JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Optimization of a Genetic Algorithm for the Functionalization of Fullerenes

Matthew A. Addicoat,*[,†,‡,¶] Alister J. Page,[§,†] Zoe E. Brain,[‡] Lloyd Flack,[∥] Keiji Morokuma,[§,⊥] and Stephan Irle*[,¶]

[‡]Department of Computer Science, Australian National University, ACT, 0200, Australia

[¶]Department of Chemistry, Graduate School of Science, Nagoya University, Nagoya 464-4602, Japan

[§]Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan

[∥]Department of Rheumatology, University of New South Wales, NSW, 2052, Australia

[⊥]Cherry L. Emerson Center for Scientic Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322, United States

Supporting Information

ABSTRACT: We present the optimization of a genetic algorithm (GA) that is designed to predict the most stable structural isomers of hydrogenated and hydroxylated fullerene cages. Density functional theory (DFT) and density functional tight binding (DFTB) methods are both employed to compute isomer energies. We show that DFTB and DFT levels of theory are in good agreement with each other and that therefore both sets of optimized GA parameters are very similar. As a prototypical fullerene cage, we consider the functionalization of the C_{20} species, since for this smallest possible fullerene cage it is possible to compute all possible isomer energies for evaluation of the GA performance. An energy decomposition analysis for both $C_{20}H_n$ and $C_{20}(OH)_n$ systems reveals that, for only few functional groups, the relative stabilities of different structural isomers may be rationalized simply with recourse to π-Hückel theory. However, upon a greater degree of functionalization, π-electronic effects alone are incapable of describing the interaction between the functional groups and the distorted cage, and both σ- and π-electronic structure must be taken into account in order to understand the relative isomer stabilities.

1. INTRODUCTION

Since their discovery, the fullerenes, and in particular C_{60} buckminsterfullerene (discovered in 1985[1]), have been the subject of intense scientific investigation. Of particular recent interest are functionalized fullerenes, which have been utilized in such diverse applications as biosensing,[2] drug-delivery,[3,4] biofuel production,[5] and solar cells.[6,7] With such interest, synthetic routes toward functionalized fullerenes have been presented on a number of occasions.[8−12] Another motivation regarding the investigation of functionalized fullerenes lies in the fact that either endohedral encapsulation or exohedral derivatization of fullerenes may result in cage structures that violate the isolated-pentagon-rule (IPR) (this issue was recently reviewed by Tan et al.[13]). A pertinent example in this respect is the recent experimental synthesis and characterization of the IPR-violating C_{2v}−C_{72} isomer *via in situ* chlorination of C_{72}, forming $C_{72}Cl_4$.[14,15]

While some synthetic methods yield a prescribed placement of functional groups,[16] most methods inevitably yield a mixture of symmetrically distinct functionalized fullerene isomers (see for instance ref 17). The prediction of the relative stabilities of such a spectrum of isomers, or more specifically, the prediction of the single most thermodynamically stable structure and nearby isomers, is a truly daunting task. This is due to the combinatorial dependence of the number of distinct isomers on the size of the cage, M, and the number of exohedral functional groups, n − without considering symmetry, the number of isomers is $M!/[n!(M - n)!]$. In the case of I_h−C_{60} (for which there are fewer distinct isomers, due to the icosohedral symmetry of the cage), the situation becomes unmanageable. For instance, while $C_{60}X_4$ has 487,635 isomers (4,165 of which are symmetrically distinct), $C_{60}X_5$ has over 5,461,512 isomers (45,010 of which are symmetrically distinct). Such large search spaces mean that rigorous, systematic/deterministic search methods are prohibitively expensive.[18,19] To this end, Bihlmeier and co-workers have, in a series of theoretical investigations,[20−22] predicted the stabilities of $C_{50}H_{2n}$, $C_{58}H_{2n}$, and $C_{60}H_{2n}$ based upon a 'screening' approach (whereby the lowest energy C_MH_{2n} isomer is employed to predict the lowest energy $C_MH_{2(n+1)}$ isomer). Although there is experimental evidence suggesting such 'even-numbered' hydrogenation of C_{60} does take place,[23] the latter strategy cannot guarantee that the $C_MH_{2(n+1)}$ is the most stable, particularly in the case of large n.

On the other hand, such large search spaces are amenable to the use of stochastic[24,25] search methods. One such method is a Genetic Algorithm (GA). GAs have been broadly used for chemical search and minimization problems, including determination of cluster geometries,[26−28] protein folding,[29] ligand docking,[30] molecular design,[31] and conformational searching of floppy molecules.[32] A GA is an 'intelligent' stochastic search method that operates using the principles of Darwinian evolution. By using a GA, the solution to the problem of interest is recast into an evolving set (population)

Received: March 7, 2012
Published: April 11, 2012

   dx.doi.org/10.1021/ct300190u | J. Chem. Theory Comput. 2012, 8, 1841−1851

of *genomes*, each of which have an associated *fitness*, which is evaluated by some *fitness function*. From a randomized initial population, successive generations are created by operations with the biological analogues of crossover (i.e., a new, child genome is generated from the complementary parts of two or more parent genomes) and mutation. Fitter individuals have a greater chance of reproducing - i.e. passing on some or all of their genes to the next generation. The process of generation, evaluation, and selection is continued until some predefined criteria, such as stable population, number of generations, or time expended are met. The parameters that define a GA, such as population size, mutation rate, and crossover probability, are themselves freely variable and must therefore be initially optimized for any particular problem (see Table 1).

Table 1. Genome Definition of the Meta-GA[a]

parameter	range	integer options	binary options
population size	5−100		
mutation rate	1−100		
crossover probability	1−100		
elitism?	0/1		
mutation operator	1−4	swap, simple inversion, integer range, integer Gaussian	swap, flip
crossover operator	1−6	1-point, 2-point, uniform, cut-and-crossfill, edge, none	1-point, 2-point, uniform, none
selection operator	1−5	tournament, alt. tournament, rank, roulette wheel, uniform	
scaling operator	1−4	linear, sigma truncation, power law, Boltzmann	

[a]Population size, mutation rate, and crossover probability are smoothly varying numeric parameters; all other parameters are category variables.

Furthermore, the mutation, crossover, and selection operators that generate new genomes may themselves also be freely chosen and can strongly affect the efficiency of the GA. The effects of additional measures such as fitness scaling and elitism further complicate the problem of choosing the most appropriate GA parameters to use.

One method of optimizing GA parameters for a given problem is to use a second genetic algorithm, a Meta-Genetic Algorithm.[33−35] In this case, the genome for the meta-GA encodes all the parameters of the GA that one wishes to optimize (for example, a genome of the form [5,0.2,0.3,1,T] could encode a GA with a population size of 5, a mutation rate of 0.2, a crossover probability of 0.3, that uses mutation operator number 1 and sets elitism to 'true'). The fitness function of the meta-GA in this case is simply the number of genomes evaluated by the encoded GA before locating a solution that is known *a priori*. Given that this approach requires the repeated solution of the target problem many thousands, or even tens of thousands of times, the first step is to select the largest example of the target problem that can be solved by means of an exhaustive search. The results of this exhaustive search are then used as an index table for the candidate GAs generated within the meta-GA. This method of GA optimization has been employed previously, both in a partial fashion[36,37] (optimizing only some parameters while holding others fixed) and recently, optimizing all parameters for a torsionally encoded molecular conformer problem.[32]

Herein we consider the optimization of a GA for locating the most thermodynamically stable isomer of an arbitrary functionalized fullerene. Since an exhaustive list of fullerene isomers and their energies is necessary for such an optimization, we use in this case the smallest fullerene cage, C_{20}, as a prototypical system. Two types of functionalization are considered here, *viz*. hydrogenation and hydroxylation. In both cases, the factors determining the relative stabilities of the functionalized isomers are examined using a range of theoretical methods, including π-Hückel molecular orbital theory, density functional tight binding (DFTB), and density functional theory (DFT). The relative performance of a broad range of GA variables, including problem-specific operators, and differing genome encodings will also be investigated to ensure that the most efficient GA is ultimately optimized. To our knowledge, there have been no previous reported applications of GAs in the context of functionalized fullerenes. We therefore believe this work to be timely in assisting the experimental characterization of an important class of molecular compounds and can rapidly be expanded toward the *a priori* prediction of low-lying functionalization isomers of carbon nanotubes, graphene flakes, etc.

2. COMPUTATIONAL DETAILS

2.1. Model Systems and Quantum Chemical Calculations. The meta-GA method relies on first exhaustively solving the largest feasible 'model' problem, and assumes that the model problem contains all the features of the larger problem. In the current context, this equates to producing an exhaustive list of all symmetrically distinct isomers of a functionalized fullerene, C_MX_n, in terms of each possible genome and the corresponding energy. To place n functional groups on a fullerene of size M yields $M!/[n!(M-n)!]$ isomers (without taking symmetry into account); for this reason, the chosen model fullerene is the smallest member of this carbon cage family, namely C_{20}. Placing more than 4 functional groups on this fullerene results in disintegration of the cage structure (see Supporting Information, Figure S2) and so the model functionalized fullerene was therefore chosen to be $C_{20}X_4$, which has 4,845 isomers in total. By taking symmetry into account, this number is reduced substantially to only 57 structures (here we have defined 'symmetrically equivalent isomers' as those whose Hückel energies agree to within 10^{-6} E_h). Selected Schlegel diagrams of symmetrically distinct isomers of $C_{20}X_n$, n = 2−4, are depicted in Figure 1 (see Figure S4 and S5 for Schlegel diagrams of all isomers).

Previous work[32] showed that in the context of searching the conformational space of floppy organic molecules, the optimized GA was ultimately invariant to the choice of quantum chemical method and basis set among conceptually similar levels of theories, as well as the target molecule. In the current context, conformational issues are irrelevant for the case of fullerene hydrogenation (X═H). On the other hand, they may be significant for fullerene hydroxylation (X═OH), since hydrogen bonding between adjacent functional groups may exist. To this end, we employ two density functional theory-based approaches in this work. One is the B3LYP hybrid functional[38,39] as implemented in Gaussian09,[40] in conjunction with the 6-31G(d) basis set. Klopper and co-workers[20−22] have previously used the BP86 functional[41−43] based on its performance for relative energetics in comparison to higher accuracy levels of theory such as CCSD(T)/CBS.[20−22] As we demonstrate in the Supporting Information however, energetics

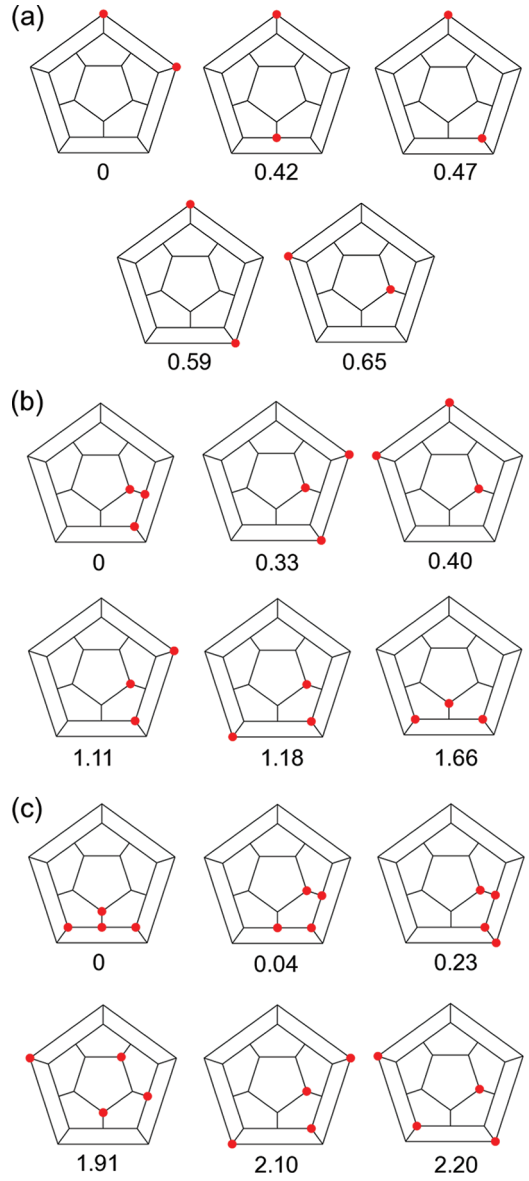

Figure 1. Schlegel diagrams of the 3 most/least stable symmetrically distinct $C_{20}X_n$ isomers for n = 2−4 according to Hückel theory. Red dots denote the position of the functional groups. Stability is defined here according to Hückel β coefficients, which are also given for each structure (in units of β). The numbers of symmetrically distinct isomers for n = 2−4 are 5, 15, and 57, respectively; complete lists of Schlegel diagrams for n = 3 and 4 are provided in Figures S4 and S5.

of C_{20} functionalization computed using B3LYP and BP86 functionals are essentially equivalent, and it is expected that B3LYP, which includes exact exchange and is more widely used, performs better for hydrogen-bonded systems[44] such as hydroxylated fullerenes. Moreover, we will show that GAs generated to describe fullerene functionalization exhibit notable invariance with respect to the DFT-based methods employed in their parametrization (see below). The other such method employed here is the computationally more economical spin-polarized self-consistent-charge DFTB[45−48] method, in conjunction with the mio-0-1 parameter sets.[47] Self-consistency with respect to atomic charges and, where applicable, spin densities was included in this DFTB method so that charge-transfer effects between the C_{20} and the functional group (notably the OH functional group) were taken into account. In the remainder of the text, the acronym "DFTB" will denote both charge-consistent and spin-polarized versions, as prescribed by the spin state: Singlet states are assumed to have closed-shell nature; doublet states are always computed using spin-polarization. For $C_{20}(OH)_4$, we employed an iterative method of geometry optimization based on alternatively minimizing the electronic gradient of the structure (at 0 K) and relaxing the geometry of the structure at higher temperature using molecular dynamics (MD). We also employ a variant of this iterative optimization method, in which MD relaxation is replaced with Hybrid Monte Carlo[49] relaxation. This iterative optimization method consists of initially optimizing the $C_{20}(OH)_4$ geometry at 0 K, after which alternating MD/HMC simulations (at T_n > 0 K) and 0 K reoptimizations, i.e. quenching, are performed. In this work, the number of iterative cycles was limited to 10. All MD calculations performed employed a time-step Δt = 0.2 fs and consisted of 2,000 iterations (totalling 4 ps). Nuclear temperature was maintained during the MD integration *via* a Nosé-Hoover chain thermostat (chain length 3)[50] coupled strongly to the degrees of freedom of the system. To investigate the efficiency of this iterative optimization algorithm, MD relaxations at 298 and 1,000 K were performed; only results obtained using the 298 K simulations are presented here (see the Supporting Information for an extended discussion of temperature effects). Following this iterative DFTB optimization, the final structure was reoptimized using B3LYP/6-31G(d).

Finally, since the addition of a functional group to any atom of a fullerene effectively removes it from the sp^2-hybridized carbon network (through the formation of an sp^3 defect), we will also compare and contrast the applicability of the Hückel molecular orbital method in the current context. In the case of a naked, unfunctionalized fullerene, the construction of the Hückel determinant is based only upon the connectivity matrix of the fullerene cage. The introduction of n functional groups X onto the cage alters this connectivity matrix due to the removal of these carbon atoms from the π-conjugated bond network of the fullerene cage (through the formation of new C−X σ bonds). It is noted that this construction is independent of the identity of X in this work. Despite the simplicity of the Hückel method, it has previously been used to analyze the aromaticity of graphene sheets.[51] We have also recently shown that it serves as an efficient and intuitive tool for understanding complex bonding behaviors observed in graphene oxide.[52] We note finally here that complete lists of Hückel, DFTB, and DFT energies of all species investigated here are included in the Supporting Information (Tables S1−S6).

2.2. GA Optimization. The procedure used to optimize the GA parameters is similar to that used previously[32,53,54] and so is only briefly reiterated here. Both the meta-GA and GA employ the PyEvolve[55] genetic algorithm framework. The genome for the meta-GA was implemented using alleles, allowing for different limits in each gene position, this is different than the "winner-takes-all" probability densities used in previous work. The ranges allowed for each gene are shown in Table 1.

Previous work[32,53,54] employed a successive reoptimization approach where each reoptimization set the most significant factor from the previous optimization at its ideal value and optimized the remaining parameters. In this work, all parameters were accounted for simultaneously. For each combination of molecule and level of theory, the meta-GA was run 500 times, with randomly generated initial conditions,

to generate 500 GA parameter sets. Each parameter set was then used 100 times to determine its efficiency, measured as the mean number of evaluations (i.e., number of generations × population size) and reliability (number of instances that located the minimum). In these calculations, each GA was terminated if it exceeded the number of evaluations equivalent to an exhaustive search.

In the language of a GA, functionalizing an arbitrary fullerene cage can be defined by indicating which carbon atoms bear the functional groups, in terms of the indices of those carbon atoms. With this information (the genome), an initial candidate can be built and subsequently optimized. There are two convenient representations that can be used to encode the location of n functional groups a C_M fullerene: either the genome can be constructed as n unique integers in the range $[1{:}M]$, where each gene directly indicates the number of the functionalized carbon atom; or the genome can be a binary string of length M with elements of 1 denoting a functionalized carbon atom, and elements of 0 denoting an unfunctionalized carbon atom. The efficacies of both options are one of the main topics in the present work.

To model the GA parameters and their performance, a generalized additive model[56] was fitted in R version 2.9.0[57] (using the mgcv package[58,59]) to the mean number of evaluations each GA required to identify explanatory variables for GA efficiency. The model used was

$$M_{\text{eff}} = f(P.\ m) + g(c) + E + M + C + S + Sc + E.\ M \quad (1)$$

where M_{eff} is the measured mean evaluations taken by 100 independent runs of the GA with unreliable GAs excluded as previously described; $g(c)$ and $f(P.m)$ are smooth nonlinear functions of the crossover rate (which is estimated by regression on a basis of splines with a roughness penalty to the fit), and population size/mutation rate, respectively; E, M, C, S, and Sc, are the effects of elitism, mutation operator, crossover operator, selection operator, and scaling operator; and $E.M$ considers the joint effect of elitism and mutation operator. In both smooth functions, the degree of smoothing used was chosen to maximize the prediction accuracy. The value of the smoothing parameter that led to this was estimated by Generalized Cross Validation. A more detailed discussion of the optimization of these GA parameters is provided in the Supporting Information.

3. RESULTS AND DISCUSSION

3.1. Functionalization of C_{20}: $C_{20}H_n$. Our discussion begins with a comparison of the complete genomes of $C_{20}H_n$ functionalization computed using Hückel (i.e., $C_{20}X_n$) and DFTB/DFT (i.e., $C_{20}H_n$); this comparison is made in Figure 2, in which the relative stabilities of $C_{20}H_n$ computed using Hückel are compared with DFTB/DFT values of ΔE

$$\Delta E = E(C_{20}) + nE(H) - E(C_{20}H_n) \quad (2)$$

which corresponds to the energy associated with the addition of n hydrogen atoms to the C_{20} cage. For $C_{20}H_2$, Figure 2(a) shows that the predicted isomer ordering (in terms of relative energy) using Hückel theory matches the predictions of DFT. The same can almost be said upon comparison of Hückel and DFTB; in this case isomers 1 and 2 (1) are reversed using DFTB. The difference in energy here is only *ca.* 0.4 kcal/mol however, one that is arguably beyond the accuracy of this method. Thus, in the case of only two disruptions to the C_{20} conjugated π-network, Hückel theory suffices the more sophisticated quantum mechanical treatments. As we have recently demonstrated in the context of graphene functionalization,[52] Hückel theory assists in the intuitive understanding of the functionalization of sp^2 conjugated systems, at a vastly cheaper computational cost. The most stable isomers of $C_{20}X_2$/$C_{20}H_2$ are those in which the disruption to π-conjugation is most localized; i.e. typically on adjacent carbon atoms. On the other hand, the least stable isomers in this case are those in which the opposite is the case; i.e. typically forming one or more diradical carbon atoms. This point will be a recurring theme in this work. We note that, in general, ΔE values predicted by DFTB are generally overbinding in comparison to those from DFT, which is a typical trend when atomic energies are employed as reference. Relative isomer energies are apparently in much better agreement between DFTB and DFT.

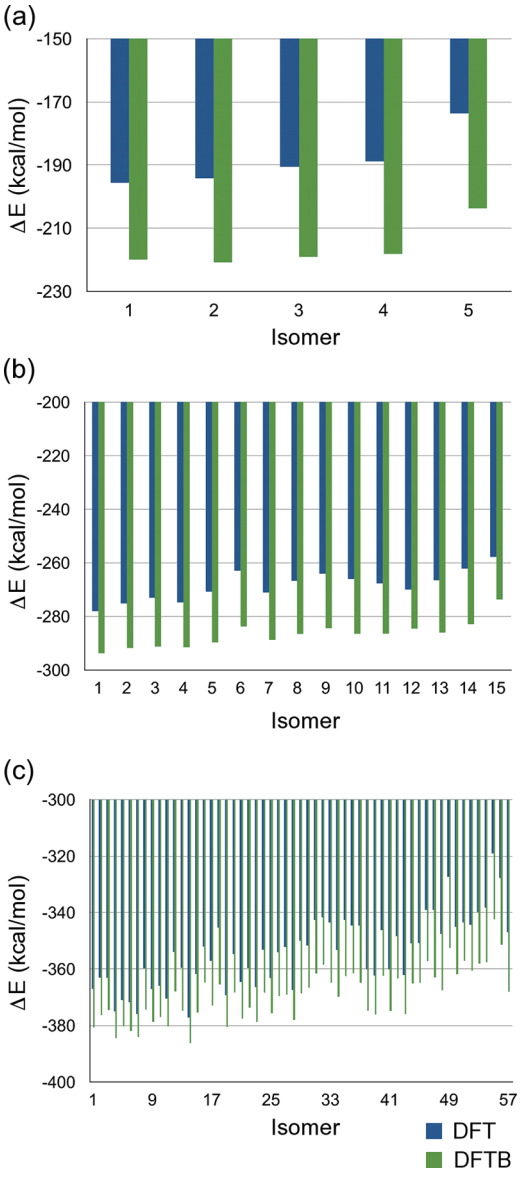

Figure 2. Comparison of relative stabilities calculated using Hückel with ΔE calculated using DFTB/DFT for (a) $C_{20}X_2/C_{20}H_2$, (b) $C_{20}X_3/C_{20}H_3$, and (c) $C_{20}X_4/C_{20}H_4$. Isomer numbers (horizontal axes) are defined in terms of ascending Hückel β coefficients. DFTB/DFT relative energies ΔE are given in kcal/mol.

This is, however, not the complete story. The comparison of Hückel, DFTB, and DFT predicted isomer ordering for $C_{20}X_3$/$C_{20}H_3$ is compared in Figure 2(b). In this case, the two density functional-based methods essentially predict the same isomer orderings in terms of energy. Again, the general trend predicted using Hückel and DFT/DFTB are also seen to be in agreement. In this case however, isomers 6 and 9 are notable exceptions; they are among the most stable according to Hückel theory yet correspond to the 3rd and 4th least stable structures using DFTB/DFT, respectively. The same phenomenon is observed, albeit more dramatically, for $C_{20}X_4$/$C_{20}H_4$ (Figure 2(c)). The general agreement between Hückel and DFTB/DFT observed for $C_{20}X_2$ and $C_{20}X_3$ is no longer present. This is illustrated succinctly by isomers 1, 14, 55, and 57 (which are the 11th, 1st, 57th, and 35th most stable structures according to DFT).

We have further elucidated these discrepancies in the manner of Zheng et al.[60] by performing a simplified energy decomposition analysis (EDA)[61,62] of fullerene functionalization at the DFT level of theory. ΔE may be decomposed such that

$$\Delta E = E_{def}(C_{20}) + E_{int} \quad (3)$$

where $E_{def}(C_{20})$ is the energy associated with deforming the equilibrium structure of the cage, respectively, and E_{int} is the interaction energy between the deformed fullerene cage and the functional group (separated by an infinite distance). $E_{def}(H)$ is of course zero, but this will not be the case for other functional groups such as OH, as we will discuss below. $E_{def}(C_{20})$ and E_{int} corresponding to $C_{20}H_4$ are given in Figure 3. $E_{def}(C_{20})$ for $C_{20}H_2$, $C_{20}H_3$, and $C_{20}H_4$ is essentially consistent at *ca.* 10−15 kcal/mol for each hydrogen added to the cage, a value that corresponds therefore to significant structural deformation. Conversely, comparison of E_{int} and ΔE shows a clear correlation: the most energetically favorable isomers are those with the largest E_{int} values. In particular, isomer 14, which is the minimum energy isomer according to DFTB/DFT, exhibits the largest E_{int} of all. Thus, the relative stabilities of these functionalized fullerenes *in general* cannot be understood in terms of a single parameter (*viz.* the disruption to π-conjugation). Instead it must be understood in terms of three: the geometrical relaxation of the functionalized cage; the disruption to π-conjugation; and higher order effects arising from the interaction with between X and the deformed fullerene cage (i.e., 'non-π' effects which cannot be modeled using Hückel molecular orbital theory). In cases where cage deformation is less prominent (as is seen here for $C_{20}X_2$ and most cases of $C_{20}X_3$), however, understanding these relative stabilities may be achieved with recourse simply to Hückel theory. An identical conclusion was recently made with respect to graphene oxidation[52] and the functionalization of carbon nanotubes.[60,63] We stress here therefore that the functionalization of these three carbon-based nanostructures are seemingly equivalent and may be discussed using the same language.

3.2. Functionalization of C_{20}: $C_{20}(OH)_n$. We turn now to the case of OH functionalization of C_{20}. An extended discussion regarding the MD- and HMC-based iterative optimization algorithms employed for $C_{20}(OH)_n$ is provided in the Supporting Information. Suffice to say that both methods ultimately predict the same minimum energy structure for each $C_{20}(OH)_n$ isomer. The annealing temperature employed also has no noticeable effect in this respect. However, the HMC-based optimization algorithm provides a more efficient exploration of the molecular potential energy surface, compared to the MD-based algorithm. Nevertheless, both MD- and HMC-based algorithms are successful in identifying transition state structures corresponding to the cleavage of one, two, and three hydrogen bonds between adjacent OH groups. DE, defined in this case as

$$\Delta E = E(C_{20}) + nE(OH) - E(C_{20}(OH)_n) \quad (4)$$

for n = 2, 3, and 4 is provided in Figure 4. This figure shows that the agreement between DFT/DFTB for $C_{20}(OH)_n$ is generally within *ca.* 20−30 kcal/mol per functional group. Nevertheless, the qualitative agreement between DFT/DFTB here is excellent; the differences in relative DFT/DFTB energies for these species are no larger than 2.0, 3.5, and 4.4 kcal/mol per functional group, respectively. Both methods therefore predict the same trend in ΔE for each series of isomers, as reported before for the case of unfunctionalized fullerene cages.[64] Consideration of Figure 4 leads to an analogous conclusion to that reached in the previous section. Namely, that agreement between the predictions of Hückel and QM methods regarding the relative stabilities of $C_{20}(OH)_n$ isomers are in agreement only for low values of n. The

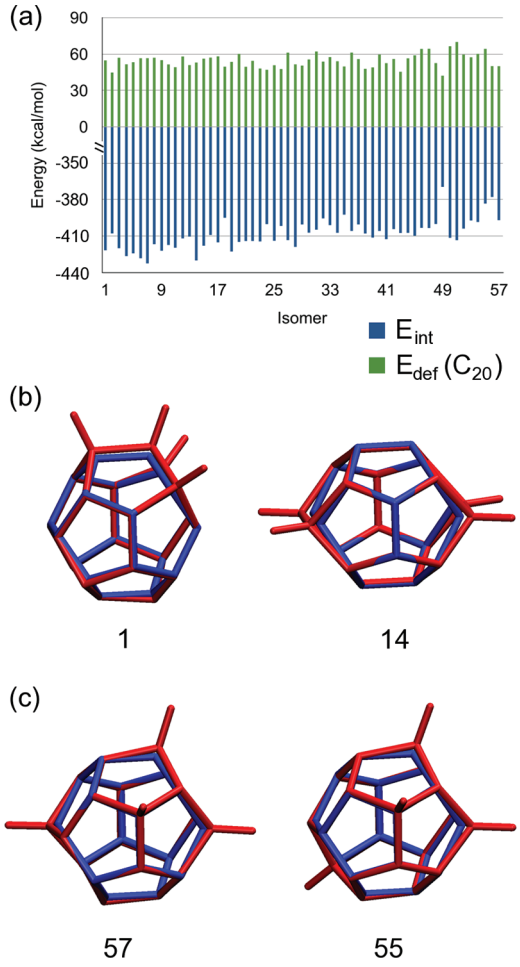

Figure 3. (a) EDA of $C_{20}H_4$ functionalization. DFT $E_{def}(C_{20})$ and E_{int} for 57 distinct isomers shows that the isomers with the most favorable ΔE are those with the largest E_{int}. (b), (c) Structural deformation of the (b) most and (c) least stable $C_{20}H_4$ isomers according to Hückel (left) and DFT/DFTB (right). Structures shown are DFT-optimized. Functionalized structures are given in red, and the pristine C_{20} cage is given in blue.

dx.doi.org/10.1021/ct300190u | J. Chem. Theory Comput. 2012, 8, 1841−1851

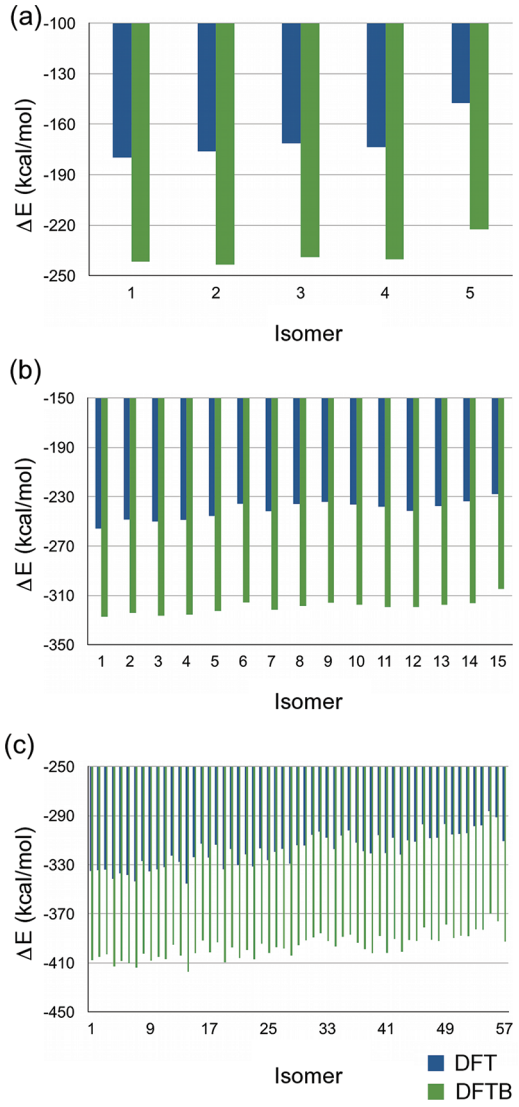

Figure 4. Comparison of relative stabilities calculated using Hückel with ΔE calculated using DFTB/DFT for (a) $C_{20}X_2/C_{20}(OH)_2$, (b) $C_{20}X_3/C_{20}(OH)_3$, and (c) $C_{20}X_4/C_{20}(OH)_4$. Isomer numbers (horizontal axes) are defined in terms of ascending Hückel β coefficients. DFTB/DFT relative energies ΔE are given in kcal/mol.

discrepancy is, in fact, more noticeable in this case. For instance, Figure 4(c) shows that according to DFT/DFTB, there are a number of structures, including the least stable Hückel structure (structure 57), that are energetically competitive with structure 1 (the most stable structure according to Hückel theory). Both DFT and DFTB predict that structure 14 is the most stable $C_{20}(OH)_4$ isomer. It is noted that this is consistent with the most stable $C_{20}H_4$ structure according to DFT/DFTB. It is therefore concluded that the principle factors determining the stabilities of hydrogenated and hydroxylated fullerenes are the same; *viz.* the disruption to π-conjugation and the relaxation of the cage structure. Effects such as hydrogen bonding present in the case of OH addition are expected to be secondary.

To elaborate upon this point, we resort to an EDA of $C_{20}(OH)_4$ in an analogous manner to that of the preceding section. In this case, eq 3 becomes

$$\Delta E = E_{\mathrm{def}} + E_{\mathrm{int}} = E_{\mathrm{def}}(C_{20}) + \sum_{i=1}^{n} [E_{\mathrm{def}}(OH)]_i + E_{\mathrm{int}} \quad (5)$$

That is, the total deformation energy may be partitioned into those corresponding to deformation of the C_{20} cage and the individual OH groups. The total interaction energy E_{int} may also be partitioned into the interaction between the deformed cage and all OH groups and the interaction between the OH groups themselves. Such partition however is not necessary in the current context as will be shown below. It is noted here that $E_{\mathrm{def}} \approx E_{\mathrm{def}}(C_{20})$, since the total $E_{\mathrm{def}}(OH)$ for any isomer never exceeds *ca.* 0.3 kcal/mol. Comparison of $E_{\mathrm{def}}(C_{20})$ and E_{int} is made in Figure 5(a). As was the case for $C_{20}H_4$, E_{int} for

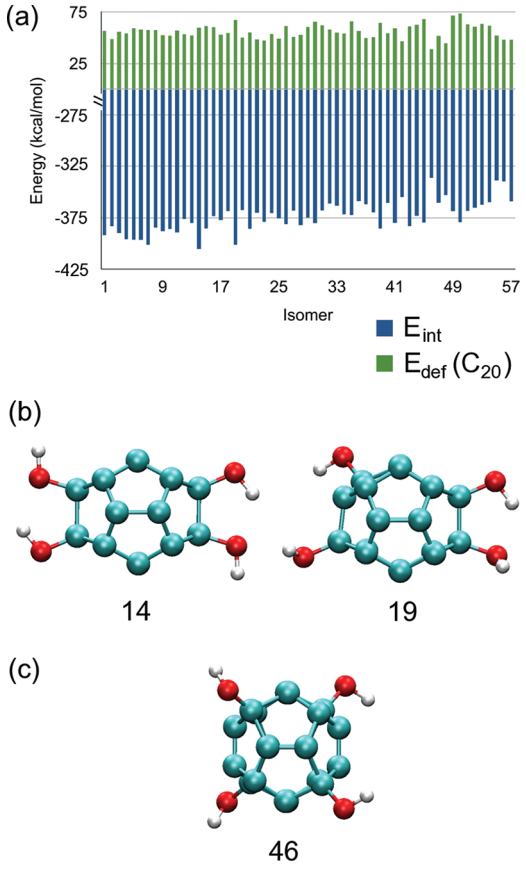

Figure 5. (a) EDA of $C_{20}(OH)_4$ functionalization. DFT $E_{\mathrm{def}}(C_{20})$ and E_{int} for 57 distinct isomers shows that the isomers with the most favorable ΔE are those with the largest E_{int}. However, hydrogen bonding between adjacent OH groups is seen to increase E_{int} substantially as shown in (b). Conversely, in the absence of these stabilizing hydrogen bonds, E_{int} is seen to decrease, as shown in (c).

$C_{20}(OH)_4$ shows a clear correlation with ΔE; *viz.* that the most stable isomers correspond to those with the largest E_{int}. Once again, $E_{\mathrm{def}}(C_{20})$ for $C_{20}(OH)_4$ shows no noticeable correlation with ΔE, as was the case for $C_{20}H_4$. It is also noted that the extent of cage deformation upon hydroxylation is essentially equivalent with hydrogenation; for $C_{20}H_4$ and $C_{20}(OH)_4$ exhibit average $E_{\mathrm{def}}(C_{20})$ values of 54 and 56 kcal/mol, respectively. As proposed above, hydrogen bonding between adjacent OH groups constitutes a secondary effect according to Figure 5. For instance, isomers 14 and 19 exhibit two and one

dx.doi.org/10.1021/ct300190u | J. Chem. Theory Comput. 2012, 8, 1841−1851

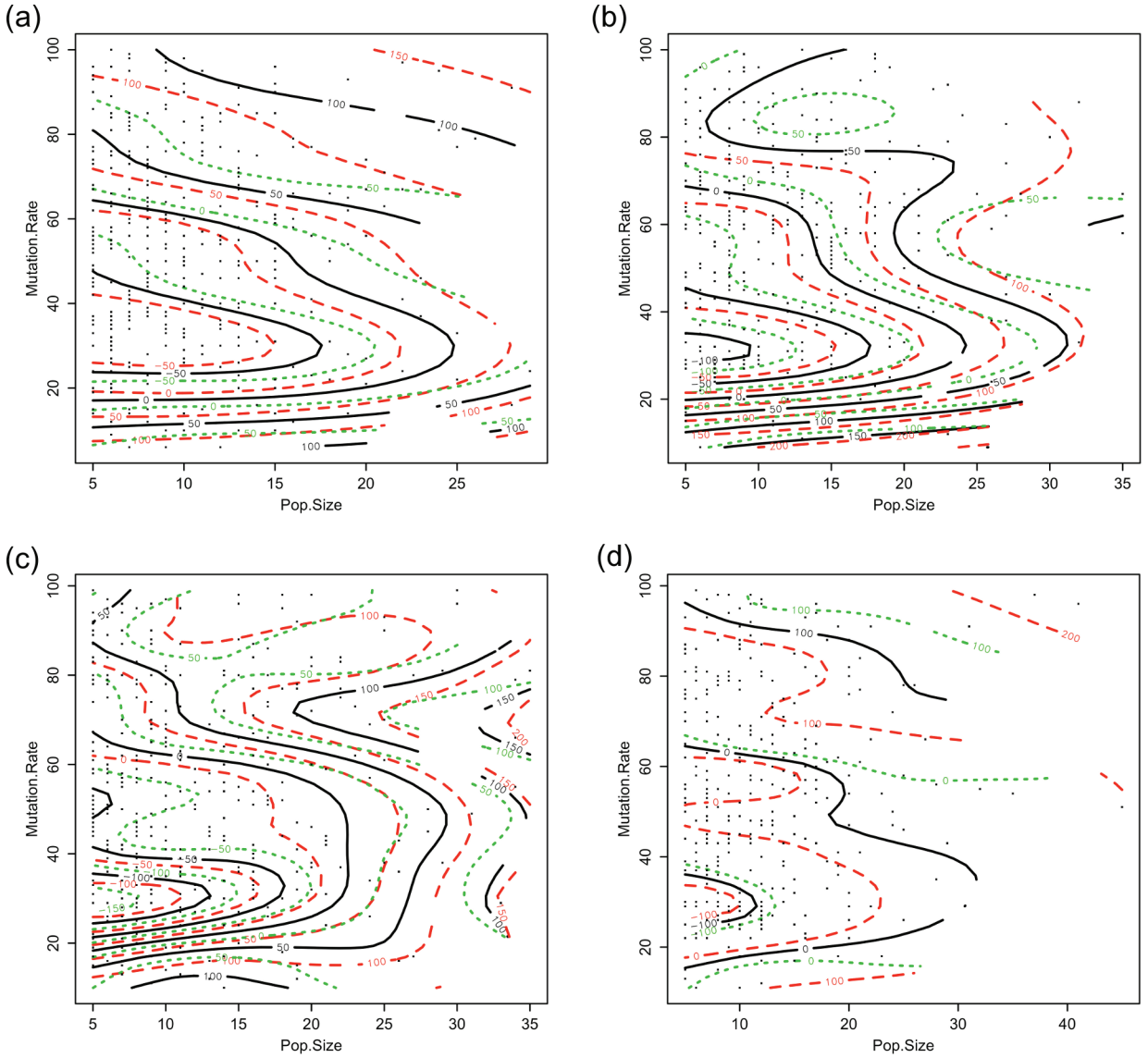

Figure 6. Effects of mutation rate and population size for (a) $C_{20}X_4$ (Hückel), (b) $C_{20}H_4$ (DFTB), (c) $C_{20}H_4$ (DFT), and (d) $C_{20}(OH)_4$ (DFTB) using the integer GA representation. Contour lines indicate M_{eff}, the measured mean evaluations as defined in eq 1. Black dots indicate the locations of data points.

hydrogen bonds between OH groups (Figure 5(b)); consequently, E_{int} is noticeably larger in these cases. Since $E_{int}(OH{-}OH) = E_{int} - E_{int}(C_{20}{-}OH)$, the stabilizing effect of hydrogen bonding in these cases is equivalent to −7.5 and −12.5 kcal/mol, respectively. On the other hand, isomer 46, which lack completely such hydrogen bonding (Figure 5(c)) exhibits an E_{int} value of −336.2 kcal/mol, one that is noticeably smaller compared to similar isomers (Figure 5(a)). $E_{int}(OH{-}OH)$ for isomer 46 is +29.3 kcal/mol; hydrogen bond stabilization therefore does not exist in this case.

3.3. GA Optimization: Comparison of Integer and Binary Representations. As it has been shown previously in the context of conformer searching,[32] population size and mutation rate have a strong effect on GA efficiency, while crossover probability is statistically insignificant. GA efficiency is defined here as the mean number of conformer evaluations required to locate a known lowest energy isomer (and its lowest energy conformer) of the functionalized fullerene. In each GA run, the number of individual evaluations is limited to the cost of an exhaustive search (4,845). The sole termination criterion used was the identification of the correct lowest energy conformer; any GA that failed to locate the lowest energy structure 100% of the time is deemed to be unreliable and so discarded. Complete details of the statistical model for each system are included in the Supporting Information (Tables S7−S16).

We turn now to comparing the two possible ways of encoding the structural positions of the functional groups, and their effect on the accuracy and efficiency of the resultant GA. Statistical analysis of the generated integer representation GAs are provided in Figure 6, in which the dependence of the GA efficiency on population size and mutation rate are shown. In each case (i.e., $C_{20}X_4$ (Hückel), $C_{20}H_4$ (DFTB/DFT), and $C_{20}(OH)_4$ (DFTB)), the most efficient GAs had a population size of *ca.* 10 individuals and a mutation rate of 0.3. Crossover rate did not have a significant effect on GA efficiency. The effect of the category variables is shown in Table 2. Elitism was found to be beneficial in every case except $C_{20}H_4$ using DFTB;

dx.doi.org/10.1021/ct300190u | J. Chem. Theory Comput. 2012, 8, 1841−1851

Table 2. Optimized Category Variables of the Integer GA and Their Effect on GA Efficiency with Respect to the Default Options[a]

parameter	$C_{20}X_4$ Hückel	$C_{20}H_4$ B3LYP/ 6-31G(d,p)	$C_{20}H_4$ DFTB	$C_{20}(OH)_4$ DFTB
evaluations	254	360	318	354
elitism = true	−62	−90	34	−37
mutation operator = integer Gaussian	196	168	209	161
crossover operator = none	−2	7	−5	23
selection operator = rank	−91	−96	−5	−83
N_{points}	355	359	318	314

[a]The default options are defined as the first option in each list, *viz.* elitism = false, mutation operator = integer range, crossover operator = 1-point, selection operator = tournament, scaling = linear. N_{points} is the number of 100% reliable GA parameter sets.

however, this result was less significant than the other cases. The swap and simple inversion mutators were found always to be unreliable, leaving only integer range and integer Gaussian mutators as alternatives. In all cases, use of the integer Gaussian mutator made the GA less efficient, requiring an extra 161−209 evaluations to locate the lowest energy structure. No crossover operator had a significant effect on GA efficiency. The rank selector was found to be the only selector that had a beneficial effect on the GA efficiency, reducing the number of evaluations required by 5−96. No scaling option had a statistically significant effect on GA efficiency, and so none of these options are considered further.

Figure 7 shows the dependence of the GA efficiency on population size and mutation rate for the binary encoded GA. Similarly to the integer representation, the most efficient GAs had a population size of *ca.* 10 individuals and a mutation rate of 0.3. There was no significant dependence on crossover rate. The effect of the category variables is shown in Table 3. Elitism was found to be beneficial in every case except $C_{20}H_4$ using

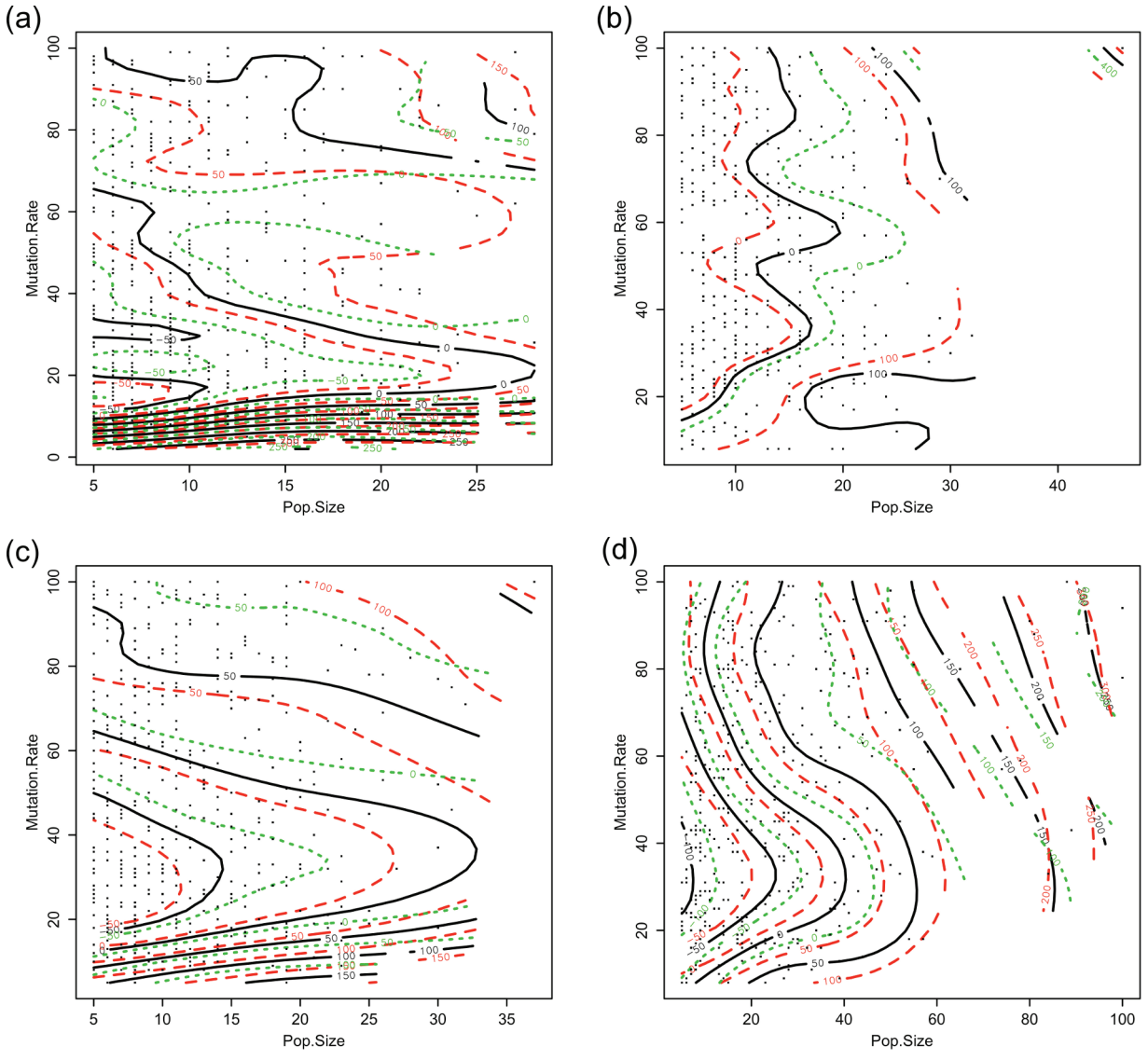

Figure 7. Smoother of mutation rate and population size for (a) $C_{20}X_4$ (Hückel), (b) $C_{20}H_4$ (DFTB), (c) $C_{20}H_4$ (DFT), and (d) $C_{20}(OH)_4$ (DFTB) using the binary GA representation. Contour lines indicate M_{eff}, the measured mean evaluations as defined in eq 1. Black dots indicate the locations of data points.

dx.doi.org/10.1021/ct300190u | *J. Chem. Theory Comput.* 2012, 8, 1841−1851

Table 3. Optimized Category Variables of the Binary GA and Their Effect on GA Efficiency with Respect to the Default Options[a]

parameter	$C_{20}X_4$ Hückel	$C_{20}H_4$ B3LYP/ 6-31G(d,p)	$C_{20}H_4$ DFTB	$C_{20}(OH)_4$ DFTB
evaluations	291	348	373	466
elitism = true	−16	−35	48	−26
mutation operator = flip	279	75	64	255
crossover operator = none	−71	−52	−74	−106
selection operator = rank	−92	−70	−24	−127
N_{points}	461	446	380	397

[a]The default options are defined as the first option in each list, *viz.* elitism = false, mutation operator = swap, crossover operator = 1-point, selection operator = tournament, scaling = linear.

DFTB; however, this result was only marginally significant. Only the swap and flip mutators were implemented for a binary string, and, as expected, the swap operator (which exchanges bits in two positions of the genome) was strongly preferred over the flip operator (which 'flips' a bit in the genome, i.e. turns a 0 to a 1 or vice versa). The latter is likely to create genomes that do not correspond to the correct number of functional groups. This problem may be alleviated somewhat by coupling it with a high mutation rate, thereby increasing the likelihood of an even number of flips. Nevertheless, where the flip operator did occur, it made the GA less efficient, requiring an extra 64−279 evaluations to locate the lowest energy isomer. In contrast to the integer representation, all choices for the crossover operator in the binary representation decreased the GA efficiency. Conversely, using no crossover whatsoever decreased the number of evaluations by 52−106. As with the integer representation, the rank selector was found to be the only selector that had a beneficial effect on the GA efficiency using the binary representation, reducing the number of evaluations required by 24−127. No scaling option had a statistically significant effect on GA efficiency.

3.4. GA Optimization: Usefulness of Problem-Specific Operators. All GA 'components' (i.e., crossover operators, mutators, selection operators) tested thus far have been standard, 'off-the shelf' routines, applicable to any representable problem. This represents a possible inefficiency in the GA, the remedy to which may be employing 'single-purpose' operators designed specifically for the underlying problem. Additionally, it is known that crossover operators that produce offspring whose phenotype is strongly different from the parent phenotype are ineffective compared to operators that maintain a high degree of similarity between parent and child phenotypes.[65] This problem is exacerbated when there are a large number of redundant representations (i.e., many genotypes correspond to a single phenotype).[66] Therefore, two additional tests were undertaken. First it was sought to test if injecting knowledge of the problem into the mutation operator could increase its effectiveness. The C_{20} fullerene was represented by its neighborlist, and functional groups were only permitted to mutate to a neighboring, unfunctionalized carbon atom. Second, two single-parent crossover operators, a shift operator and a flip operator were implemented. The shift operator selects a random point on the genome and shifts the beginning of the genome to the selected point, 'wrapping' around the end as necessary. The flip operator selects two random points of the genome and reverses the intervening section. All three operators were implemented for binary represented genomes only. GA optimization was repeated including the new operators in addition to the standard operators.

None of the new operators increased the efficiency of the derived GA, and therefore the derived GA remains the same. The neighborlist mutator was found to be less efficient than the integer range mutator, requiring 82−92 additional evaluations. The two single-parent crossover operators were the most efficient crossover operators tested, the single-parent shift operator required 61−67 less evaluations than the default 1-point crossover, and the single parent flip operator required 88−91 evaluations less. Neither of these operators, however, were more efficient than employing no crossover, which in these optimizations saved 97−104 evaluations. Complete details of these models are contained in the Supporting Information.

4. CONCLUSION

We have optimized a GA for the purpose of identifying the most stable isomers of arbitrary hydrogenated and hydroxylated fullerenes, and verified the predictions by performing exhaustive calculations of all possible isomers. The most efficient GAs do not employ crossover, thus reducing the GA to an Evolutionary Algorithm (EA). The optimized GA identifies without fail in each case the lowest energy isomer within ≈60 generations (≈300 calculations). As was previously shown in the context of conformational searching of floppy organic molecules,[32] the parameters defining the optimized GA were largely insensitive to the quantum chemical method employed in its optimization (in this case, DFTB and DFT). Indeed, the optimized GA reported here bears striking resemblance to that reported previously,[32] despite the disparity in the underlying chemical phenomena in question. We therefore recommend an EA with a small population, of 5−20 individuals, a moderate (≈0.3) mutation rate and rank selection with elitism as a suitable GA for general chemical search problems. We note here that the GA reported in this work was optimized using a pentagon-only fullerene cage. This is not expected to effect results in the context of larger fullerene cages that also contain hexagons. This follows from the fact that differences in reactivity are simply the manifestations of relative energies of two different functionalized structures. Since the GA reported here is optimized to converge the thermodynamically most stable structure, these differences in reactivity should be automatically taken into account. The efficiency of the GA reported here is evidently not decreased for larger, hexagon-containing fullerenes, following supplementary calculations using I_h-C_{60}.

An EDA of the factors underpinning hydrogenation and hydroxylation of the fullerenes themselves revealed that, for only few functional groups, one may resort to a simple π-Hückel description of electronic structure in order to understand the most stable functionalization pattern. However, this convenience breaks down relatively quickly in the case of C_{20} (at perhaps 3, and definitely 4 functional groups, which represents 25% functionalization of the cage). This was demonstrated to arise from the presence of non-π interaction between the functional groups and the deformed cage structure, which, by definition, are beyond simple π-Hückel theory. Nevertheless, the efficacy of such a simple, intuitive method in this context should be noted, as should the similarity between apparent trends in the energetics of functionalized fullerenes,

CNTs and graphene. It is evident from the latter observation that these seemingly disparate phenomena can be described in a common language.

■ ASSOCIATED CONTENT

Supporting Information

Extended discussions of the effect of temperature on the iterative MD/HMC optimization algorithm and statistical analysis of GA performance; comparison of $C_{20}H_4$ energies calculated using DFT/DFTB; Schlegel diagrams for all unique isomers of $C_{20}X_3$ and $C_{20}X_4$; lists of absolute energies for all unique isomers of $C_{20}H_n$ and $C_{20}(OH)_n$, n = 2, 3, 4 using Hückel/DFTB/DFT; extended details of the GA parametrization in both integer and binary representations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: M.A.A. madd@rsc.anu.edu.au; S.I. sirle@iar.nagoya-u.ac.jp.

Author Contributions

†These authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was in part supported by a CREST (Core Research for Evolutional Science and Technology) grant in the Area of High Performance Computing for Multiscale and Multiphysics Phenomena from the Japanese Science and Technology Agency (JST). S.I. acknowledges support from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan under the Strategic Programs for Innovative Research (SPIRE) and the Computational Materials Science Initiative (CMSI). Computer simulations were performed using The Academic Center for Computing and Media Studies (ACCMS) at Kyoto University the Institute for Molecular Science (IMS) in Okazaki, Japan and the National Computational Infrastructure (NCI) in Canberra, Australia. A.J.P. acknowledges the Kyoto University Fukui Fellowship. M.A.A. acknowledges the JSPS Fellowship Program.

■ REFERENCES

(1) Kroto, H.; Heath, J.; O'Brien, S.; Curl, R.; Smalley, R. *Nature* **1985**, *318*, 162−163.

(2) Chung, D.-J.; Seong, M.-K.; Choi, S.-H. *J. Appl. Polym. Sci.* **2011**, *122*, 1785−1791.

(3) Partha, R.; Conyers, J. *Int. J. Nanomed.* **2009**, *4*, 261−275.

(4) Montellano, A.; Da Ros, T.; Bianco, A.; Prato, M. *Nanoscale* **2011**, *3*, 4035−4041.

(5) Gao, J.; Wang, Y.; Folta, K.; Krishna, V.; Bai, W.; Indeglia, P.; Georgieva, A.; Nakamura, H.; Koopman, B.; Moudgil, B. *PloS One* **2011**, *6*, e19976.

(6) Hammond, M. R.; Kline, R. J.; Herzing, A. A.; Richter, L. J.; Germack, D. S.; Ro, H.-W.; Soles, C. L.; Fischer, D. A.; Xu, T.; Yu, L.; Toney, M. F.; DeLongchamp, D. M. *ACS Nano* **2011**, *5*, 8248−8257.

(7) Lee, R.-H.; Lee, L.-Y.; Huang, J.-L.; Huang, C.-C.; Hwang, J.-C. *Colloid Polym. Sci.* **2011**, *289*, 1633−1641.

(8) Economopoulos, S. P.; Karousis, N.; Rotas, G.; Pagona, G.; Tagmatarchis, N. *Curr. Org. Chem.* **2011**, *15*, 1121−1132.

(9) Ni, L.; Chang, W.; Hou, H.-L.; Li, Z.-J.; Gao, X. *Org. Biomol. Chem.* **2011**, *9*, 6646−6653.

(10) Mamlouk-Chaouachi, H.; Heinrich, B.; Bourgogne, C.; Guillon, D.; Donnio, B.; Felder-Flesch, D. *J. Mater. Chem.* **2011**, *21*, 9121−9129.

(11) Itami, K. *Chem. Rec.* **2011**, *11*, 226−235.

(12) Nagatsuka, J.; Sugitani, S.; Kako, M.; Nakahodo, T.; Mizorogi, N.; Ishitsuka, M. O.; Maeda, Y.; Tsuchiya, T.; Akasaka, T.; Gao, X.; Nagase, S. *J. Am. Chem. Soc.* **2010**, *132*, 12106−12120.

(13) Tan, Y.-Z.; Xie, S.-Y.; Huang, R.-B.; Zheng, L.-S. *Nat. Chem.* **2009**, *1*, 450−460.

(14) Ziegler, K.; Mueller, A.; Amsharov, K. Y.; Jansen, M. *J. Am. Chem. Soc.* **2010**, *132*, 17099−17101.

(15) Tan, Y.-Z.; Zhou, T.; Bao, J.; Shan, G.-J.; Xie, S.-Y.; Huang, R.-B.; Zheng, L.-S. *J. Am. Chem. Soc.* **2010**, *132*, 17102−17104.

(16) Nakae, T.; Matsuo, Y.; Nakamura, E. *Org. Lett.* **2008**, *10*, 621−623.

(17) Kuvychko, I. V.; Streletskii, A. V.; Shustova, N. B.; Seppelt, K.; Drewello, T.; Popov, A. A.; Strauss, S. H.; Boltalina, O. V. *J. Am. Chem. Soc.* **2010**, *132*, 6443−6462.

(18) Foloppe, N.; Chen, I. *Curr. Med. Chem.* **2009**, *16*, 3381−3413.

(19) Izgorodina, E. I.; Lin, C. Y.; Coote, M. L. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2507−2516.

(20) Bihlmeier, A.; Klopper, W. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1050−1059.

(21) Bihlmeier, A. *J. Chem. Phys.* **2011**, *135*, 044310.

(22) Bihlmeier, A.; Tew, D. P.; Klopper, W. *J. Chem. Phys.* **2008**, *129*, 114303.

(23) Luzan, S. M.; Tsybin, Y. O.; Talyzin, A. V. *J. Phys. Chem. C* **2011**, *115*, 11484−11492.

(24) Saunders, M. *J. Am. Chem. Soc.* **1987**, *109*, 3150−3152.

(25) Addicoat, M.; Metha, G. *J. Comput. Chem.* **2009**, *30*, 57−64.

(26) Assadollahzadeh, B.; Bunker, P. R.; Schwerdtfeger, P. *Chem. Phys. Lett.* **2008**, *451*, 262−269.

(27) Nunez, S.; Johnston, R. L. *J. Phys. Chem. C* **2010**, *114*, 13255−13266.

(28) Llanio-Trujillo, J. L.; Marques, J. M. C.; Pereira, F. B. *J. Phys. Chem. A* **2011**, *115*, 2130−2138.

(29) Oakley, M. T.; Wales, D. J.; Johnston, R. L. *J. Phys. Chem. B* **2011**, *115*, 11525−11529.

(30) Fuhrmann, J.; Rurainski, A.; Lenhof, H.-P.; Neumann, D. *J. Comput. Chem.* **2010**, *31*, 1911−1918.

(31) Pfeffer, P.; Fober, T.; Huellermeier, E.; Klebe, G. *J. Chem. Inf. Model.* **2010**, *50*, 1644−1659.

(32) Brain, Z. E.; Addicoat, M. A. *J. Chem. Phys.* **2011**, *135*, 174106.

(33) Grefenstette, J. J. *IEEE Trans. Systems, Man Cybern.* **1986**, *16*, 122−128.

(34) Friesleben, B.; Hartfelder, M. In *Artificial Neural networks and Genetic Algorithms*; Albrecht, R., Reeves, C., Steele, N., Eds.; Springer-Verlag: Heidelberg, 1993; pp 392−399.

(35) de Landgraaf, W. A. *Parameter Calibration Using Meta-Algorithms*, M.Sc. thesis, Artificial Intelligence Vrije Universiteit, Amsterdam, 2006

(36) Haupt, R. *Antennas and Propagation Society International Symposium, 2000. IEEE*; 2000; Vol. 2, pp 1034 −1037

(37) Zhang, Y.; Sakamoto, M.; Furutani, H. *Natural Computation, 2008. ICNC '08. Fourth International Conference*; 2008; Vol. 1; pp 70 −75

(38) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(39) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(40) 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. et al. Gaussian 09 Revision B.1; Gaussian Inc.: Wallingford, CT, 2009.

(41) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200−1211.

(42) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(43) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822−8824.

(44) Rao, L.; Ke, H.; Fu, G.; Xu, X.; Yan, Y. *J. Chem. Theory Comput.* **2009**, *5*, 86−96.

(45) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947−12957.
(46) Aradi, B.; Hourahine, B.; Frauenheim, T. *J. Phys. Chem. A* **2007**, *111*, 5678−5684, PMID: 17567110.
(47) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260−7268.
(48) Kohler, C.; Seifert, G.; Frauenheim, T. *J. Chem. Phys.* **2005**, *309*, 23−31.
(49) Duane, S.; Kennedy, A.; Pendleton, B. J.; Roweth, D. *Phys. Lett. B* **1987**, *195*, 216−222.
(50) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys.* **1992**, *97*, 2635−2643.
(51) Matito, E.; Feixas, F.; Solà, M. *J. Mol. Struct.: THEOCHEM* **2007**, *811*, 3−11.
(52) Page, A. J.; Buu, P. Q.; Witek, H.; Irle, S.; Morokuma, K. Submitted for publication.
(53) Brain, Z. E.; Addicoat, M. A. *Genetic and Evolutionary Computation Conference, GECCO 2010, Proceedings*; 2010; pp 823−824
(54) Brain, Z.; Addicoat, M. *Artificial Life XII, Twelfth International Conference on the Synthesis and Simulation of Living Systems*; 2010; pp 378−385.
(55) Perone, C. S. *SIGEVOlution* **2009**, *4*, 12−20.
(56) Wood, S. *Generalized additive models: an introduction with R*; CRC Press: 2006; Vol. *66*
(57) R Development Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, 2008.
(58) Wood, S. *J. Roy. Stat. Soc. B* **2000**, *62*, 413−428.
(59) Wood, S. *J. Roy. Stat. Soc. B* **2011**, *73*, 3−36.
(60) Zheng, G.; Wang, Z.; Irle, S.; Morokuma, K. *J. Am. Chem. Soc.* **2006**, *128*, 15117−15126, PMID: 17117863
(61) Morokuma, K. *J. Chem. Phys.* **1971**, *55*, 1236−1244.
(62) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325−340.
(63) Wang, Z.; Irle, S.; Zheng, G.; Morokuma, K. *J. Phys. Chem. C* **2008**, *112*, 12697−12705.
(64) Zheng, G.; Irle, S.; Morokuma, K. *Chem. Phys. Lett.* **2005**, *412*, 210−216.
(65) Rothlauf, F.; Goldberg, D. *Parallel Problem Solving from Nature PPSN VI* **2000**, 395−404.
(66) Rothlauf, F.; Goldberg, D. *Evol. Comput.* **2003**, *11*, 381−415.