algorithms is briefly presented and reviewed. As an exemplary application of a parallelized implementation of such an algorithm, neutral pure water clusters are globally optimized. In contrast to previous studies, the sophisticated and quantitatively reliable TTM2-F potential is employed. Significant qualitative differences to the earlier results are found, implicating a breakdown of simple water models for water clusters of non-trivial size.

PACS. 02.60.Pn Numerical optimization – 36.40.Mr Spectroscopy and geometrical structure of clusters – 34.20.Gj Intermolecular and atom-molecule potentials and forces

1 Introduction

In the realm of large clusters, many properties can be extrapolated with cluster size [1]. Depending on the observable under study, such smooth extrapolations to smaller clusters start to fail in the size region of $n = 10^3 – 10^7$ and are replaced by seemingly irregular behavior, indicating that cluster properties start to depend on cluster structure. Information on inner structure, however, is usually not directly obtainable from cluster experiments – therefore, theory is needed.

Any theoretical treatment of clusters has to deal with the fact that configuration space increases exponentially with cluster size, as exemplified by Lennard-Jones (LJ) clusters [2,3]. For example, the count of local minima in LJ clusters reaches forbidding numbers very quickly: serious estimates for only 100 atoms are in the range of 10^{40}. Clearly, this is a serious problem for any approach, including molecular dynamics (MD), even if it is often not sufficiently appreciated. It also rules out exact global optimization methods as an alternative route to low-energy structures, since such methods have to cover configuration space completely.

Dropping the aim of exactness, however, leads to practical global optimization algorithms that yield global and low-energy local minima with sufficient reliability in practice. Compared to MD, algorithms of this type offer a complementary, cheaper, and more direct access to experimentally relevant cluster structures. For systems of non-trivial size, though, even these methods are too expensive to be performed with sufficiently accurate *ab initio* treatments of the molecular electronic structure. Therefore, semiempirical and empirical approaches to molecular structure have to be employed; these can then be coupled indirectly with *ab initio* treatments [4,5].

2 Global optimization method

We have used the paradigm of evolutionary algorithms (EA) to create a global cluster geometry optimization method [6–8], incorporating contributions from other authors [9]. The method has been described in detail in these articles, and this class of methods has been reviewed recently [10]; therefore we will give only a brief sketch here.

The algorithm is initialized by creating a set of random cluster geometries, which are then locally optimized. All possible pairs of clusters are then promoted into a larger, intermediate set. Within each pair, each cluster is cut in two parts (not necessarily exact halves), and after swapping one of the parts with the other cluster, two new clusters are formed by re-assembling the parts, replacing the old clusters. Each cluster in the intermediate set is then changed further by small random dislocations of a small subset of its atoms (or molecules), followed by a local optimization. The larger, intermediate set is then reduced to the size of the original set, by an adaptive collection of rules that discards clusters too close to each other in energy or in geometry. In a final postprocessing stage, several attempts are made to improve each cluster geometry further, for example by moving the atom (or molecule) with the smallest contribution to the overall cluster energy to the best vacancy in the cluster. This postprocessing completes one iterative cycle. The iteration is continued by

---

[a] e-mail: `hartke@phc.uni-kiel.de`

again forming all possible pairs of clusters, as described above. The iteration is terminated after a preset maximum number of steps, or after the best energy found so far has not improved during a preset number of steps. At no place in the algorithm, any assumptions about the inner or outer structure of the clusters are made or used.

This algorithm reliably finds all global and low-energy local minima in various practical applications: at first, the algorithm was tested for various benchmark systems of atomic and molecular clusters. In the standard benchmark case of LJ clusters, all currently accepted global minima are found at least up to a number of particles $n = 150$ [7], including notoriously difficult cases like $n = 75, 76, 77$, with a surprisingly low scaling n^3 of computational effort with cluster size. Results of a grid search up to $n = 309$ using prior information [11] could be improved upon in several cases, *e.g.* $n = 185, 186, 187$ [12]. For neutral water clusters using the TIP4P potential [13], the results of a previous global optimization study [14] could be reproduced up to $n = 21$ [8] and extended to $n = 22$.

Non-benchmark applications have ranged from DFT studies on silicon clusters [5] and mixed empirical/*ab initio* calculations on mercury clusters [15] to alkali cation microhydration clusters [16,17]. Structural transition aspects of these results have been reviewed recently [18].

As another advantage for larger clusters or more expensive potentials, this algorithm is embarrassingly parallel: most of the computational time is spent in local optimizations that are entirely independent of each other, and the amount of communication is very small (only one energy and one cluster geometry per task). Therefore, in the present case, we have used a massively parallel machine with a simple master-slaves model. This ensures automatic load balancing if the number of processes is significantly smaller than the number of clusters.

3 Application example: water clusters

3.1 Previous work

Neutral pure clusters of water have been in the focus of interest of experimental and theoretical cluster research [19] for such a long time that comprehensive coverage of the literature is not possible here; see references in the works cited. For smaller clusters $(H_2O)_n$, $n = 2-10$, various series of experimental studies and theoretical calculations have largely come to qualitative or even quantitative agreement [20–22]. The single exception is the notorious case of the water hexamer [20,23,24] which is still not fully closed.

For larger clusters, volume and quality of available data become scarce quickly. There are theoretical predictions of IR spectra in the OH-stretch region, by MP2 calculations with a fairly small basis set [25], for selected isomers of $n = 11, 12$. Actual global geometry minimizations were done up to $n = 21, 22$ on the simple TIP4P potential [8,14]. Although the best TIP4P structures for each cluster size agree qualitatively with the results mentioned above for $n = 2-10$, some doubts remain, bearing in mind the very simple and likely quantitatively insufficient form of the TIP4P potential: in particular, these studies could not detect any appreciable tendency of water clusters towards incorporation of water into the inside of the structures. Instead, with the single exception of $n = 19$, all globally optimal TIP4P clusters from $n = 2$ to $n = 22$ have all water molecules at the surface of the cluster. Cages around central water molecules in the manner of solvation shells seem not to be favored, instead there is an astonishing tendency towards arrangements of face-sharing cubes and pentagonal prisms. Furthermore, there is no discernible growth pattern [8]: the nature of the global minimum structure changes with each addition of a water molecule. All this is in stark contrast to chemical intuition, and to related findings, *e.g.* for hydronium ion [26] and alkali cation [17] microhydration clusters, where both more systematic growth and cage closures at $n = 16-20$ have been found.

3.2 Purpose of present work, water models

In this study, we are putting the global geometry optimization work for neutral pure water clusters with the TIP4P potential to a test, by conducting global optimization on a totally different, more complicated and quantitatively more reliable potential, TTM2-F.

In the TIP4P potential [13], water monomers have a fixed internal structure (hence, there is also no need for any description of intramolecular forces). Monomers interact with each other solely through pairwise, additive terms of two types: (1) a LJ term between the oxygen atoms, summarily modelling dispersion and exchange-repulsion, and (2) Coulomb terms between fixed partial point charges on a dummy site and on the hydrogens. This simplicity makes the model fast to evaluate; therefore, it can and has been employed in many large-scale MD studies with explicit water solvent treatments. Nevertheless, and in spite of the parameters being fit to bulk values, the model captures the essentials of water-water interactions, since it shows good performance not only across various solid and liquid phase properties [13] but also for qualitative structures of small clusters, as mentioned above.

The TTM2-F potential by Burnham and Xantheas [27] features non-rigid monomers, with the highly accurate, empirically adjusted *ab initio* results by Partridge and Schwenke [28] as intramolecular potential and intramolecular dipole moment surface. Partial charges and polarizabilities of the monomers are "smeared out", overcoming divergence problems of more standard multipole schemes. Many-body effects are incorporated by an iterative evaluation of the induction energy over all molecules in the cluster. All this leads to an increase in computational effort by a factor of 20, compared to TIP4P, but also to an impressive accuracy for a wide range of quantities of clusters, solids, and fluids [27,29], in many cases quantitatively agreeing with MP2 results in the complete basis set (CBS) limit, in particular also for small water clusters.

In spite of the parallel implementation, the computational expense of the TTM2-F potential allowed only

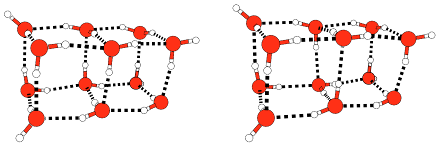

Fig. 1. $n = 12$: (a) parallel *vs.* (b) alternating homodromy.

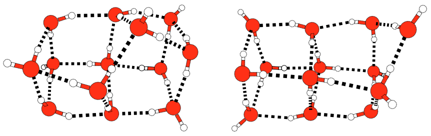

Fig. 2. $n = 13$: (a) central *vs.* (b) terminal addition.

3 independent runs per cluster size, within our computational budget. Agreement of the final results enhances the probability of having found the global minimum with a stochastic algorithm. In case of disagreements, we performed further runs to arrive again at least at this level of reliability. Nevertheless, we anticipate that we cannot go beyond $n = 30$ within reasonable real time.

3.3 Results

For up to and including $n = 11$, TIP4P and TTM2-F global minimum structures agree qualitatively. A first minor difference appears at $n = 12$: the two potentials differ in their preference for two orientational isomers of the same positional structure (two “fused” cubes, *cf.* Fig. 1): the structure can be viewed as a stack of three 4-rings. Each ring is “homodromic” (all donor OH bonds point in the same direction around the ring periphery), which attaches an orientation to the ring. TIP4P prefers to have the same orientation in all three rings, whereas TTM2-F prefers alternating orientations (in Ref. [25], the possibility of different homodromy isomers for $n = 12$ was already noted, but the alternating version was apparently not considered). Note that the energy differences are exceedingly small: 0.2 kJ/mol in TIP4P and 0.5 kJ/mol in TTM2-F. This is well below the accuracy of the models, and even below the accuracy of the best *ab initio* treatments. Actually, these differences likely have no physical relevance. However, in an abstract sense, within each of these two models, they are real and reproducible. Hence it is reassuring that our global optimization method is able to detect them, in spite of their smallness.

For $n = 13$, the first (minor) difference in positional isomers appears: in both cases, the structures can be generated from the $n = 12$ double-cube. In the structure favored by TIP4P, a water molecule of the central 4-ring is replaced by two molecules, tilted away from the connection lines between the neighboring molecules. TTM2-F favors a simple addition of an edge-capping molecule to one of the terminal 4-rings, Figure 2. While for TIP4P the energy difference is even less meaningful than before (0.002 kJ/mol), the TTM2-F energy difference (4.23 kJ/mol) is still small but not insignificant anymore.

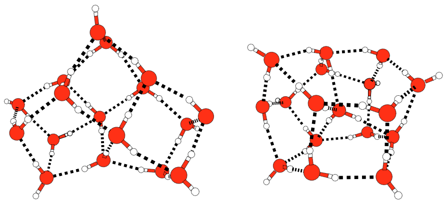

Fig. 3. $n = 17$: (a) all-surface *vs.* (b) centered cage.

For $n = 14$, TIP4P and TTM2-F result in the exact same global minimum. For $n = 15$ (stack of two pentagonal prisms) and $n = 16$ (stack of three cubes), the positional structure is the same, but as in the case of $n = 12$ TIP4P favors parallel homodromy of rings whereas TTM2-F favors alternating antiparallel homodromy. Up to this size, all global and low-energy local isomers in both water models have all molecules at the surface of the cluster.

The first drastic structural difference occurs for $n = 17$: TIP4P favors a distorted structure that can loosely be described as a bent and twisted derivative of the $n = 16$ straight stack of cubes, with one additional molecule inserted at the bend, Figure 3. Although it may not be apparent from the figure at first sight, this structure still has all molecules at the surface of the cluster. TTM2-F, however, favors for the first time a cage structure around one central water molecule. The cage is far from perfect and devoid of symmetry: it is slightly oblate and consists of four 4-rings, four 5-rings and two 6-rings. Nevertheless, the structural difference is significant and not accidental: the energy difference is 2.07 kJ/mol in TIP4P. It is 4.57 kJ/mol in TTM2-F, and at least one other centered cage isomer (with three 4-rings, six 5-rings, and one 6-ring) falls in between the two structures shown in Figure 3.

In the size range $n = 18-20$, as for $n = 14-16$, TIP4P and TTM2-F have global minimum structures that agree in water molecule positions but differ slightly in preference for various orientational isomers. The best $n = 19$ structure is a water-centered cage in both models, whereas $n = 18$ and $n = 20$ are all-surface (two cubes fused to a pentagonal prism, and three fused pentagonal prisms, respectively). Note, however, that various water-centered cage structures are closer in energy to the all-surface global minimum within the TTM2-F model than within the TIP4P model.

For $n = 21$, there is again a major qualitative difference in positional isomers: the best TIP4P structure is again of the cuboid-prismoid variety and quite hard to visualize (Fig. 4) – but it definitely still has all molecules at the surface of the cluster. Finding this structure is a big challenge to global optimization algorithms. They tend to settle much more easily for the second-best isomer, a fusion of two cubes and two pentagonal prisms, with one hydrogen bond missing – another all-surface isomer. The cage structure of TTM2-F is at best the third-lowest isomer in TIP4P, 1.48 kJ/mol higher in energy.

The best TTM2-F structure is again a cage around one central water molecule. This time, the cage consists

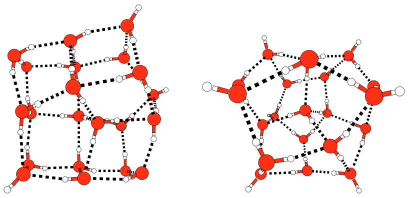

Fig. 4. $n = 21$: (a) all-surface *vs.* (b) centered dodecahedron.

solely of (twelve) 5-rings. Castleman [30] has promoted perfect dodecahedral cage structures with various internal molecules and ions as magic cluster structures. As for other systems [17], the present dodecahedron is not a perfectly regular one, though: the outer shape is not quite spherical (principal moments of inertia: 669.0, 687.3, 705.2) and all of the 5-rings are strongly distorted. Furthermore, there are several other centered cage isomers (with the same and with different 4-, 5-, and 6-ring counts) and the second-best TIP4P isomer mentioned above between this pseudo-dodecahedral structure and the all-surface isomer favored by TIP4P, which is higher in energy by 7.55 kJ/mol in TTM2-F.

According to our as yet uncompleted calculations for still larger systems, the case $n = 21$ seems to be typical at least for slightly larger clusters: TIP4P continues to favor all-surface structures consisting of cubes and pentagonal prisms, while the best TTM2F structures are water-centered cages.

4 Conclusions

As demonstrated in references [27,29], TTM2-F structures and energies for small water clusters can be expected to be quantitatively close to MP2-CBS results, in contrast to TIP4P which is at best qualitatively correct. Hence, where TIP4P and TTM2-F differ for larger clusters, it is a reasonable assumption that TIP4P is more likely to be wrong. Therefore, it has to be concluded that TIP4P seems to fail qualitatively for clusters larger than $n = 12$. In particular, it misses the important structural transition from all-surface to water-centered geometries, occurring in the size range $n = 17-21$ within the TTM2-F model.

Clusters as small as these are not good models for bulk systems (neither liquid or solid), but they are far less removed from them than simple dimers or trimers. In this sense, this study should caution practitioners using simple water models against overinterpreting their results with respect to short-range structural order.

We are currently extending this work towards larger clusters, checking possible explanations for the structural differences found, performing comparative DFT and *ab initio* calculations to further support the TTM2-F model for larger clusters, and searching for property signatures that would enable experimental detection of the all-surface to water-centered transition.

The author thanks Dr. Christian Burnham for supplying various versions of his TTM2-R and TTM2-F programs prior to publication. A Cray T3E computer time grant by the high-performance computing center Stuttgart (HLRS) is gratefully acknowledged.

References

1. J. Jortner, Z. Phys. D: At. Mol. Clust. **24**, 247 (1992)
2. M.R. Hoare, P. Pal, Adv. Phys. **20**, 161 (1971)
3. L.T. Wille, *Annual Reviews of Computational Physics VII*, edited by D. Stauffer (World Scientific, Singapore, 2000), p. 25
4. B. Hartke, Chem. Phys. Lett. **258**, 144 (1996)
5. B. Hartke, Theor. Chem. Acc. **99**, 241 (1998)
6. B. Hartke, J. Phys. Chem. **97**, 9973 (1993)
7. B. Hartke, J. Comput. Chem. **20**, 1752 (1999)
8. B. Hartke, Z. Phys. Chem. **214**, 1251 (2000)
9. D.M. Deaven, K.M. Ho, Phys. Rev. Lett. **75**, 288 (1995)
10. R.L. Johnston, C. Roberts, *Soft Computing Approaches in Chemistry*, edited by H. Cartwright, L. Sztandera (Physica-Verlag, Heidelberg, 2001)
11. D. Romero, C. Barrón, S. Gómez, Comput. Phys. Commun. **123**, 87 (1999)
12. B. Hartke, *Nonconvex optimization and its applications*, edited by J.D. Pinter (Kluwer, Dordrecht, to appear)
13. W.L. Jorgensen, J. Chem. Phys. **77**, 4156 (1982); W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. **79**, 926 (1983)
14. D.J. Wales, M.P. Hodges, Chem. Phys. Lett. **286**, 65 (1998)
15. B. Hartke, H.-J. Flad, M. Dolg, Phys. Chem. Chem. Phys. **3**, 5121 (2001)
16. B. Hartke, A. Charvat, M. Reich, B. Abel, J. Chem. Phys. **116**, 3588 (2002)
17. F. Schulz, B. Hartke, Chem. Phys. Chem. **3**, 98 (2002)
18. B. Hartke, Angew. Chem. **114**, 1534 (2002); Int. Ed. **41**, 1468 (2002)
19. *Clusters of atoms and molecules*, edited by H. Haberland, Springer Series in Chemical Physics (Springer, Berlin, 1994), Vol. 52
20. J. Kim, D. Majumdar, H.M. Lee, K.S. Kim, J. Chem. Phys. **110**, 9128 (1999)
21. J. Sadlej, V. Buch, J.K. Kazimirski, U. Buck, J. Phys. Chem. A **103**, 4933 (1999)
22. E.M. Mas, R. Bukowski, K. Szalewicz, G.C. Groenenboom, P.E.S. Wormer, A. van der Avoird, J. Chem. Phys. **113**, 6687, 6702 (2000)
23. K. Liu, M.G. Brown, C. Carter, R.J. Saykally, J.K. Gregory, D.C. Clary, Nature **381**, 501 (1996)
24. K. Nauta, R.E. Miller, Science **287**, 293 (2000)
25. J. Sadlej, Chem. Phys. Lett. **333**, 485 (2001)
26. M.P. Hodges, D.J. Wales, Chem. Phys. Lett. **324**, 279 (2000)
27. C.J. Burnham, S.S. Xantheas, J. Chem. Phys. **116**, 5115 (2002)
28. H. Partridge, D.W. Schwenke, J. Chem. Phys. **106**, 4618 (1997)
29. C.J. Burnham, S.S. Xantheas, J. Chem. Phys. **116**, 1479, 1493, 1500 (2002)
30. P.M. Holland, A.W. Castleman Jr, J. Chem. Phys. **72**, 5984 (1980)