JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

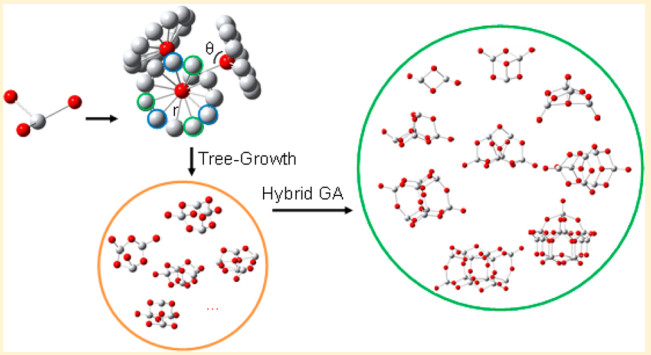
Tree Growth—Hybrid Genetic Algorithm for Predicting the Structure of Small $(TiO_2)_n$, n = 2−13, Nanoclusters

Mingyang Chen and David A. Dixon*

Department of Chemistry, The University of Alabama, Shelby Hall, Box 870336, Tuscaloosa, Alabama 35487-0336, United States

Supporting Information

ABSTRACT: The initial structures for the search for the global minimum of TiO_2 nanoclusters were generated by combining a tree growth (TG) algorithm with a hybrid genetic algorithm (HGA). In the TG algorithm, the clusters grow from a small seed to the size of interest stepwise. New atoms are added to the smaller cluster from the previous step, by analogy to new leaves grown by a tree. The addition of the new atoms is controlled by predefined geometry parameters to reduce the computational cost and to provide physically meaningful structures. In each step, the energies for the various generated structures are evaluated, and those with the lowest energies are carried into the next step. The structures that match the formulas of interest are collected as HGA candidates during the various steps. Low energy candidates are fed to the HGA component to search for the global minimum for each formula of interest. The lowest energy structures from the HGA are then optimized by using density functional theory to study the dissociation energies of the clusters and the evolution in the structure as the size of the cluster increases. The optimized geometries of the $(TiO_2)_n$ nanoclusters for n = 2−13, do not show the character of a TiO_2 bulk crystal with a hexacoordinate Ti. The average clustering energy $\langle \Delta E_n \rangle$ converges slowly to the bulk value for rutile. The TiO_2 dissociation energies for $(TiO_2)_n$ clusters approach the bulk value for rutile more quickly but show larger variations. The $(TiO_2)_{12}$ cluster appears to be quite stable, and the $(TiO_2)_{13}$ cluster is quite unstable on a relative scale.



INTRODUCTION

Small clusters play an important role in the formation of nanoparticles and can exhibit interesting behaviors. The study of small clusters can aid the understanding of nanoparticle formation leading to crystalline materials. Even for small clusters, it is usually difficult to find the global minimum because the number of local minima can be very large, even if only the bond torsions are considered. For a chain such as a peptide, there can be at least N^x local minima, where N is the number of local minima obtained by rotating about a bond and x is the number of the dihedral angles, which leads to Levinthal's paradox for protein folding.[1]

Currently, molecular dynamics methods,[2] simulated annealing,[3] Monte Carlo methods,[4] basin hopping methods,[5] and genetic algorithm (GA) based[6] methods[7] are the main techniques for global minimum searches of small clusters. GA is a widely applied optimization algorithm in many scientific areas. It is used for drug design, protein conformation searches, and global minimum searches of clusters. In GA, crossover of the structures with the best *fitness* (energy based in our case) in the Nth generation gives the $(N + 1)$th generation, and mutations can be introduced into the new generation to enlarge the search area on the potential energy surface (PES). By performing the operation repeatedly, low-lying local minima can be generated, with the possibility that one is the global minimum. The origin of the GA approach is that the global minimum can contain the substructures from several initial structures and intermediate structures. Therefore, the choice of the initial structures for GA can affect the area of the PES that is explored and affects both our ability to locate the global minimum and the convergence to the global minimum.

The current work employs a hybrid genetic algorithm (HGA),[8,9] which is similar to the regular GA global optimization, except that, in each generation, the structures are externally optimized before the *fitness* evaluation. The initial structures for small size clusters can be generated by multiple approaches. One can either obtain starting geometries from crystal structures, or run molecular dynamics on randomly placed atoms for a given force field. The substructures derived from crystal structures are usually not versatile enough to initiate a GA calculation, as GA welcomes as many "good genes" as possible from the initial structures. Molecular dynamics approaches employ the time evolution of structures where chemical bonds are formed between atoms after a certain number of time steps, and the efficiency of the approach largely depends on the quality of the starting geometries, the quality of the force field, and the cost of the energy evaluation. However,

Received: February 8, 2013

   dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

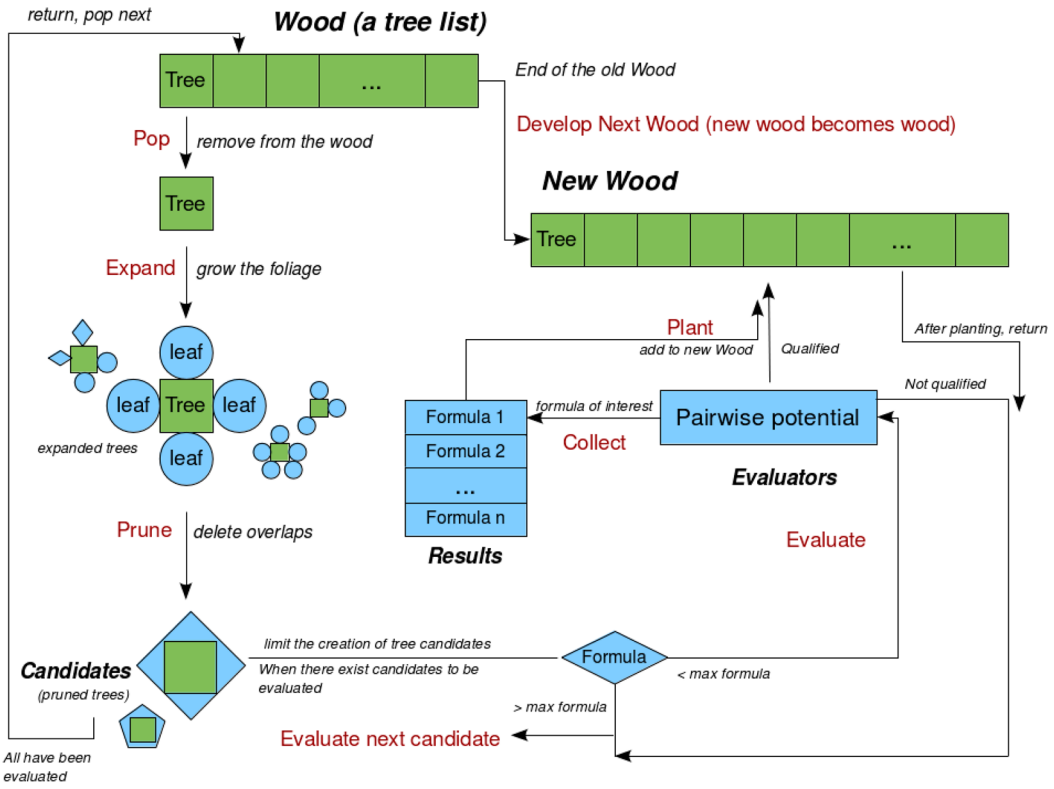


Figure 1. Schematic diagram of the TG algorithm.

for most small size clusters, geometry parameters such as bond lengths can be approximately predicted from small molecules. Since we already have estimates of the bond lengths and the oxidation states an atom can have in a cluster, we can use this information to build the initial structures instead of carrying out hundreds of thousand optimization calculations. In this work, we use a tree growth (TG) algorithm that can take advantage of the so obtained geometry parameters to systematically generate the initial structures for HGA.

We chose as the initial application of this cluster builder method the study of TiO_2 nanoclusters as we have been studying such clusters for catalytic applications, including photocatalysis.[10−13] TiO_2 has been extensively studied as a photocatalyst for solar water splitting to generate H_2 and O_2. Since the first report of this phenomenon using UV light in 1971,[14,15] Hamad et al.[9] used a simulated annealing-Monte Carlo basin hopping (SA) and the hybrid genetic algorithm (HGA) to search for the global minimum of $(TiO_2)_n$ for $n = 1-15$ and optimized the SA/HGA structures at the B3LYP/DZVP level. The authors proposed that there is less than one terminal oxygen atom in $(TiO_2)_n$ for $n > 7$ and all of the lowest energy $(TiO_2)_n$ clusters contain at least one 6-coordinated Ti center for $n > 10$.[9] Qu and Kroes optimized $(TiO_2)_n$ for $n = 1-9$ at B3LYP/LANL2DZ level and argued that the lowest energy $(TiO_2)_n$ structures contain one or two terminal oxygen atoms.[16] We have studied the clusters for up to $n = 4$ with density functional theory and coupled cluster (CCSD(T)) methods.[10] The current work extends this up to $n = 13$ with the discovery of a number of new global minima.

COMPUTATIONAL APPROACHES AND METHODS

Tree Growth (TG) Algorithm. In the TG algorithm (Figure 1), a cluster is grown from an initial small seed structure, which can be a few atoms or a smaller cluster. A molecular structure of interest can be considered to be a tree, and a seed can be seen as the initial stage of a tree. Atoms in a cluster are the tree leaves and are either "dead" or "living". Dead leaves are inactive nodes (atomic sites) that are not subject to expansion by adding atoms to the site. Living leaves are the active nodes to which new atoms can be attached during future growth cycles. The living leaves (i.e., the active nodes) will turn into dead leaves (i.e., inactive nodes) after an expansion as they have already been used in the cluster growth process. The growth process is stepwise, where new leaves (atomic sites) grow on the older branches (previously active sites that become inactive after growth) of a cluster tree layer by layer until the cluster reaches the size of interest. In other words, new atoms are attached to the active nodes (usually the outermost atoms) of the smaller cluster generated from a previous step. Different growth patterns result in different new structures or trees, and we label the new group of structures (trees in our analogy) generated in a cluster growth cycle a "wood". The addition of the new atoms is controlled by predefined geometry parameters to minimize the number of structures that are generated and to make sure that atoms are not overlapped so that structures that are not chemically correct are excluded. In each growth step, the energies (see below for a description of the energy function) for the various generated structures are evaluated, and those that are energetically lower as compared to structures of the same molecular formula are carried into next growth cycle. Structures that match the predefined molecular formulas of interest (e.g., cluster size or number of atoms) are collected during the cycle. Multiple molecular formulas can be input, so with a single run, a set of clusters with different size and formula can be returned. The tree growth algorithm parameters for growing TiO_2 clusters are given in Table S1 in the Supporting Information.

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

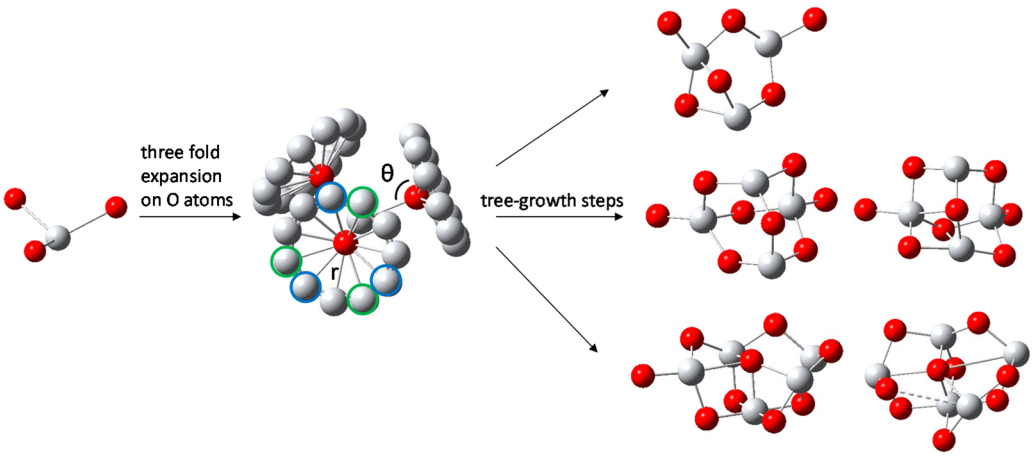


Figure 2. Expansion of each of the O atom active nodes on TiO_3 by three Ti in three folds, then rotating the Ti_{center}−O bond by 30° each step.

Initialization. During the initialization step, a set of global parameters is provided as input to the TG algorithm program including the formula for the maximum cluster size, the basic cluster unit, structures serving as initial seed structures, radial expansion parameters, maximum coordination number for each atom type, etc. The maximum cluster size is the threshold for how large the a cluster is grown. The basic cluster unit is provided so that the algorithm can filter the expansion results and return only clusters with molecular formulas of interest. The initial structures or seeds can be a single atom or a small cluster depending on the goal of the structure search. Radial expansion parameters describe how an active node is expanded radially, as discussed below. The maximum coordination number of an atom limits the number of nearest neighbors.

Expanding and Pruning. The set of active structures (list of trees in a wood) are expanded one at a time. To expand a cluster (tree), its active nodes are expanded first using two methods. The first method is to expand the active node radially and then rotate the expanded site about the bond between the active node and its parent node to generate more configurations. In this part of the expansion algorithm, the clusters are rigid, which means that the bond lengths and bond angles are all fixed, and the rotations of a group about a bond are limited to certain degrees. Therefore, the radial expansion step is determined by four parameters: the expansion angle θ, the expansion bond length r, the maximum number of new sites for a given active node f, and the rotational angle φ. θ is the angle between the newly generated atomic site or group of atoms (the leaf), the active node, and the parent node of the active node. r is the distance between the newly generated atomic site or group of atoms and the active node. φ is $1/(f(m + 1))\cdot 360°$ where m is the number of rotations of the newly generated atomic site or group of atoms about the bond between the active node and its parent node (the rotation axis) to generate the positions of the new sites (leaves). In certain cases, other input parameters such as whether to grow leaves along a rotation axis are available, for example, growth of an octahedron. Figure 2 is an example of the first method, where the current tree is TiO_3 with three oxygen being the active nodes and a center Ti being the parent node of the active nodes. The active nodes are expanded by adding Ti atoms to an O using the parameter values: $\theta = 109°$, $r = 2.0$ Å, $f = 3$, and $\varphi = 30°$. The Ti atoms circled in blue are the new sites (leaves) grown from the O active node before rotation of its position about the $Ti_{original}$−O axis, and the Ti atoms circled in green are the new sites (leaves) generated by a rotation of 30°.

The second expansion method is to search for bridging points of the existing active nodes. For a given active node, are there any active or inactive nodes close enough together to be bridged by a new atom or group? First, the algorithm searches the vicinity of the active node and finds if there is a set of qualified atoms that can be bridged by a new atom or group of atoms. In general, these are nodes sharing the same atom type as the targeted active node. Then, edge-bridging and face-bridging sites are occupied by a new atom or group between the active node and the atoms from the qualified set. The number of edge bridging sites can be infinite as we are simply positioning sites about a line between two atoms. Our approach is to choose one of the edge-bridging sites and then to create multiple sites by rotating the first site around the edge-line between the two atoms by a number of degrees set to a predefined value g.

Each active node in the growing cluster (tree) can have $m + 1$ radial expansions for each radial expansion parameter and one bridge expansion. Each radial expansion contains f new sites. Each bridge expansion generates up to g sites for each edge, if only the edge bridging sites are considered. To reduce the computational cost, bridging sites that are too close to the atoms in the original structure are removed, and a total number of G ($G < g \times$ number of edges) new sites is obtained in the bridge expansion. The sites from the bridge expansion are added to each of the radial expansions to obtain the raw complete sets that consist of $f + G$ available sites. New sites from each raw expansion are discarded if they are either overlapping or too close to the existing nodes in the original cluster or if the addition of a new site exceeds the maximal coordination number of a node. We use a preset parameter dr as the bond length tolerance to evaluate if a new atom B is getting too close to the existing atom A. If $r(B - A)$ is less than $r -$ dr, then atom B is redundant (Figure 3). Counting the coordination number of an atom also involves dr. Only adjacent atoms within the range $[r -$ dr, $r +$ dr) of an atom are considered to be a neighbor of a growth atomic site.

The raw expansion results consisting of $f + G$ sites are a complete set of available sites rather than the sites actually filled by new atoms. Each subgroup of a complete set is added to the original structure to generate a new structure. For an active node, a full expansion might generate several complete sets of the available radial and bridge sites from rotations and radial

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

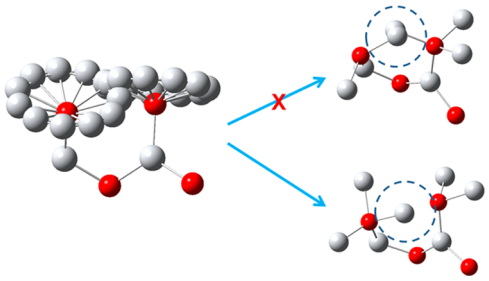


Figure 3. Rejection and acceptance in a 2-node expansion. The dashed circle indicates the overlap threshold $r - dr$.

expansions. The bridge sites are shared by all of these complete sets on an active node and the radial expansion subset of a complete set is orthogonal to a radial expansion subset of other complete sets, which means that a newly generated structure cannot have atoms in radial sites from different complete sets. To enumerate all of the subgroups of a complete set with $f + G$ available sites, a range of k-combinations (k = 0 to $f + G$) for the complete set are used to obtain the final expansion sets (Figure 4a). As shown by the example in Figure 4a, the growth of the Ti active node generates three complete sets of available sites (shown in red, yellow, and purple) orthogonal to each other by a radial expansion with $\theta_{O-Ti-O} = 120°$, $r_{Ti-O} = 1.8$ Å, $f = 2$, and $\varphi = 60°$, $g = 0$, and each of the complete set contains two O sites ($f + G = 2$). The final expansion sets consist of a total number of 10 newly generated cluster structures, in each of which the newly added oxygen atoms belong to the same complete set (shown in the same color). The 0-combination of each complete set gives the same structure as the original cluster, the 1-combination of each complete set gives two newly generated structures by adding one oxygen to the original structure for each color, and the 2-combination of each complete set gives one newly generated structure by adding two oxygen to the original structure for each color.

In many cases, the numbers of k's can be reduced, as the kth-combination with a large k could lead to structures that are overgrown (i.e., have too many atoms). Therefore, k should be no greater than the smaller value between the total number of the atoms in each complete expansion (which is less than $f + G$) and the number of atoms needed to generate the molecular formula. We take the tree growth to generate $(TiO_2)_5$ as an example. In the expansion result of the current structure Ti_3O_7 as an example, each possible expansion is a subgroup of a complete set with $f + G$ available sites, but not all of these

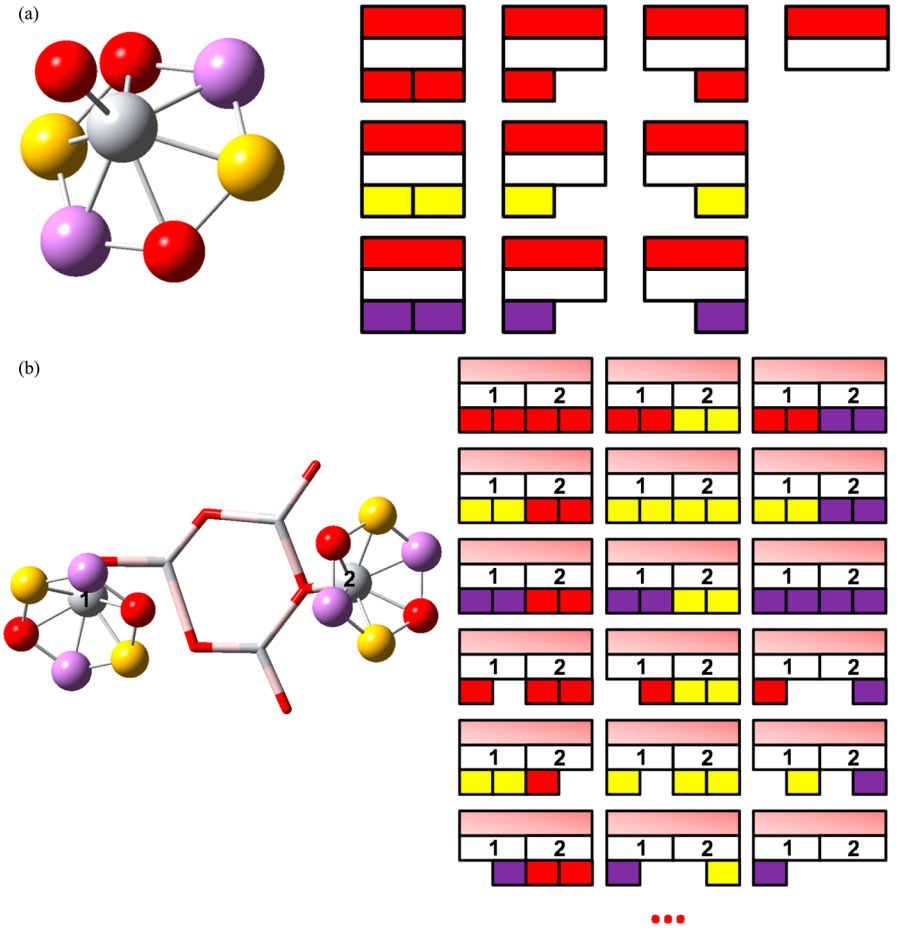


Figure 4. (a) Structures generated by expanding one Ti node with O as the parent node. (b) Selected structures generated by expanding two Ti nodes. The red, yellow, and purple represent the oxygen atoms generated by direct radial expansion (red) as well as by rotating these atoms around the radial expansion axis, O−Ti (yellow and purple). The graphical representations of the new structures are on the right of the ball and stick structures. Each scheme represents a generated structure, in which the upper block represents the parent atom or structure; the middle block represents the current Ti node to expand, the numbers on the middle block in b correspond to the active node 1 and 2, and the bottom block represents the atomic growth sites filled by O. The 2-node expansion in this example generated 100 new structures, and not all structures are shown.

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

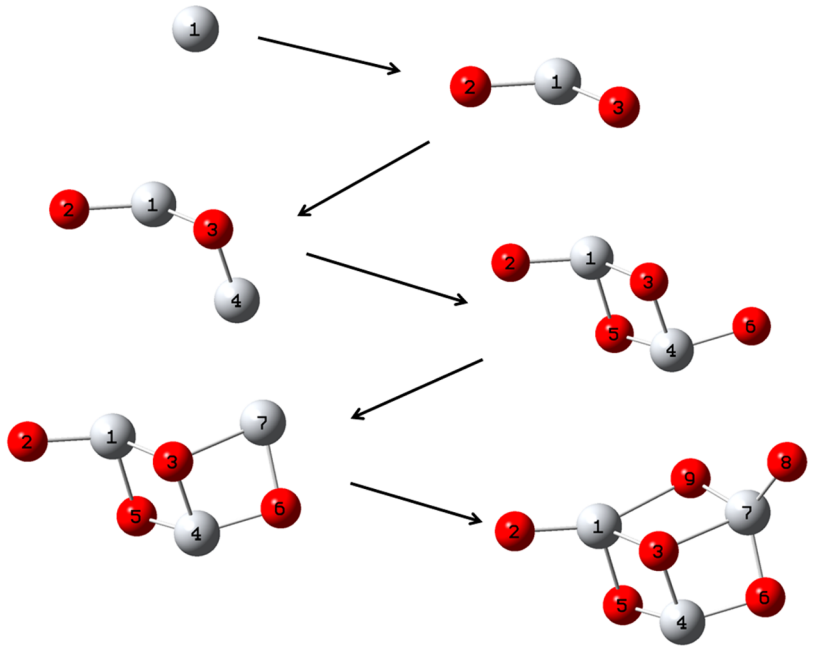


Figure 5. Expansion route to generate $(TiO_2)_3$.

possible expansions are valid for the growth of the cluster to generate the desired $(TiO_2)_5$. Any expansion with more than two new Ti atoms or three new O atoms, or with more than five new atoms will result in a structure exceeding the formula of the maximally allowed structure $(TiO_2)_5$. Therefore, only the k-combinations, $k = 0-5$, of the complete sets need be evaluated.

The k-combination evaluation gives a large number of growth patterns for each active node, which scales factorially to the maximum value of k. Expansion of multiple active nodes at one time further increases the computational cost scaling as C^x, where C is the cost for the expansion of a single node, and x is number of nodes to expand at a time. As a consequence, during each cycle, we only expand at most two active nodes to meet computer memory restrictions (Figure 4b). The expanded active nodes are marked as inactive nodes after an expansion. We then choose one expansion from each active node's final expansion set and add the chosen expansions to the original cluster to get a new structure (an unpruned new tree). These new clusters (trees) are pruned by deleting overlapping sites. Then, the pruned new clusters are added to the candidate pool to be evaluated.

Figure 5 is an example of the expansion used to generate $(TiO_2)_3$, where the final structure is generated step by step from an initial "Ti" seed. For larger clusters, especially for the cases where the cluster geometry varies by a small amount as n increases, smaller clusters can be used as the seeds for generating larger clusters.

Formula Evaluation. First, the formula of a candidate from the expanding and pruning step is compared with the maximal goal formula, and the candidates that exceed the maximal goal formula are removed. The energies of the remaining candidates are evaluated using a two-body Morse-like potential (see below). The energy of the candidate structure is saved into a global list of unique structures if the energy does not appear in the list, otherwise it is discarded. If the formula of the candidate meets any of the goal formulas, it is collected in the results database, which is grouped by goal formulas. For each formula, the tuples of the cluster structure and energy are stored in a sorted list with fixed length, and the energy of the new candidates is compared with the highest energy in the sorted list if the list is already full. If the new candidate has lower energy than the highest energy in the list, the old tuple in the sorted list is then replaced by the new candidate and its energy and the new candidate is labeled as qualified; otherwise, the new candidate is discarded. This helps to limit the number of expansions in later cycles to avoid memory stack overflow due to the combinatorial nature of the problem.

The qualified candidate structures (trees) are grouped together (planted into a new wood) for the next structure expansion cycle. After the current group of structures (wood) is fully expanded, a new group is generated for a new cycle of expansion starts following the same steps. Figure 2 shows the $(TiO_2)_n$ structures obtained from the expansion methods after a few cycles without optimization.

Energy Evaluation. The total potential energy for the raw clusters (in this case Ti_xO_y) are evaluated using eq 1 with parameters taken from the literature (Table S2 in the Supporting Information)[17,18]

$$U(r) = \frac{Z_i Z_j}{r_{ij}} + U_{\text{Morse}} + \frac{C_{12}}{r_{ij}^{12}} \tag{1}$$

with z_i and z_j the charge on the ions i and j, r_{ij} the distance between two atoms (ions) i and j, U_{Morse} the Morse potential, and C_{12}, a parameter to calculate repulsive interactions (Table S2). U_{Morse} is evaluated using

$$U_{\text{Morse}} = D[e^{-2\alpha(r_{ij}-r0)} - 2\, e^{-\alpha(r_{ij}-r0)}] \tag{2}$$

with D the well depth of the Morse curve, r_0 the minimum, and α the width of the potential curve (Table S2, Supporting Information). Other potential energy functions can be chosen depending on the type of molecular structures to be generated.

There are other types of evaluation methods that can be used to filter the candidate structures. For example, the volume of the cluster can be used to screen more condensed structures.

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

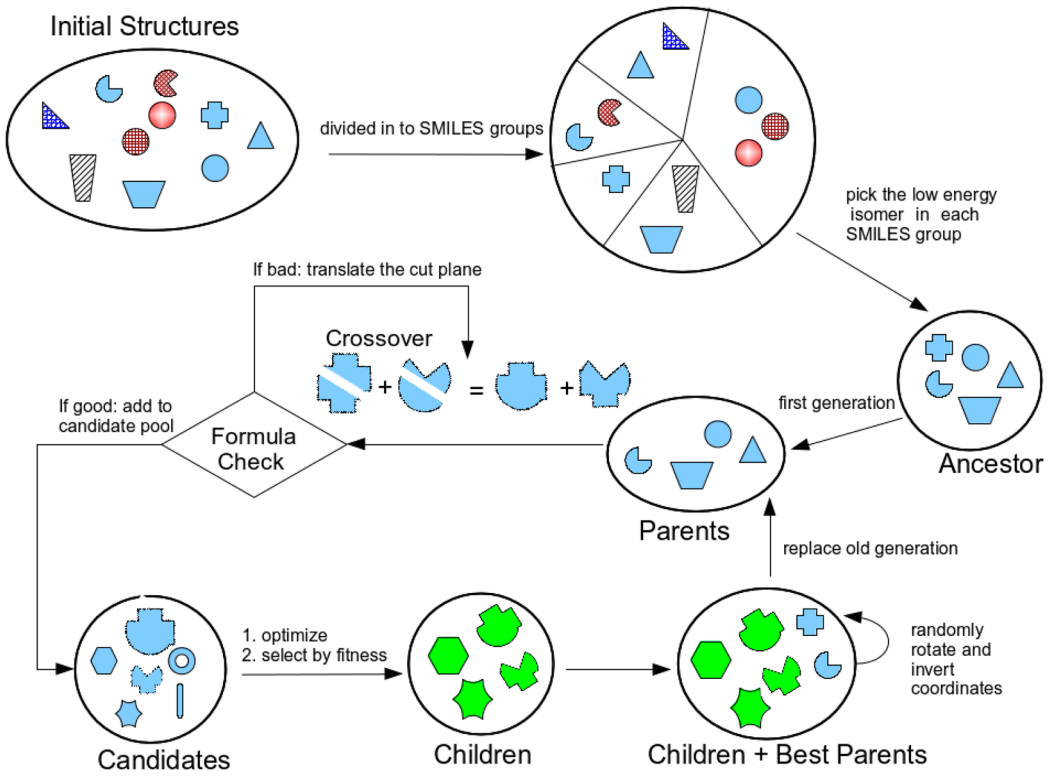


Figure 6. Schematic diagram of the HGA.

Several evaluation methods can be applied for a finer screen. Proper hybrid evaluation methods, created from multiple evaluators, can help filter useful structures. For example, the use of E/V as a hybrid evaluator can be used to filter lower energy condensed structures, where E is the potential energy and V is the size.

Hybrid Genetic Algorithm. The procedure of the HGA is illustrated in Figure 6 and employs methods similar to that described by Johnston[7] and used by Hamad et al.[9] The first (ancestor) generation is obtained from the initial candidates generated by the TG algorithm described above. The geometries of the initial candidates are parsed into simplified molecular-input line-entry system (SMILES)[19] strings. SMILES is a line notation that can be used to represent molecular structures and atomic connectivity. The unique SMILES string for a molecule can be obtained by canonicalization.[19] In the current work, structures with the same canonicalized SMILES string are grouped together. The lowest energy structure of each SMILES group is then chosen to be the member of the first generation for HGA. During the HGA process, each generation of structures is treated with crossover, optimization, evaluation, selection, and mutation operations.

Crossover. The crossover operation is applied to the parent generation pairwise: (a) two structures A and B are picked from the current generation of structures, and a plane is placed cutting through the geometry center of each structure, with these two planes parallel to each other. (b) The upper part of A is added to the lower part B, and vice versa, to determine if the new structures have the desired molecular formula (which is usually same as that of A and B). If the formula is different, the cutting planes in A and B are transformed by the same small amount but in different directions to generate two new structures. The new formula is then checked. If one of the new structures has the desired molecular formula, this structure is added to the candidate pool. If the initial structures A and B have the same molecular formula as the desired molecular formula, each crossover operation will generate two child candidates. (c) Once all of the pairwise combinations of structures in the parent generation are mated, proceed to the next optimization step. In our approach, there is a global counter for the crossovers to reduce the computational cost in the later optimization steps. The counter records the energy of the structures A and B. It helps to lower the frequency of crossovers between structure pairs with the same energies as A and B by allowing only a maximum number of crossovers. If the counter for an energy pair reaches the maximum, the crossover between the structure pair will be skipped.

Optimization. In this version of our approach, we employ semiempirical molecular orbital (SEMO) methods[20−24] to optimize the candidates obtained from the crossover step. Classical force field potentials can also be used if the necessary parameters are available and such parameters do not bias the search by neglecting types of structures. Such methods are computationally inexpensive and can give reasonable geometry parameters and energy orderings if properly chosen. For the $(TiO_2)_n$ clusters, the optimizations are carried out using a SEMO method with the PM6[25] parameters and the AMPAC suit of programs.[26] If an optimized structure has the same energy or is within 0.001 hartree (0.6 kcal/mol) of a previously optimized structure, the structure is eliminated.

Evaluation and Selection. The fitness of a structure is evaluated using $(E - E_0)/(E_{max} - E_0)$, where E is the energy of the structure, E_{max} the maximum energy of the optimized candidates, and E_0 the lowest energy of all of the structures from past generations and the current generation. The lower the ratio $(E - E_0)/(E_{max} - E_0)$, the better the fitness of a structure. If $(E_{max} - E_0)$ is too large, it is set to 0.5 hartree. If $E_{max} - E_0$ is less than a threshold number (e.g., 0.002 hartree),

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

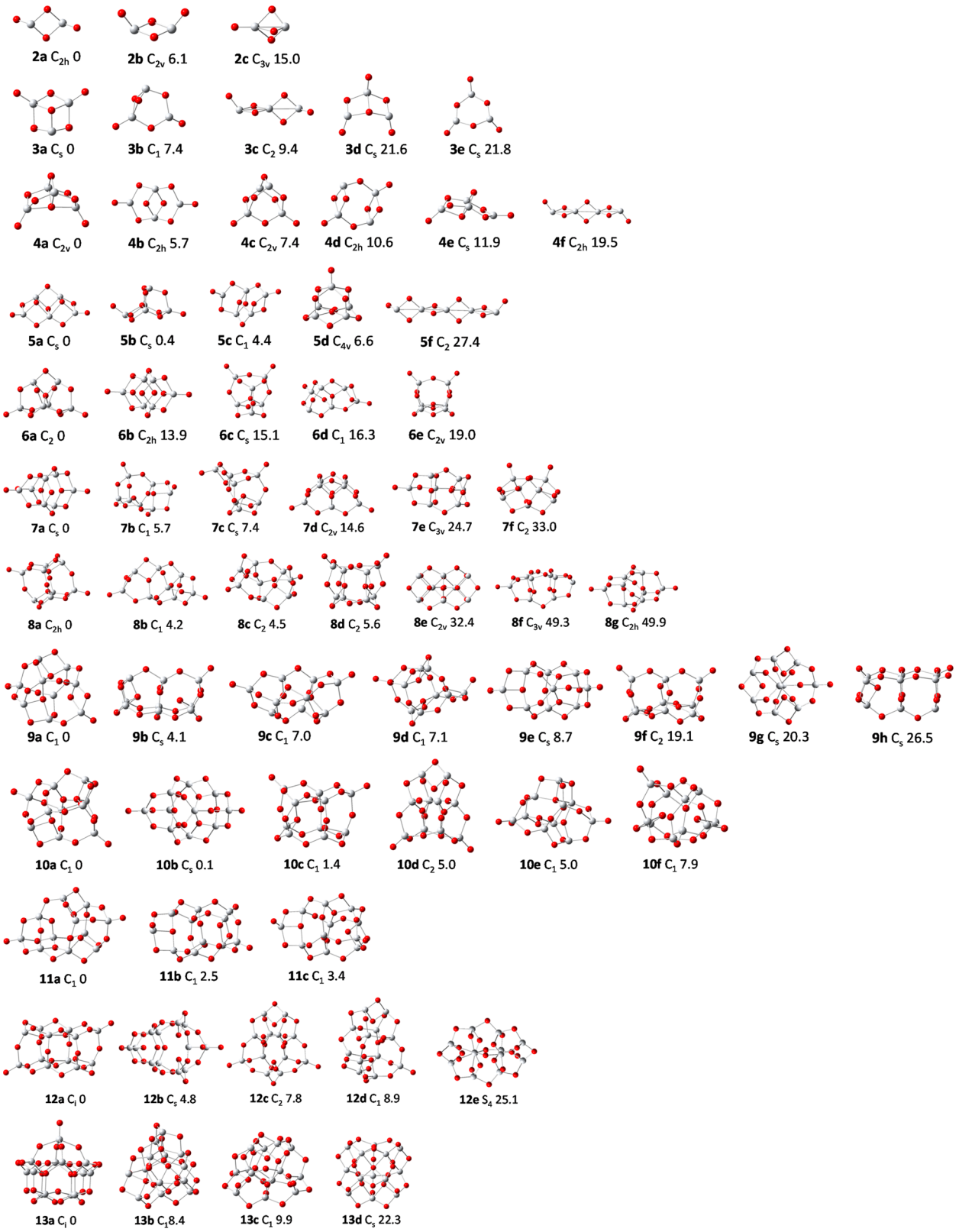


Figure 7. Low energy TiO_2 structures and relative energies to the lowest energy structure (kcal/mol) for n = 2−13.

the HGA procedure terminates, and the structure with energy E_0 is potentially a global minimum. The optimized structure is selected to be a member of the child generation if the fitness is less than a randomly generated floating point number in the range from 0 to 1; otherwise, it is discarded. Structures with the best fitness from the parent generation are also added into the child generation to increase the probability to reuse the "good genes" of the parent generation. Then, the child generation is truncated to a certain size so that only the lower energy structures are carried into the mutation step.

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Mutation. The structures in the selected child generation are mutated to generate randomness in structure crossovers. Mutations are achieved by randomly reorienting the structures and/or, inverting the coordinates through the geometrical center of the structure. After the mutation operation, the child generation replaces the previous parent generation to start a new HGA circle. The goal of all of these structure generations is to appropriately sample the space of geometry parameters in the minimal amount of computational time to find the low energy structures.[27]

DFT Calculations. The low energy structures for each $(TiO_2)_n$ generated by the HGA calculations are further optimized at the density functional theory using the hybrid exchange-correlation functional B3LYP[28,29] with the DZVP2[30] basis set. All of the calculations are carried out using the Gaussian 03/09 suite of programs.[31] The low energy structures from previous workers were also calculated to compare with our results. These structures are either generated from our TG-HGA cluster builder or built by hand. We could not reproduce some of the structures by Hamad et al.[9] because of the poor quality of the images and the lack of atomic coordinates in the Supporting Information, so we calculated the total energy for our low energy structures using the same method and basis set (B3LYP/6-311G** in the Gaussian 03/09 software suite) as used by Hamad et al.[9] to compare energies with those for their lowest energy TiO_2 clusters.

RESULTS AND DISCUSSION

Lowest Energy $(T_iO_2)_n$ Nanoclusters. The optimized geometries for the $(T_iO_2)_n$ clusters at B3LYP/DZVP2 level are shown in Figure 7. Table 1 provides a comparison of the low energy structures found in the current work with previous work. The lowest energy structure for $(TiO_2)_2$ is predicted to have C_{2h} symmetry with two terminal oxygen atoms (**2a**), 6.1 kcal/mol lower in energy than a structure with C_{2v} symmetry and two terminal oxygen atoms (**2b**) and 15.2 kcal/mol lower in energy than a C_{3v} structure with one terminal oxygen atom (**2c**). Our lowest energy isomer prediction for $(TiO_2)_2$ agrees with the previous CCSD(T) prediction of Li and Dixon,[10] as well as the B3LYP prediction of Qu and Kroes.[16] The predicted C_{2h} structure is 35.1 kcal/mol lower in energy than the D_{2h} structure predicted by Hamad et al.[9] at the B3LYP/6-311G**//B3LYP/DZVP level. The lowest energy isomer of $(TiO_2)_3$ is predicted to be a C_s structure with two terminal oxygen atoms (**3a**), in agreement with the previous predictions of the ground state by Li and Dixon[10] using CCSD(T)//B3LYP and by Qu and Kroes[16] using B3LYP. A C_1 $(TiO_2)_3$ structure (**3b**) previously predicted to be ground state by Hamad et al.[9] is found to be 7.4 kcal/mol higher in energy at the B3LYP level. The trimer chain rhombus structure (**3c**) in C_2 symmetry is found to be 9.4 kcal/mol higher than the most stable isomer **3a**, and the ring-like trimer (**3e**) is 21.8 kcal/mol higher in energy than **3a**.

The lowest energy isomer found for $(TiO_2)_4$ matches the C_{2v} structure (**4a**) found by Li and Dixon[10] and by Hamad et al.[9] A C_{2h} structure (**4b**) is predicted to be 5.7 kcal/mol higher in energy than **4a**. Qu and Kroes reported another C_{2v} structure (**4c**) with a 3-coordinate Ti atom to be 6.7 kcal/mol lower in energy at the B3LYP/LANL2DZ level. However, we found **4c** to be 7.4 kcal/mol higher in energy than **4a** at the B3LYP/DZVP2 level. The chain-like isomer (**4f**) is predicted to be 19.5 kcal/mol higher than **4a**.

At the B3LYP/DZVP2 level, the lowest energy structures for $(TiO_2)_5$ are predicted to be a C_s structure (**5a**) with three 3-coordinate O atoms and a Cs structure (**5b**) with a 4-coordinate O atom only 0.4 kcal/mol higher in energy. However, **5b** is found to be 1.5 kcal/mol lower in energy than **5a** using B3LYP with a larger aug-cc-pVTZ basis set. Hamad et al.[9] and Qu et al.[16] previous reported **5b** as the lowest isomer structure. The lowest energy **5a** structure has not been reported previously. The chain-like structure **5f** is found to be ~27 kcal/mol higher in energy than **5a** and **5b**.

The lowest energy structure predicted for $(TiO_2)_6$ matches the lowest energy structure predicted by Qu et al.[16] It is 33.3 kcal/mol lower in energy than the lowest energy isomer proposed by Hamad et al.[9] at the B3LYP/6-311G** level. A C_{2h} structure (**6b**), 13.9 kcal/mol higher in energy than **6a**, is found to be slightly lower in energy than the second lowest energy structure (**6c**) predicted by Qu et al.[16] A C_1 structure (**6d**) is found to be 16.3 kcal/mol higher in energy than **6a**. The third lowest energy structure predicted by Qu et al.[16] (**6e**) is 19.0 kcal/mol higher in energy than **6a**. The chain-like $(TiO_2)_6$ structure (not shown) is ~60 kcal/mol less stable than **6a**, and the energy differences between the chain structure and the most stable isomer became even larger for $n > 6$.

A C_s isomer with two terminal O atoms (**7a**) is predicted to be the most stable structure for $(TiO_2)_7$, 5.7 kcal/mol lower in energy than the second most stable structure in C_1 symmetry (**7b**). The three lowest energy structures in the work of Qu et al.,[16] **7d**, **7c**, and **7e**, are calculated to be 14.6, 7.4, and 24.7 kcal/mol higher in energy than **7a**, respectively. The lowest energy structure reported by Hamad et al.[9] is 33.0 kcal/mol higher in energy than **7a**.

The lowest energy structure of $(TiO_2)_8$ is a C_{2h} structure with two terminal O atoms and two 3-coordinate O atoms (**8a**), followed by a C_1 structure (**8b**) 4.2 kcal/mol higher in energy. Two C_2 structures (**8c** and **8d**) are calculated to be 4.5 and 5.6 kcal/mol less stable than **8a**. The lowest energy structure predicted by Hamad et al.[9] (**8e**) is calculated to be 32.4 kcal/mol less stable than **8a** at the B3LYP/DZVP2 level. The three lowest energy structures obtained by Qu et al.[16] are

Table 1. Survey of the Low Energy Structures $(TiO_2)_n$[a]

n	this work[b]	Hamad et al.[c]	Qu and Kroes[d]	Li and Dixon[e]
2	**2a**	D_{2h}, 35.1	**2a**	**2a**
3	**3a**	**3b**, 7.4	**3a**	**3a**
4	**4a**	**4a**	**4c**, 7.4[f]	**4a**
5	**5a**, **5b**	**5b**	**5b**	
6	**6a**	33.3	**6a**	
7	**7a**	33.0	**7d**, 14.6 **7c**, 7.4	
8	**8a**	**8e**, 32.4	**8f**, 49.3	
9	**9a**	38.3	**9b**, 4.1	
10	**10a**, **10b**	8.8		
11	**11a**	51.5		
12	**12a**	74.7		
13	**13a**	63.4		

[a]Values after "," are the relative energies with respect to the most stable isomer in kcal/mol. Bold numbers refer to the structures in Figure 7 [b]This work: B3LYP//TG-HGA(PM6). [c]Ref 9: B3LYP/6-311G**//B3LYP/DZVP. [d]Ref 16: B3LYP/LANL2DZ. [e]Ref 10: CCSD(T). [f]7.4 kcal/mol higher in energy than **4a** at the B3LYP/DZVP2 level and 6.7 kcal/mol lower than **4a** at the B3LYP/LANL2DZ level. Note that CCSD(T) gives **4a** as the lowest energy structure.

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 2. Number of Bond Types, Average Bond Lengths, Average Coordination Number (CN), and Ti=O Stretching Frequencies for $(TiO_2)_n$, n = 1−13, at the B3LYP/DZVP2 Level

n	Ti—O	Ti=O	avg. r(Ti—O)	avg. r(Ti=O)	CN (Ti)	Ti=O freq[a]
1	0	2	0.00	1.65	2	997 (447), 1019 (42)
2	4	2	1.86	1.64	3	1024 (651), 1043 (0)
3	9	2	1.92	1.63	3.66	1034 (635), 1063 (172)
4	14	2	1.92	1.62	4	1055 (279), 1069 (424)
5b	18	2	1.90	1.62	4	1053 (482), 1061 (111)
5a	19	2	1.94	1.63	4.2	1056 (553), 1062 (190)
6	22	2	1.90	1.62	4	1057 (512), 1063 (108)
7	26	2	1.88	1.62	4	1063 (472), 1067 (142)
8	30	2	1.87	1.62	4	1058 (548), 1061 (0)
9	36	2	1.90	1.62	4.33	1056 (331), 1059 (223)
10	39	2	1.88	1.62	4.1	1056 (373), 1060 (154)
11	43	2	1.88	1.62	4.1	1058 (589), 1060 (15)
12	48	2	1.88	1.62	4.16	1059 (537), 1060 (0)
13	60	1	1.93	1.62	4.69	1066 (150)
rutile			1.966		6	
anatase			1.956		6	

[a]Frequencies in cm^{-1}. Infrared intensity in parentheses in km/mol.

49.3 (**8f**), 62.8 (not shown), and 49.4 (**8g**) kcal/mol higher in energy than **8a**.

A C_1 structure with two terminal O atoms and two 3-coordinate O atoms (**9a**) is found to be the lowest energy isomer for $(TiO_2)_9$, and is 4.1 kcal/mol lower in energy than the lowest energy structure (**9b**) predicted by Qu et al.,[16] and 19.2 and 26.5 kcal/mol lower than the second (**9f**) and third (**9h**) lowest energy structure predicted by Qu et al.[16] Other low energy structures include two C_1 structures (**9c** and **9d**) and a C_s structure (**9e**), all of which have two terminal O atoms and are within 10 kcal/mol of **9a**. The lowest energy structure reported by Hamad et al.,[9] without terminal O atoms, is found to be 38.3 kcal/mol higher in energy than **9a**.

Structures **10a** with C_1 symmetry and **10b** with C_s symmetry are predicted to be the lowest energy isomers for $(TiO_2)_{10}$, with a negligible energy difference between these two isomers at the B3LYP/DZVP2 level. An additional C_1 isomer is found to be only 1.4 kcal/mol higher in energy than structures **10a** and **10b**. Additional low energy structures within a few kcal/mol of **10a** also include **10d**, **10e**, and **10f**. The lowest energy structure reported in the Hamad et al.[9] study is found to be 8.8 kcal/mol higher in energy than **10a**.

The lowest energy isomer **11a** for $(TiO_2)_{11}$ is found to be a C_1 structure with two terminal O atoms. Two low energy structures **11b** and **11c** with one terminal oxygen atom in C_1 symmetry are 2.5 and 3.4 kcal/mol higher in energy. The lowest energy $(TiO_2)_{11}$ structure predicted by Hamad et al.[9] is 51.5 kcal/mol higher than **11a** at the B3LYP/6-311 G ** level.

The most stable structure **12a** for $(TiO_2)_{12}$, with similar tubular building blocks as **8a**, is 4.8 kcal/mol lower in energy than **12b** with C_s symmetry. Structures **12c** in C_2 symmetry and **12d** in C_1 symmetry are also within 10 kcal/mol of **12a**. The lowest energy $(TiO_2)_{12}$ predicted by Hamad et al.[9] is 74.7 kcal/mol higher in energy than **12a**.

Structure **13a** in C_i symmetry is predicted to be the lowest energy isomer for $(TiO_2)_{13}$ and is the first lowest energy structure with only a single terminal oxygen atom. Structure **13a** is calculated to be 63.4 kcal/mol lower in energy than the lowest energy $(TiO_2)_{13}$ predicted by Hamad et al.[9] Two C_1 structures (**13b** and **13c**) are within 10 kcal/mol of **13a**.

Common Structural Motifs Moving to the Bulk. The lowest energy structures for $(TiO_2)_n$, n = 2−12, have two terminal Ti=O bonds. For n = 2 − 12, the average Ti=O bond distance is ∼1.62 Å, and the Ti=O stretching mode frequencies are around 1060 cm^{-1} (Table 2). $(TiO_2)_{13}$ has a single terminal Ti=O bond with a vibrational frequency of 1066 cm^{-1}. The 1060 cm^{-1} band region can serve as a diagnostic band for TiO_2 small clusters with terminal Ti=O bonds as it does not exist in bulk TiO_2.

The general trend is that the coordination number for Ti increases as the size of the TiO_2 cluster increases. $(TiO_2)_{13}$ has the highest average coordination number of 4.69 among the studied clusters, but it is still far below the coordination number of 6 in the bulk crystals. A 6-coordinate Ti atom is not found in our predicted lowest energy $(TiO_2)_n$ structures, with the exception that $(TiO_2)_9$ has a 6-coordinate Ti atom if the bond distance threshold is increased to 2.35 Å (otherwise, it is 5-coordinate). The average Ti−O bond distances show correlations with the Ti coordination numbers. Structures with higher Ti coordination numbers have longer Ti−O bonds. The TiO_2 cluster with the highest Ti coordination number, $(TiO_2)_{13}$ has the longest Ti−O bonds with an average r(Ti−O) of 1.93 Å among all of the studied clusters, and this is still shorter than the Ti−O bonds in rutile (two 1.982 Å and four 1.947 Å) and anatase (two 1.980 Å and four 1.947 Å).[32,33]

The lowest energy $(TiO_2)_8$ and $(TiO_2)_{12}$ structures mainly contain Ti_2O_4 4-member rings and Ti_4O_8 8-member rings, which are the building blocks of the TiO_2 anatase 100 surface, even though $(TiO_2)_8$ and $(TiO_2)_{12}$ have structures that are different from the TiO_2 anatase 100 surface.[34] In the tree growth step, we included a radial O−Ti−O expansion with θ = 90°, f = 4, and m = 2 and allowed the axial growth in the expansion to include a search for structures with 6-coordinate Ti in a pseudo-octahedral site as found in bulk TiO_2. However, we did not find any low energy structures containing 6-coordinate Ti atoms in a pseudo-octahedral site. Thus, these nanoclusters are not large enough to have the bulk-like 6-coordinate Ti atom even in the center of the cluster.

Average Clustering Energies. The clustering energies for the $(TiO_2)_n$ clusters for n = 2−13 are defined by eq 3

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$\Delta E_n = nE(TiO_2) - E((TiO_2)_n) \quad (3)$$

and the normalized clustering energies are given by eq 4

$$<\Delta E_n> = \Delta E_n/n \quad (4)$$

The reaction energies for the dissociation reactions $(TiO_2)_n \rightarrow (TiO_2)_x + (TiO_2)_{n-x}$ are obtained from eq 5

$$\Delta E_d = E((TiO_2)_{n-x}) + E((TiO_2)_x) - E((TiO_2)_n) \quad (5)$$

The normalized clustering energy ($\langle\Delta E_n\rangle$, Table 3) increases as the size of the cluster increases. For example, $\langle\Delta E_{13}\rangle$ is calculated to be 112 kcal/mol, approximately twice the value of $\langle\Delta E_2\rangle$. $\langle\Delta E_n\rangle$ as a function of n was fit for $(TiO_2)_n$ clusters using a sigmoidal function (Figure 8). To fit $\langle\Delta E_n\rangle$, we placed the $\langle\Delta E_n\rangle$ of TiO_2 rutile (156.5 kcal/mol) for $n \sim 1000-2000$, where the region between $n = 1000-2000$ is represented as an asymptotically flat line to represent bulk TiO_2. The experimental value $\langle\Delta E_\infty\rangle$ for bulk rutile is 156.5 kcal/mol obtained from the following heats of formation: experimental $\Delta H_f°(TiO_2$, rutile, 298 K) = −224.4 ± 0.3 kcal/mol[35] and calculated[11] $\Delta H_f°(TiO_2$, gas, 298 K) = −67.9 kcal/mol at the CCSD(T) complete basis set level with zero point energy, core valence, scalar relativistic, and spin orbit corrections. (The experimental value[35] for $\Delta H_f°(TiO_2$, gas, 298 K) is −73.0 ± 3 kcal/mol.) The fitting curve is semiquantitative as our largest value of n is 13. We found that $\langle\Delta E_n\rangle$ for $(TiO_2)_n$ converges toward the bulk value with $\langle\Delta E_{30}\rangle$ estimated to be ∼130 kcal/mol, 84% of of the bulk value, $\langle\Delta E_{50}\rangle$ estimated to be ∼140 kcal/mol, 90% of the bulk value, and $\langle\Delta E_{250}\rangle$ estimated to be ∼150 kcal/mol, 96% of the bulk value. The convergence of $\langle\Delta E_n\rangle$ to the bulk value suggests that the core structural unit in bulk TiO_2, e.g., 6-coordinate Ti, has energetics comparable to the values for the smaller clusters. As noted, even at $n = 13$, the structure is not yet large enough to contain a hexacoordinate Ti.

Table 3. Clustering Energies ΔE_n and Normalized Clustering Energies $\langle\Delta E_n\rangle$ for $(TiO_2)_n$ $n = 2-13$

n	ΔE_n	$\langle\Delta E_n\rangle$
2	112.4	56.2
3	225.7	75.2
4	342.5	85.6
5	456.5	91.3
6	594.2	99.0
7	713.2	101.9
8	838.9	104.9
9	957.0	106.3
10	1083.4	108.3
11	1206.8	109.7
12	1341.7	111.8
13	1449.1	111.5
rutile		156.5

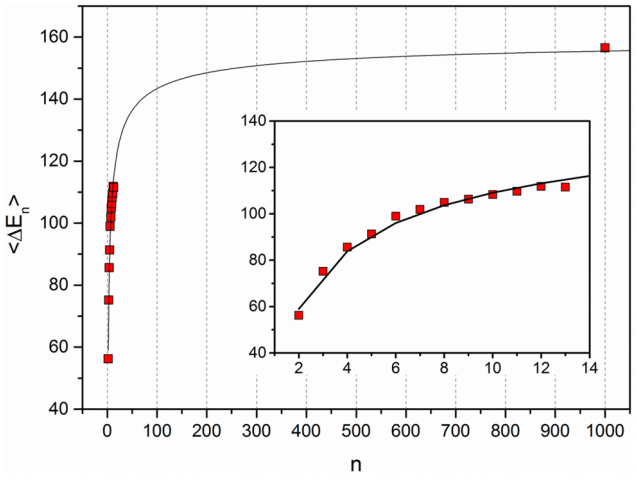


Figure 8. Plot of $\langle\Delta E_n\rangle$ (in kcal/mol) vs n for $(TiO_2)_n$ clusters. The curve is a sigmoidal function.

The dissociation energies for the $(TiO_2)_n$ clusters are shown in Table 4. The monomer dissociation energy increases as the size of the cluster increases in general, with a few exceptions, such as $(TiO_2)_{13}$. The monomer dissociation reaction can also be compared with that for the bulk of 156.5 kcal/mol. In general, the even n clusters have a higher energy to lose TiO_2 and the odd n clusters have lower values. For $n = 12$, the energy to lose a TiO_2 is 86% of the bulk value. However, the energy to lose TiO_2 from $(TiO_2)_{13}$ is the smallest value, 108 kcal/mol. This represents the stability of the $(TiO_2)_{12}$ cluster and the apparent instability of the $(TiO_2)_{13}$ cluster consistent with the loss of a Ti═O bond for $n = 13$ without forming a hexacoordinate Ti. The reactions to divide $(TiO_2)_{2m}$ into two $(TiO)_m$ or divide $(TiO_2)_{2m+1}$ into $(TiO_2)_m$ and $(TiO_2)_{m+1}$ tend to be the most endothermic for most of the clusters.

CONCLUSIONS

A hybrid TG-HGA algorithm has been developed and applied to the study of neutral TiO_2 clusters as an example. The quality of our global minimum search algorithm is mainly determined by the potential used in the last HGA step. The use of a poor potential in the HGA step will fail to find reasonable low energy structures no matter which search algorithm is used, and this is the most important feature in any such search algorithm. Our TG algorithm has its own advantages: (1) it generates preoptimized structures with geometry restrictions rather than randomized coordinates. The geometry restrictions provide positive bias for the initial geometries and the resulting structures can be close to the global minimum or contain substructures that are part of the global minimum, which improve the computational efficiency of the genetic algorithm; (2) it not only can generate structures for a given cluster, it can also generate cluster structures with smaller size than the given goal during the process, which is suitable for a systematic study in a cluster series; (3) the quality of the preoptimized structures can be systematically improved by employing better force field parameters and by including better expansion patterns. The predicted structures and energetics at higher level (such as HGA and DFT// HGA) can be used to improve the force field parameters and the expansion patterns for the TG steps in an iterative fashion. We are currently exploring such an approach. The SMILES string pattern also plays an important role in our method as compared to others. The TG algorithm can generate structures similar to each other, but it is hard to eliminate redundant structures just from their preoptimized coordinates and/or energies. The SMILES string pattern provides a way to recognize such structures as being a unique pattern to reduce the redundancy and increase the variety in the initial structure pool for HGA. It is possible that the TG algorithm could miss a cluster's global minimum if it is very different from the one from which it is grown. However, the TG algorithm merely provides a new set of structures for further optimization. The

dx.doi.org/10.1021/ct400105c | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 4. Reaction Energies for the Dissociation Reactions $(TiO_2)_n \rightarrow (TiO_2)_x + (TiO_2)_{n-x}$ in kcal/mol

n	$TiO_2 + (TiO_2)_{n-1}$	$(TiO_2)_2 + (TiO_2)_{n-2}$	$(TiO_2)_3 + (TiO_2)_{n-3}$	$(TiO_2)_4 + (TiO_2)_{n-4}$	$(TiO_2)_5 + (TiO_2)_{n-5}$	$(TiO_2)_6 + (TiO_2)_{n-6}$
2	112.4					
3	113.4	[a]				
4	116.8	117.8	[a]			
5	114.0	118.5	[a]	[a]		
6	137.6	139.3	142.8	[a]	[a]	
7	119.0	144.3	144.9	[a]	[a]	
8	125.7	132.3	156.6	153.8	[a]	[a]
9	118.2	131.5	137.1	158.0	[a]	[a]
10	126.3	132.1	144.5	146.6	170.3	[a]
11	123.5	137.4	142.2	151.1	156.1	[a]
12	134.8	145.9	158.9	160.2	171.9	153.3
13	107.5	130.0	140.0	149.6	153.7	141.8

[a]Not a unique value.

smaller clusters are not necessarily even minima in the TG algorithm, so our approach does a reasonable job of sampling the coordinate space. It is also not necessary that the global minimum for stage $n + 1$ be grown from the global minimum for stage n in the growth process (see Figure 5 for an example).

New lowest energy structures of $(TiO_2)_n$ that have not previously been reported were found using the TG-HGA algorithm. The optimized geometries of the $(TiO_2)_n$ nanoclusters for $n = 2-13$, do not show the character of a TiO_2 bulk crystal with a hexacoordinate Ti. The average clustering energy $\langle \Delta E_n \rangle$ is converging slowly to the bulk value for rutile. The TiO_2 dissociation energies for $(TiO_2)_n$ clusters approach the bulk value for rutile more quickly but show larger variations. The $(TiO_2)_{12}$ cluster appears to be quite stable, and the $(TiO_2)_{13}$ cluster is quite unstable on a relative scale.

■ ASSOCIATED CONTENT

Supporting Information

Tree growth algorithm parameters for TiO_2 clusters. Parameters for cluster potential evaluation in tree-growth algorithm. Cartesian coordinates in angstroms for the most stable structures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: dadixon@bama.ua.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Argonne National Laboratory (ANL) LDRD program. D.A.D. also thanks the Robert Ramsay Chair Fund of The University of Alabama and Argonne National Laboratory for support.

■ REFERENCES

(1) (a) Levinthal, C. *J. Chim. Phys.* **1968**, *65*, 44−45. (b) Levinthal, C. In *Mossbauer Spectroscopy in Biological Systems*, Proceedings of a Meeting held at Allerton House, Monticello, IL, Debrunner, P., Tsibris, J. C. M.; Munck, E., Eds.; University of Illinois Press: Urbana, IL, 1969, pp 22−24. (c) Zwanzig, R.; Szabo, A.; Bagchi, B. *Proc. Nail. Acad. Sci. U.S.A.* **1992**, *89*, 20−22.

(2) Ramakrishna, M. V.; Bahel, A. *J. Chem. Phys.* **1997**, *106*, 8149.

(3) Kirkpatrick, S.; Gellat, J. C. D.; Vecchi, M. P. *Science* **1983**, *220*, 671−680.

(4) Rusyniak, M. J.; Ibrahim, Y. M.; Wright, D. L.; Khanna, S. N.; El-Shall, M. S. *J. Am. Chem. Soc.* **2003**, *125*, 12001−12013.

(5) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A* **1997**, *101*, 5111−5116.

(6) Holland, J. *Adaptation in Natural and Artificial Systems*; University of Michigan Press: Ann Arbor, MI, 1975; Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley: Reading, MA, 1989; Mitchell, M. *An Introduction to Genetic Algorithms*; MIT Press: Cambridge, MA, 1998; Coley, D. A. *An Introduction to Genetic Algorithms for Scientists and Engineers*; World Scientific: Singapore, 1999.

(7) Johnston, R. L. *Dalton Trans.* **2003**, *22*, 4193−4207.

(8) Hartke, B. In *Structure and Bonding (Berlin)*; Johnston, R. L., Ed.; Springer-Verlag: Heidelberg, 2004; Vol. *110*, p 33.

(9) Hamad, S.; Catlow, C. R. A.; Woodley, S. M.; Lago, S.; Mejías, J. A. *J. Phys. Chem. B* **2005**, *109*, 15741−15748.

(10) Li, S.; Dixon, D. A. *J. Phys. Chem. A* **2008**, *112*, 6646−6666.

(11) Li, S.; Hennigan, J. M.; Dixon, D. A.; Peterson, K. A. *J. Phys. Chem. A* **2009**, *113*, 7861−7877.

(12) Wang, T.-H.; Navarrete-López, A. M.; Li, S.; Dixon, D. A.; Gole, J. L. *J. Phys. Chem. A* **2010**, *114*, 7561−7570.

(13) Wang, T.-H.; Fang, Z.; Gist, N. W.; Li, S.; Dixon, D. A.; Gole, J. L. *J. Phys. Chem. C* **2011**, *115*, 9344−9360.

(14) Fujishima, A.; Honda, K. *Nature* **1972**, *238*, 37−38.

(15) Fujishima, A.; Honda, K. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 1148−1150.

(16) Qu, Z.-W.; Kroes, G.-J. *J. Phys. Chem. B* **2006**, *110*, 8998−9007.

(17) Pedone, A.; Malavasi, G.; Menziani, M. C.; Cormack, A. N.; Segre, U. *J. Phys. Chem. B* **2006**, *110*, 11780−11795.

(18) Swamy, V.; Gale, J. D. *Phys. Rev. B* **2000**, *62*, 5406−5412.

(19) (a) Anderson, E.; Veith, G. D.; Weininger, D. *SMILES: A Line Notation and Computerized Interpreter for Chemical Structures*, Report No. EPA/600/M-87/021; US EPA Environmental Research Laboratory-Duluth, Duluth, MN, 1987. (b) Weininger, D. *J. Chem. Inform. Model.* **1988**, *28*, 31−36. (c) Weininger, D.; Weininger, A.; Weininger, J. L. *J. Chem. Inform. Model.* **1989**, *29*, 97−101. (d) Weininger, D. *J. Chem. Inform. Model.* **1990**, *30*, 237−43. (e) Helson, H. E. In *Reviews in Computational Chemistry*; Lipkowitz, K. B. Boyd, D. B., Eds.; Wiley-VCH: New York, 1999; Vol. *13*, pp 313−398. (f) http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html (accessed April 15, 2013).

(20) Atkins, P.; Friedman. R. *Molecular Quantum Mechanics*, 5th ed.; Oxford University Press: Oxford U.K., 2011.

(21) Cramer. C. *Essentials of Computational Chemistry*; John Wiley: New York, 2002.

(22) Stewart. J. J. P. In *Reviews in Computational Chemistry*; Lipkowitz, K. B. Boyd, D. B., Eds.; VCH: New York, 1990; Vol. *1*, pp 45−81.

(23) Clark. T. *A Handbook of Computational Chemistry. A Practical Guide to Chemical Structure and Energy Calculations*; J. Wiley & Sons: New York, 1985.

(24) Thiel. W. In *Modern Methods and Algorithms in Quantum Chemistry*; Grotendorst, J., Ed.; NIC Series; John von Neumann Institute for Computing, Forschungszentrum Jülich: Jülich, Germany, 2000; Vol. *3*, pp 233−255.
(25) Stewart, J. J. P. *J. Molecular Model* **2007**, *13*, 1173−1213.
(26) *AMPAC 8*; Semichem, Inc.: Shawnee, KS, 2004.
(27) An alpha version of the implementation of the TG-HGA algorithm is available from the authors.
(28) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040−1046.
(29) (a) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789. (b) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200−206.
(30) Sosa, C.; Andzelm, J.; Elkin, B. C.; Wimmer, E.; Dobbs, K. D.; Dixon, D. A. *J. Phys. Chem.* **1992**, *96*, 6630−6636.
(31) 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, Jr., J. A.; 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, N. J.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision B.1 Gaussian, Inc.: Wallingford, CT, 2009.
(32) Meagher, E. P.; Larger, G. A. *Can. Mineral.* **1979**, *17*, 77−85.
(33) Horn, M.; Schwerdtfeger, C. F.; Meagher, E. P. *Z. Kristallogr.* **1972**, *136*, 273−281.
(34) Koch, R.; Lipton, A. S.; Filipek, S.; Renugopalakrishnan, V. *J. Mol. Model.* **2011**, *17*, 1467−1472.
(35) Chase, M. W. NIST-JANAF Themochemical Tables, 4th Edition. *J. Phys. Chem. Ref. Data, Monograph 9* **1998**, 1−1951.