

## Discrete Optimization of Electronic Hyperpolarizabilities in a Chemical Subspace

B. Christopher Rinderspacher,<sup>\*,†,‡</sup> Jan Andzelm,<sup>†</sup> Adam Rawlett,<sup>†</sup> Joseph Dougherty,<sup>†</sup>  
David N. Beratan,<sup>‡</sup> and Weitao Yang<sup>‡</sup>

*Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, Maryland 21005,  
and Department of Chemistry, Duke University, 124 Science Dr,  
Durham, North Carolina 27708*

Received June 26, 2009

**Abstract:** We introduce a general optimization algorithm based on an interpolation of property values on a hypercube. Each vertex of the hypercube represents a molecule, while the interior of the interpolation represents a virtual superposition (“alchemical” mutation) of molecules. The resultant algorithm is similar to branch-and-bound/tree-search methods. We apply the algorithm to the optimization of the first electronic hyperpolarizability for several tolane libraries. The search includes structural and conformational information. Geometries were optimized using the AM1 Hamiltonian, and first hyperpolarizabilities were computed using the INDO/S method. Even for small libraries, a significant improvement of the hyperpolarizability, up to a factor of ca. 4, was achieved. The algorithm was validated for efficiency and reproduced known experimental results. The algorithm converges to a local optimum at a computational cost on the order of the logarithm of the library size, making large libraries accessible. For larger libraries, the improvement was accomplished by performing electronic structure calculations on less than 0.01% of the compounds in the larger libraries. Alternation of electron donating and accepting groups in the tolane scaffold was found to produce candidates with large hyperpolarizabilities consistently.

### 1. Introduction

In recent years, organic molecules have garnered increasing attention as components of high-hyperpolarizability materials, partly due to the variety of synthetically accessible compounds, cost, and ease of processing.<sup>1,2</sup> Applications for materials with high hyperpolarizabilities are found in telecommunication and optics.<sup>3</sup> The dominant nonlinear response of organic molecules often finds its origin in the conjugated  $\pi$ -system, which facilitates the electronic polarizability. The design of such molecules *in silico* is complicated by the fact that chemical space, even constrained to smaller organic compounds, is combinatorially complex. The number of organic molecules of medium size is estimated<sup>4</sup> to be on the order of  $10^{200}$ . Enumeration is therefore unfeasibly costly,

and other methods for property optimization need to be developed. Including conformational searching further complicates molecular design.

Methods for optimization in discrete spaces have been studied extensively and recently reviewed.<sup>5</sup> Optimization methods include integer programming, as in branch-and-bound techniques (including dead-end elimination<sup>6</sup>), simulated annealing,<sup>7</sup> and genetic algorithms.<sup>8</sup> These algorithms have found renewed interest and application in molecular and materials design.<sup>9–12</sup> Recently, new approaches have been explored to embed discrete chemical space in continuous spaces to take advantage of continuous optimization techniques. These include, in particular, activities in our group on the linear combination of atomic potentials (LCAP)<sup>13–15</sup> method and the approach of von Lilienfeld,<sup>16–19</sup> using a grand-canonical ensemble strategy. Here, we further employ continuous optimization methods aimed at discovering structures with optimal properties.

\* Corresponding author e-mail: berend.rinderspacher@arl.army.mil.

<sup>†</sup> Army Research Laboratory.

<sup>‡</sup> Duke University.

The problem of discrete optimization in chemical space can be tackled by embedding the discrete space in a virtual continuous space, parametrized by a set of continuous variables. This strategy establishes a continuous path from one molecule to another. Such a space can be constructed by defining molecules as a succession of gradual replacements of an atom or molecular fragment by another. These fragment or atom placements may be arbitrary, but the satisfaction of valency rules may be desirable. For example, a hydrogen in CH<sub>4</sub> might be replaced by a halogen or a methyl group, each corresponding to a specific geometry (or ensemble of geometries), energy(ies), and property value(s). It is possible to construct a continuous transition between Hamiltonians for the chemical structures as was done for LCAP.<sup>13</sup> Equation 1 illustrates the procedure.

$$H(\lambda) = \sum_i \lambda_i H_i, \quad \sum_i \lambda_i = 1, \quad 0 \leq \lambda_i \leq 1 \quad (1)$$

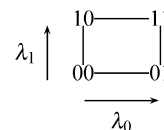
Each Hamiltonian  $H_i$  acts only on its own molecular subspace  $\Omega_i$ , and  $H$  acts on the direct sum of these spaces  $\oplus_i \Omega_i$ . In eq 1, the summation constraint implies the mutual exclusivity of the groups in the library (e.g., in the above example, as the hydrogen component increases toward 1, the halogen component decreases toward 0). In this approach, the groups are still linked through the wave function. Therefore, it is possible that all optima are at nonphysical configurations (e.g., half hydrogen and half halogen in the same location). Starting with allowed values of  $\lambda_i$  ( $0 \leq \lambda_i \leq 1$ ), it is possible to compute the numerical derivative of a property  $P$ . We now explore the application of this idea for discrete optimization of the first hyperpolarizability using differences of property values to replace continuous gradients.

## 2. Methods

**Linear Interpolation of Discrete Spaces.** Analogous to LCAP optimization, any property can in principle be interpolated in a virtual continuous space. We call the interpolated space “virtual” since noninteger  $\lambda_i$ -values correspond to intermediate or “alchemical” species. In general, given a library with  $N$  molecules with property values  $P_s$  for molecule  $s$ ,  $\lceil \log_2 N \rceil$  (the smallest integer larger than  $\log_2 N$ ) variables may be used to embed the discrete library in the continuous space. In LCAP, intermediate species contain contributions of each subspecies as well as cross-terms that arise from coupling *via* the wave function, which is one source of “virtual” optima. The values for intermediate species in this scheme are not contaminated by these cross-terms and depend only on the values at the real molecules. For example, assume a library consisting of methane, ethane, propane, and butane in exactly that order (see Figure 1). It is possible to interpolate among the 4 molecules using the parameters  $\lambda_0$  and  $\lambda_1$ . A (quadratic) polynomial interpolating the ground state energies (for example) is the following:

$$E(\lambda_0, \lambda_1) = E_0(1 - \lambda_0)(1 - \lambda_1) + E_1\lambda_0(1 - \lambda_1) + E_2(1 - \lambda_0)\lambda_1 + E_3\lambda_0\lambda_1 \quad (2)$$

Molecule	s	$\lambda_1\lambda_0$
CH <sub>4</sub>	0	0,0
C <sub>2</sub> H <sub>6</sub>	1	0,1
C <sub>3</sub> H <sub>8</sub>	2	1,0
C <sub>4</sub> H <sub>10</sub>	3	1,1



**Figure 1.** Simple example for interpolation. The bits  $\lambda_1\lambda_0$  represent the molecule number  $s = 2\lambda_1 + \lambda_0$  in the binary system.

This energy equation has a well-defined minimum when constrained to the square domain. Due to the domain constraints (only values on the square are allowed), the components of the gradient at the vertices pointing outside the square have to be dropped in pure gradient methods. At the minimum the gradient  $g$  points outside the square of definition for all components and thus is zero in the function's domain. Similarly, the Hessian  $J$  decomposes into normals, which point into the square or away. Again the constraints mandate that components of  $\Delta x = -J^{-1}g$  pointing out of the square are dropped for methods depending on  $\Delta x$ . Interpolation using a single variable for this set of compounds would produce a third degree polynomial, but homogeneous solutions to third order polynomials are not trivial, and the optimum is not guaranteed to correspond to a true molecule, i.e.,  $\lambda \in \{0, 1, 2, 3\}$ .

The preceding example highlights the dependence of the property polynomial on the ordering of the molecules. Generalization of the example to a library  $\mathcal{L}$  of size  $N$  leads to eqs 3 and 4. Equation 3 describes the bit-string (binary) representation of a number  $s$  with bit  $s(i)$  at the  $i$ th position.

$$s = \sum_{i=0}^{\lceil \log_2 N \rceil} s(i) \times 2^i, \quad s(i) \in \{0, 1\} \quad (3)$$

$$\tilde{P}(\lambda) = \sum_{s=0}^{N-1} P_s \prod_{i=0}^{\lceil \log_2 N \rceil} ((1 - \lambda_i)^{s(i)} \lambda_i^{1-s(i)}) \quad (4)$$

Equation 4 defines the property interpolation  $\tilde{P}$  based on the bit-strings. We differentiate between the interpolation function  $\tilde{P}$  and the set of discrete property values  $P_s$  to emphasize the domain of definition. The former is defined on the “virtual”, continuous hypercube  $([0, 1]^{\lceil \log_2 N \rceil})$ , while the latter is defined on the discrete space  $\mathcal{L}$ . This polynomial is continuous on the hypercube and has order  $\lceil \log_2 N \rceil$  and  $\lceil \log_2 N \rceil$  variables.

**Derivatives of  $\tilde{P}$ .** In order to use conventional optimization algorithms on continuous spaces, it is necessary to find the derivatives of  $\tilde{P}$ .

$$\frac{\partial \tilde{P}}{\partial \lambda_j}(\lambda) = \sum_{s=0}^{N-1} P_s (-1)^{s(j)} \prod_{b \neq j}^{\lceil \log_2 N \rceil} ((1 - \lambda_b)^{s(b)} \lambda_b^{1-s(b)}) \quad (5)$$

$$\frac{\partial^2 \tilde{P}}{\partial \lambda_k \partial \lambda_l}(\lambda) = \sum_{s=0}^{N-1} P_s (-1)^{s(k)+s(l)} \prod_{b \notin \{k,l\}}^{\lceil \log_2 N \rceil} ((1 - \lambda_b)^{s(b)} \lambda_b^{1-s(b)}) \quad (6)$$

Equations 5 and 6 show first and second order analytical derivatives of  $\tilde{P}$ . The derivative of  $\tilde{P}$  at  $\lambda$  corresponding to

the molecule with number  $s$  in the library  $\mathcal{L}$  can be computed from nearest bit-string neighbors (see eqs 7–10).  $s^{(j)}$  denotes the neighbor which differs only by the  $j$ th bit, while  $s^{(k,l)}$  identifies the neighbor which differs only in the  $k$ th and  $l$ th bits.

$$s^{(j)} = s + (-1)^{s^{(j)}} \times 2^j \quad (7)$$

$$s^{(k,l)} = s + (-1)^{s^{(k)}} \times 2^k + (-1)^{s^{(l)}} \times 2^l, \quad k \neq l \quad (8)$$

$$\lambda_i = s(i), \quad \frac{\partial \tilde{P}}{\partial \lambda_j}(\lambda) = (-1)^{s^{(j)}}(P_s - P_{s^{(j)}}) \quad (9)$$

$$\frac{\partial^2 \tilde{P}}{\partial \lambda_k \partial \lambda_l}(\lambda) = (-1)^{s^{(k)}}(-1)^{s^{(l)}}(P_s - P_{s^{(k)}} - P_{s^{(l)}} + P_{s^{(k,l)}}), \quad l \neq k, \quad \lambda_i = s(i) \quad (10)$$

The highly nonlinear, but continuous, function  $\tilde{P}$  allows the development of optimization methods by substituting derivatives by finite differences in continuous optimization methods. In this case, the analytical property derivatives for a molecule (i.e., at the vertices of the hypercube, where  $\lambda_i = s(i)$  for vertex  $s$ ) are simple (finite) property value differences, unlike in LCAP. The derivatives of LCAP need not be on straight lines pointing from one physical (non-“alchemical”) molecule to another, although the property values of each real molecule are the same for either optimization scheme. Formally,  $\tilde{P}$  is very similar to the Bayesian clustering approach, but no stochastic interpretation is needed in this case.<sup>30</sup> This framework also unifies some previous approaches.<sup>15,20</sup> Balamurugan et al.<sup>20</sup> have applied a best-first approach (BFA) to chemical optimization, which chooses the first substituent at a substitution site that improves the property. This method resembles the optimization algorithm employed in the latter sections in that the property improves at every step, but BFA uses the property value instead of the derivative. Keinan et al.<sup>15</sup> have used an algorithm which represents the steepest-descent method applied to  $\tilde{P}$ , as well as a line-search in which the direction of largest change is exclusively used. Unlike the other algorithms described, the steepest-descent method potentially jumps through the hypercube. While the following algorithm and Keinan’s line-search both traverse the edges of the hypercube constantly improving the property value, Keinan computes all single substitutions at every step.

**Comparison with Dead-End Elimination.** To compare our approach (eq 4) with dead-end-elimination algorithms (DEE), we consider the minimization of a pairwise additive property function comprised of single-parameter contributions  $P_i^{(\mu)}$  acting on site  $i$  with occupation  $\mu$  and double-parameter contributions  $P_{ij}^{(\mu,\nu)}$  acting on sites  $i$  and  $j$  with occupation  $\mu$  and  $\nu$  (eq 11).

$$P_s = \sum_i P_i^{s(i)} + \sum_{i < j} P_{ij}^{s(i),s(j)} \quad (11)$$

$$\begin{aligned} \tilde{P}(\lambda) = & \sum_i (P_i^{(0)}\lambda_i + P_i^{(1)}(1 - \lambda_i)) + \\ & \sum_{i < j} (P_{ij}^{(0,0)}\lambda_i\lambda_j + P_{ij}^{(1,0)}(1 - \lambda_i)\lambda_j + \\ & P_{ij}^{(0,1)}\lambda_i(1 - \lambda_j) + P_{ij}^{(1,1)}(1 - \lambda_i)(1 - \lambda_j)) \end{aligned} \quad (12)$$

Collecting all terms, we find a quadratic dependence of  $\tilde{P}$  on the pairwise terms  $P_{ij}$  with the parameters  $\lambda_i$  (eq 12). Consequently, the derivatives are linear with respect to  $\lambda_i$  (eq 13).

$$\frac{\partial \tilde{P}}{\partial \lambda_i} = P_i^{(0)} - P_i^{(1)} + \sum_{j \neq i} [(P_{ij}^{(0,0)} - P_{ij}^{(1,0)})\lambda_j + [P_{ij}^{(0,1)} - P_{ij}^{(1,1)}](1 - \lambda_j)] \quad (13)$$

From eq 13, a pruning argument for minimization can be derived, which is equivalent to the first-order DEE pruning rule applied to the special case of only two options at each site. Whenever the gradient with respect to a parameter  $\lambda_i$  is negative for all values of  $\lambda$  in the hypercube, then  $\lambda_i = 1$  minimizes  $\tilde{P}$ . This condition is precisely met when inequality 14 is fulfilled.

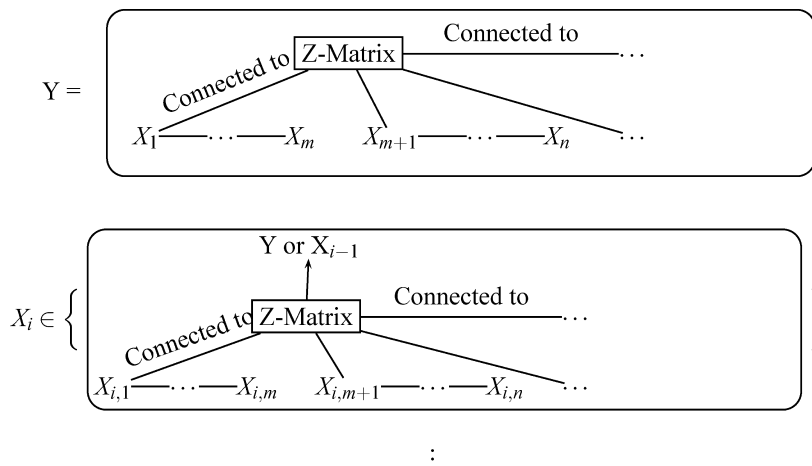
$$\frac{\partial \tilde{P}}{\partial \lambda_i} < 0 \Leftrightarrow P_i^{(0)} - P_i^{(1)} < \sum_{j \neq i} \min\{P_{ij}^{(1,0)} - P_{ij}^{(0,0)}, P_{ij}^{(1,1)} - P_{ij}^{(0,1)}\} \quad (14)$$

Conversely, a positive gradient (eq 15) implies that  $\lambda_i = 0$  minimizes  $\tilde{P}$ . Thus, it has been demonstrated that  $\tilde{P}$  naturally leads to DEE-like algorithms.

$$\frac{\partial \tilde{P}}{\partial \lambda_i} > 0 \Leftrightarrow P_i^{(0)} - P_i^{(1)} > \sum_{j \neq i} \max\{P_{ij}^{(1,0)} - P_{ij}^{(0,0)}, P_{ij}^{(1,1)} - P_{ij}^{(0,1)}\} \quad (15)$$

**2.1. Library Construction and Ordering.** The choice of enumeration of the library  $\mathcal{L}$  determines the assignment of specific molecules to  $\lambda$ . Consequently, this choice greatly influences the characteristics of  $\tilde{P}$ , such as its smoothness. Considering the example of Figure 1, the energy rises in all directions only for  $C_4H_{10}$ . But if  $CH_4$  and  $C_3H_8$  exchange places in the order, then going from  $C_3H_8$  to either of its two neighbors ( $C_2H_6$  or  $CH_4$ ) increases the energy, so that a “hurdle” has to be overcome to reach  $C_4H_{10}$ . Just exchanging the position of two neighboring molecules in the library changes the sign of the derivative at the corresponding  $\lambda$ . If the Hessian of the pairwise-additive property function is positive-semidefinite, the corresponding  $\tilde{P}$  is convex and optimization quickly reaches the global minimum. Using steepest gradient or Newton–Raphson algorithms locates property extrema (minima). It is beneficial to find an ordering of the library that produces a convex property surface. The linearity in each parameter  $\lambda_i$  implies convexity of  $\tilde{P}$  with respect to that parameter.

Assuming that molecules of similar structure have similar properties, a measure of similarity may be used to decrease the ruggedness/convexity of  $\tilde{P}$ . One choice to facilitate smooth property surfaces is the enumeration of molecules



**Figure 2.** Substitution pattern hierarchy. Y contains a Z-matrix that has several open “valences”. The first can be filled with substituents found in  $X_1$ , which are connected to substituents found in  $X_2$ , etc. The second is filled from  $X_m$  in the same manner. The  $X_i$  themselves are taken from a set of substitution patterns of the same kind as Y. Each instance is anchored to Y at the appropriate valence. The substitutions are terminated by Z-matrices that have no open valences.

by subsequent substitutions from a starting compound (see Figure 2). Returning to the example in Figure 1, Y contains the Z-matrix of  $\text{CH}_3$  with the connectivity information for  $X_1$ , which consists of the Z-matrix of H and  $\text{CH}_2X_{1,1}$  and connectivity information for  $X_{1,1}$  (level 1), which contains H and  $\text{CH}_2X_{2,1}$  (level 2), which finally contains H and  $\text{CH}_3$  (level 3). Evidence provided by the LCAP approach supports the supposition of smoothness when using substitutions.<sup>21</sup> The substitutions may be defined recursively; therefore, each level of a hierarchy of substitutions consists of a molecular fragment or atom to be connected to the next higher level, a list of substitution sites, and a set of subsequent levels for each site (see Figure 2). Each element of the set of subsequent levels is identified with a coefficient between 0 and 1, and the sum of these coefficients for each set must equal 1 (see eq 16). For a case in which more than two possible substitutions are available at a site, the bit-string representation must be extended to allow mixed numeric bases  $b_k$ . The general properties discussed in the preceding sections remain unchanged in this alternative interpolation (eq 18). The advantage of this description is the increase of convexity throughout a single substitution site.

$$\sum_j \lambda_{ij} = 1, \quad j \in \{0, \dots, b_i - 1\} \quad (16)$$

$$s = \sum_i \left( \prod_{k=0}^{i-1} b_k \right) \sum_{j \in b_i} s(i, j) \times j, \quad s(i, j) \in \{0, 1\}, \quad \sum_j s(i, j) = 1 \quad (17)$$

$$\tilde{P}(\{\lambda_{ij}\}_{j \in \{0, \dots, b_i\}, i}) = \sum_{s=0}^{N-1} P_s \left( \prod_i \prod_{j=0}^{b_i-1} \lambda_{ij}^{s(i,j)} \right) \quad (18)$$

**Inclusion of Multiple Conformational States.** For each molecule, it is important to find low-energy conformers for the property optimization to be physically meaningful. For each molecule in the molecular library, another optimization can be started with the (second) library consisting of the

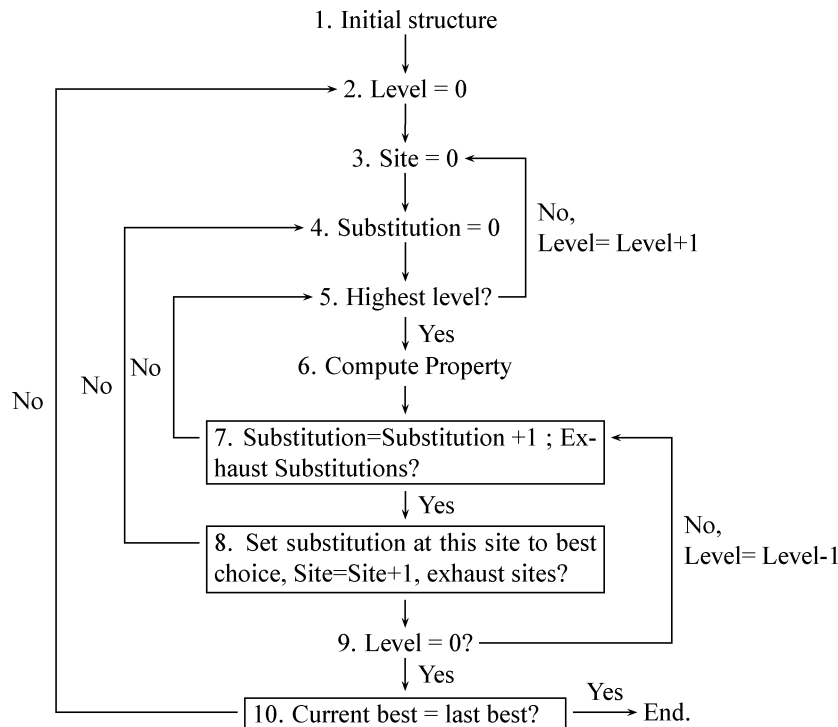
corresponding conformers. Each dihedral degree of freedom can be treated as a substitution site at the lowest level with a number of rotations as possible substitutions, as is commonly done in conformational searches.<sup>6,22</sup> In this manner, the conformational search can be introduced as the lowest level in the previously described substitution hierarchy. Thus, the conformational search precedes property computation in property optimizations. More general constraints on the optimal molecule can be introduced *via* alternate methods, like Lagrange multipliers or stochastic algorithms. Lagrange multipliers can be implemented using (soft) penalty functions with weightings that increase throughout the optimization.

**Algorithm.** Here, a line search algorithm is used; in particular, each parameter  $\lambda_i$  is followed to a minimum in that direction before varying the next parameter  $\lambda_{i+1}$ . Maximization *via* this algorithm can be achieved for instance by minimizing the negative objective function. This line search algorithm is an implicit branch-and-bound algorithm. A flowchart for the employed recursive algorithm appears in Figure 3, and application of the algorithm to a small example will be discussed in section 3 under the subsection Framework A (see also the accompanying Figure 6).

Since  $\tilde{P}(\lambda)$  is locally convex, this algorithm converges locally. The line-search steps 4–7 in Figure 3 correspond to a linear tree search or branch-and-bound algorithm. The computational complexity is on the order  $O(\log N)$  in the library size  $N$  due to the linear dependence on the  $\log N$  variables. In contrast to conventional branch-and-bound methods, no structures are explicitly excluded from the search space. Since each molecule chosen in step 8 in Figure 3 is strictly better in the sense of property optimization than its predecessor, the algorithm quickly converges to a local property value minimum in the library.<sup>20</sup>

All property minima for this algorithm are minima for the steepest-descent derived method and *vice versa*. This algorithm traverses the library in a smoother fashion compared to the steepest-descent derived method, successfully employed by Keinan et al.,<sup>15</sup> because the molecules are





**Figure 3.** Flowchart of the algorithm.

traversed variationally by single substitutions. While on one hand the steepest-descent based approach can sidestep barriers in the immediate vicinity efficiently, due to the simultaneous change of potentially several bits, the variational nature of this line search guarantees convergence, which is particularly useful on rugged property surfaces.

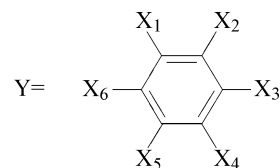
For the sake of computational accessibility, all geometries were optimized using the semiempirical Austin model 1 (AM1) method as implemented in *Gaussian03*.<sup>23</sup> The static electronic hyperpolarizability was computed using the INDO/S method as implemented in CNDO by Reimers et al.<sup>24</sup> using the sum-over-states expression in eq 19. The configuration interaction (CI) space was spanned by up to 100 unoccupied or occupied orbitals to accommodate for the large number of electrons in some of the investigated systems

$$\beta_{ijk} = \sum_{\nu\kappa} \frac{\langle 0|x_i|\nu\rangle\langle\nu|x_j - \mu_j|\kappa\rangle\langle\kappa|x_k|0\rangle}{E_{0\nu}E_{0\kappa}} \quad (19)$$

$$\beta_i = \frac{1}{3} \sum_j (\beta_{ijj} + \beta_{jij} + \beta_{jji}) \quad (20)$$

$$\beta_\mu = \frac{\vec{\mu}}{\|\vec{\mu}\|} \cdot \vec{\beta}, \beta_0 = \|\vec{\beta}\| \quad (21)$$

where  $E_{0\nu}$  is the excitation energy from the ground state to the  $\nu$ th excited state,  $\vec{\beta}$  is the static electronic hyperpolarizability with components  $\beta_i$  and corresponding hyperpolarizability tensor elements  $\beta_{ijk}$ ,  $\beta_0$  is the isotropic hyperpolarizability,  $\beta_\mu$  is the hyperpolarizability component in direction of the ground state dipole moment,  $\vec{x}$  is the dipole operator with components  $x_i$ , and  $\vec{\mu}$  is the ground state dipole moment with components  $\mu_i$ .



**Figure 4.** Top level of the substitution scheme (see Figure 2).

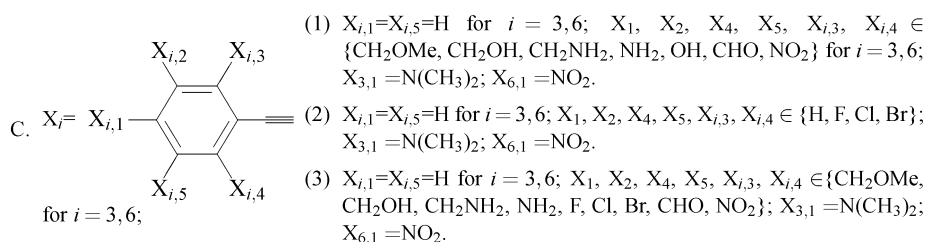
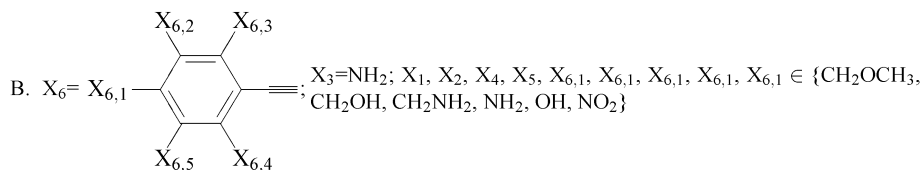
Figures 4 and 5 summarize the tolane-based system studies. Tolane spectroscopic properties are favorable for applications, so their first and second hyperpolarizabilities have been studied extensively.<sup>25,26</sup> In addition, these structures are readily modified<sup>27</sup> and present a large number of possible derivatives. Tolanes therefore present a particularly rich testbed for these optimization studies.

### 3. Results and Discussion

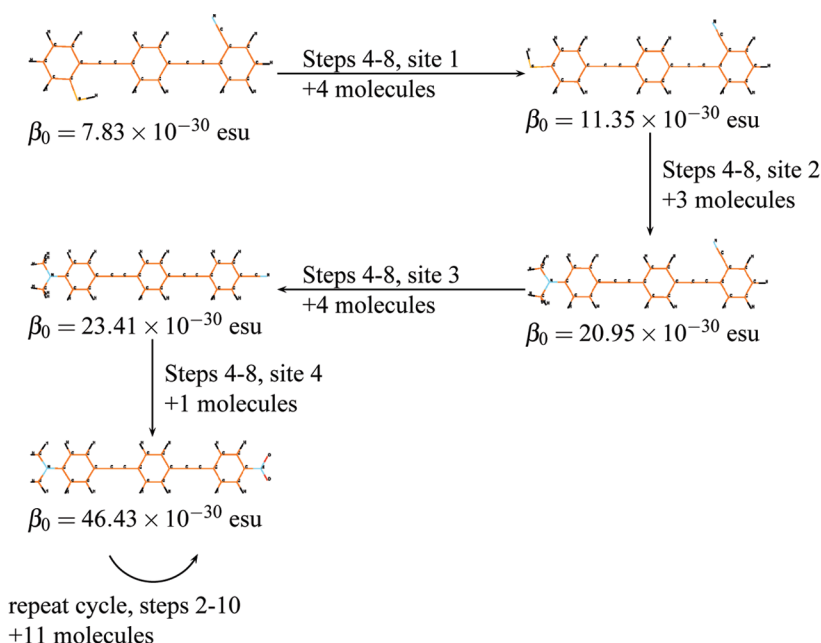
Overall, five different tolane libraries were investigated (general structure in Figure 2). The first three sets of molecules are optimized with respect to their static isotropic hyperpolarizability  $\beta_0$  (eq 21), while the remaining sets are optimized with respect to the component of the hyperpolarizability in direction of the dipole  $\beta_\mu$  (eq 21).

**Framework A.** Validation of the algorithm was performed on the structure framework A in Figure 5. Figure 6 shows the progress of the algorithm. There are 200 molecules in this library, but hyperpolarizabilities of only 24 different molecules were computed during the optimization, the minimum number of molecules required for the algorithm to finish the optimization. Regardless of the starting structure, the algorithm consistently finishes with the global hyperpolarizability optimum (Figure 6), which has also been confirmed experimentally.<sup>28</sup> For comparison, if the library

- A. Y as in Figure 4.  $X_1=X_2=X_4=X_5=H$ ;  $X_i=\equiv X_{i,1}-X_{i,2}$  for  $i=3,6$ ;  $X_{i,1} \in \{o-, m-, p-, o', m', \text{phenyl}\}$ ;  $X_{3,2} \in \{OH, SH, NH_2, NMe_2\}$ ;  $X_{6,2} \in \{CN, NO_2\}$



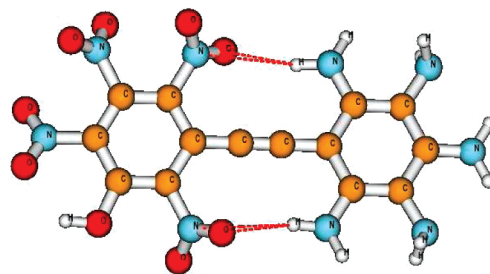
**Figure 5.** Tolane libraries investigated. Terminology as in Figure 2 with the top level as in Figure 4.



**Figure 6.** Progress of the optimization algorithm. The steps refer to the steps in Figure 3. The number of molecules indicated is the number of previously unvisited molecules for which the property is computed in performing the steps. Carbons are marked in orange, hydrogens in white, oxygens in red, and nitrogens in light blue.

is searched randomly, the expected number of computed molecules before finding the global minimum is 200 molecules. If repeats are avoided, then still 101 molecules would need to be computed on average in order to obtain the same result.

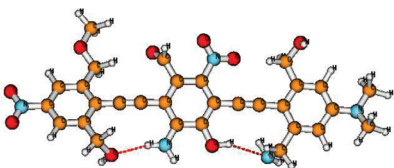
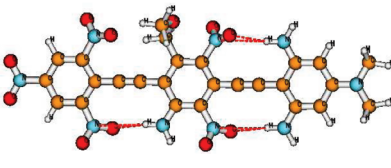
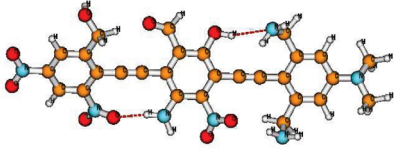
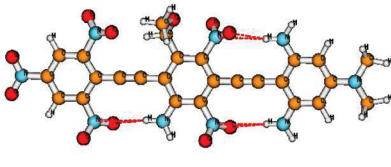
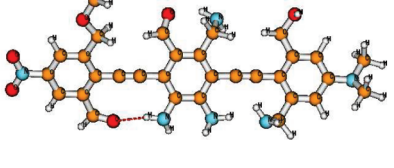
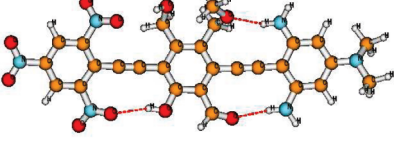
**Framework B.** The static hyperpolarizability  $\beta_0$  of framework B in Figure 5 optimizes to an unstable, perhaps explosive structure with mostly nitro- and amino-substituents (Figure 7). The final computed  $\beta_0$ -value was  $131.9 \times 10^{-30}$  esu after 121 computed structures from  $6^8 \approx 1.7 \times 10^6$  possible molecules. Additionally, conformational analysis was performed. CHO and OH were allowed two possible orientations in the plane of the tolane. For  $CH_2OH$  and  $CH_2NH_2$ , 3-fold rotation around the C–O and C–N bonds, respectively, was included, while only 2-fold rotations around the bonds connecting to the tolane framework were allowed.



**Figure 7.** Final structure of framework B. Carbons are marked in orange, hydrogens in white, oxygens in red, and nitrogens in light blue.

**Framework C-1.** The static hyperpolarizability for compounds in C-1 of Figure 5 was optimized starting from three different randomly chosen initial structures. A total of  $7^8 \approx 5.8 \times 10^6$  possible molecules exist in this family. Confor-

**Table 1.** Starting and Final Structures of Framework C-1 of Figure 5<sup>a</sup>

Run	Initial structure	Final structure
1		
2		
3		

<sup>a</sup> Carbons are marked in orange, hydrogens in white, oxygens in red, and nitrogens in light blue.

**Table 2.** Starting and Final Hyperpolarizabilities and Number of Computed Molecules for Framework C-1 in Figure 5<sup>a</sup>

run	initial $\beta_0/10^{-30}$ esu	final $\beta_0/10^{-30}$ esu	molecules computed
1	55.1	214.6	157
2	71.0	214.6	109
3	49.9	216.9	169

<sup>a</sup> See also Table 1.

mational considerations were treated as in framework B. Two of the three runs converged to the same structure ( $\beta_0 = 214.6 \times 10^{-30}$  esu), while the third converged to a second structure with comparable hyperpolarizability ( $\beta_0 = 216.9 \times 10^{-30}$  esu, see Tables 1 and 2). All three runs finished after computing less than 0.1% of all possible molecules and achieved 3- to 4-fold improvements of the hyperpolarizability. Comparing the two structures, we see some common motifs emerge: the variable fragments  $X_{3,3}$  and  $X_{3,4}$  contain nitro-groups, while  $X_{6,3}$  and  $X_{6,4}$  are occupied by amino-groups. Furthermore, positions  $X_2$  and  $X_4$  are occupied by electron acceptors, and sites  $X_1$  and  $X_5$  are occupied by electron donors. It is notable that not all positions are occupied by the “strongest” donors or acceptors in the substitution set, i.e.,  $\text{NH}_2$  and  $\text{NO}_2$ , respectively.

**Framework C-2.** Halogen substituents do not necessitate extensive conformational analysis, so they allow the evaluation of the optimization method without added constraints. The structures C-2 in Figure 5 were optimized for the hyperpolarizability in the direction of the dipole moment ( $\beta_\mu$ , see eq 21). Entries a and c in Table 3 show the results of two optimizations of framework C-2 in Figure 5 starting from the same initial structure with all substitutions set to hydrogens. In this case, convergence to a hyperpolarizability maximum is confirmed to be logarithmic in the library size; i.e., squaring the library size from 256 to 65536 leads to roughly twice the number of computed molecules.

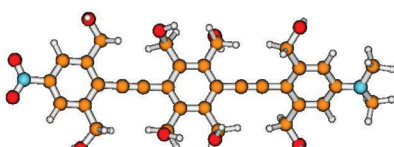
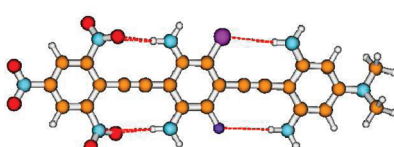
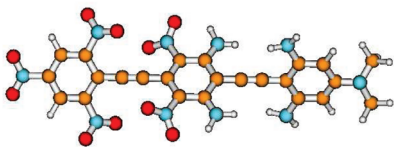
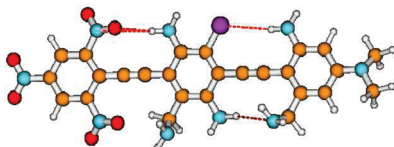
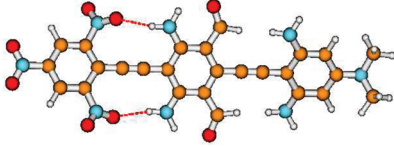
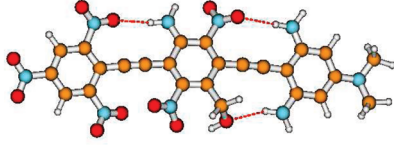
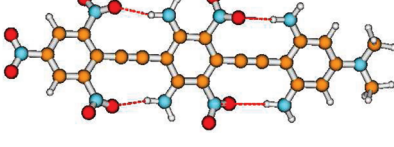
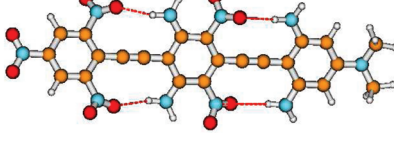
**Table 3.** Optimized Structures for Frameworks C-2 in Figure 5

	compound	$\beta_\mu/10^{-30}$ esu	molecules computed	library size
a	$X_1, X_5, X_{3,3}, X_{3,4} = \text{H}$ $X_{6,3}, X_{6,4} = \text{Br}$ $X_2 = \text{C1}$ $X_4 = \text{F}$	84.1	67	65536
b	$X_1, X_2, X_4, X_5, X_{3,3}, X_{3,4} = \text{H}$ $X_{6,3}, X_{6,4} = \text{Br}$	77.4	69	65536
c	$X_1, X_5 = \text{Br}$ $X_{3,3}, X_{3,4} = \text{H}$ $X_{6,3}, X_{6,4} = \text{Br}$ $X_2, X_4 = \text{F}$	83.5	28	256
d	$X_1, X_5, X_{3,3}, X_{3,4} = \text{H}$ $X_{6,3}, X_{6,4} = \text{Br}$ $X_2, X_4 = \text{Br}$	83.2	28	256

The stability of the optimization procedure was tested by constraining substitutions to be symmetric with respect to the mirror plane perpendicular to the plane of the backbone (runs (c) and (d) in Table 3), as well as starting from different initial structures: runs (a) and (c) were started with all substituents set to hydrogen, while run (b) starts from  $X_{6,3} = \text{Br}$  and  $X_{3,4} = \text{F}$ , and run (d) starts from  $X_{6,3} = X_{6,4} = \text{Br}$  and  $X_{3,3} = X_{3,4} = \text{F}$ . The hyperpolarizabilities of the initial structures were within 4 units of  $50 \times 10^{-30}$  esu. Since the procedure is not a global optimization algorithm, it is possible to end at different local maxima, here each run ended in a different structure with corresponding hyperpolarizabilities ( $\beta_\mu/10^{-30}$  esu = 84.1, 77.4, 83.5, 83.2, respectively, see Table 3). Nonetheless, the optimizations lead to significant and comparable improvements between runs. The found maxima all place bromine in the  $X_{6,3}$  and  $X_{6,4}$  positions, implying that a large fraction of the gain in  $\beta_\mu$  arises from bromine to amino charge transfer interactions.

**Framework C-3.** In combination with parts of libraries of C-1 and C-2 in Figure 5, structures C-3 in Figure 5 were subjected to optimization of the static hyperpolarizability in

**Table 4.** Starting and Final Structures of Framework C-3 in Figure 5<sup>a</sup>

Run	Initial structure	Final structure
1		
2		
3		
4		

<sup>a</sup> Carbons are marked in orange, hydrogens in white, oxygens in red, nitrogens in light blue, bromine in dark red, fluorine in dark blue, and chlorine in purple.

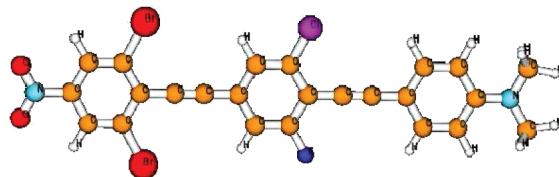
**Table 5.** Starting and Final Hyperpolarizabilities and Number of Computed Molecules for Framework C-3 in Figure 5<sup>a</sup>

run	initial $\beta_{\mu}/10^{-30}$ esu	final $\beta_{\mu}/10^{-30}$ esu	no. comp
1	37.0	181.5	181
2	55.6	171.6	153
3	139.8	191.6	161
4	173.3	173.3	65

<sup>a</sup> See also Table 4.

the direction of the dipole moment ( $\beta_{\mu}$ ). Four optimizations from different starting configurations were performed (see Tables 4 and 5 for results). The “unbiased” first optimization leads to a 5-fold increase in  $\beta_{\mu}$  ( $37.0 \rightarrow 181.5 \times 10^{-30}$  esu). The final structure (see Table 4) indeed is a mixture of the results for C-1 and C-2 in Figure 5. The second optimization was started with a structure concentrating equal numbers of donors on one side and acceptors on the other, analogous to the final structure of framework B in Figure 5. This starting structure exhibited only a marginally larger hyperpolarizability ( $55.6 \times 10^{-30}$  esu) than the “unbiased” starting structure, but optimized to an alternating donor–acceptor arrangement ( $171.6 \times 10^{-30}$  esu) that failed to reach the optimum found in the first optimization. The low hyperpolarizability is presumably due to the benzene rings twisting out of plane and reducing conjugation.

A biased starting point, with alternating donor and acceptor groups, leads to a marginally increased final hyperpolarizability ( $191.6 \times 10^{-30}$ ) over the first optimization. The attempt to exceed this value by substituting the “strongest” electron donors and acceptors,  $\text{NH}_2$  and  $\text{NO}_2$ , fails despite the fact that this structure is indeed a local maximum ( $173.3$

**Figure 8.** Largest  $\beta_{\mu}$  structure for framework C-2 in Figure 5. Compare to entry a in Table 3. Carbons are marked in orange, hydrogens in white, oxygens in red, nitrogens in light blue, bromine in dark red, fluorine in dark blue, and chlorine in purple.

$\times 10^{-30}$  esu). All four optimization runs finish computing less than 0.001% out of the possible  $9^8 \approx 4.3 \times 10^7$  molecules.

#### 4. Summary and Conclusion

We have introduced an embedding of discrete molecular spaces in a continuous space, similar to the embedding of discrete Hamiltonians in LCAP.<sup>21</sup> From the embedding, an optimization based on differentiation in the continuous space was developed. The embedding is based on the chemically intuitive ordering of molecules by substitutions. Assuming that single substitutions are small perturbations, the ordering also increases smoothness in the resultant continuous space. Although the framework is very general, it is limited to properties that can be derived and computed by defining substitution patterns as well as computational accessibility, such as binding problems, linear spectra, or stress–strain curves of molecules.



The theoretical framework transforms a discrete optimization problem into a continuous optimization problem, which then gives rise to a discrete optimization strategy. The theoretical complexity of the used line-search algorithm is  $O(\log N)$  in the library size  $N$ , and applications of the algorithm to a variety of conditions confirm the method's effectiveness. A design strategy for tolans of alternating donors and acceptors along a conjugated framework is suggested by the optimization results. Choosing a set of initial structures increases the likelihood of finding the global optimum. Further applications and improvements are under study including an extension to second-order derivative methods, probabilistic methods<sup>29,30</sup> and dynamic ordering of the parameters to achieve overall convexity.

**Acknowledgment.** We would like to thank S. Keinan, B. Desinghu, G. Lindsay, A. Chafin, and M. Davis for helpful discussions. We are thankful to DARPA Predicting Real Optimized Materials through ARO (W911NF-04-1-0243) and the Army Research Laboratory for funding.

### References

- Andrekson, P. A.; Westlund, M. *Laser Photonics Rev.* **2007**, *1*, 231–248.
- Bergmann, G.; Ellis, C.; Hindmarsh, P.; Kelly, S. M.; O'Neill, M. *Mol. Cryst. Liq. Cryst.* **2001**, *368*, 4439–4446.
- Dalton, L. R.; Sullivan, P. A.; Bale, D. H.; Bricht, B. C. Theory-inspired nano-engineering of photonic and electronic materials: Noncentro symmetric charge-transfer electro-optic materials. In *3rd Nano and Giga Forum*; Pergamon-Elsevier Science Ltd.: Oxford, U.K., 2007; pp 1263–1277.
- van Deursen, R.; Reymond, J.-L. *Chem. Med. Chem.* **2007**, *2*, 636–640.
- Michalewicz, Z.; Fogel, D. B. *How to Solve It: Modern Heuristics*; Springer Verlag: Berlin, 2002.
- Gordon, D. B.; Mayo, S. L. *J. Comput. Chem.* **1998**, *19*, 1505–1514.
- Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. *Science* **1983**, *220*, 671–680.
- Muhlenbein, H.; Gorgeschleuter, M.; Kramer, O. *Parallel Comput.* **1988**, *7*, 65–85.
- Franceschetti, A.; Dudiy, S. V.; Barabash, S. V.; Zunger, A.; Xu, J.; van Schilfgaarde, M. *Phys. Rev. Lett.* **2006**, *97*, 047202.
- Dudiy, S. V.; Zunger, A. *Phys. Rev. Lett.* **2006**, *97*, 046401.
- Franceschetti, A.; Zunger, A.; van Schilfgaarde, M. *J. Phys.: Condens. Matter* **2007**, *19*, 242203.
- Piquini, P.; Graf, P. A.; Zunger, A. *Phys. Rev. Lett.* **2008**, *100*, 186403.
- Wang, M. L.; Hu, X. Q.; Beratan, D. N.; Yang, W. T. *J. Am. Chem. Soc.* **2006**, *128*, 3228–3232.
- Keinan, S.; Hu, X. Q.; Beratan, D. N.; Yang, W. T. *J. Phys. Chem. A* **2007**, *111*, 176–181.
- Keinan, S.; Paquette, W. D.; Skoko, J. J.; Beratan, D. N.; Yang, W. T.; Shinde, S.; Johnston, P. A.; Lazo, J. S.; Wipf, P. *Org. Biomol. Chem.* **2008**, *6*, 3256–3263.
- von Lilienfeld, O. A.; Lins, R. D.; Rothlisberger, U. *Phys. Rev. Lett.* **2005**, *95*, 153002.
- von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *J. Chem. Phys.* **2005**, *122*, 014113.
- von Lilienfeld, O. A.; Tuckerman, M. E. *J. Chem. Phys.* **2006**, *125*, 154104.
- Marcon, V.; von Lilienfeld, O. A.; Andrienko, D. *J. Chem. Phys.* **2007**, *127*, 064305.
- Desinghu, B.; Yang, W.; Beratan, D. N. *J. Chem. Phys.* **2008**, *129*, 174105.
- Xiao, D.; Yang, W.; Beratan, D. N. *J. Chem. Phys.* **2008**, *129*, 44106.
- Izgorodina, E. I.; Lin, C. Y.; Coote, M. L. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2507–2516.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 03 Technical Report*; Gaussian Inc.: Wallingford, CT, 2004.
- Tejerina, B.; Reimers, J. *CNDO/INDO*; 2007 (accessed 08/08/2009); DOI 10254/nanohub-r3352.5.
- Liu, C. P.; Liu, P.; Wu, K. C. *Acta Chim. Sinica* **2008**, *66*, 729–737.
- Oliva, M. M.; Casado, J.; Hennrich, G.; Navarrete, J. T. L. *J. Phys. Chem. B* **2006**, *110*, 19198–19206.
- Traber, B.; Oeser, T.; Gleiter, R. *Eur. J. Org. Chem.* **2005**, 1283–1292.
- Nguyen, P.; Lesley, G.; Marder, T. B.; Ledoux, I.; Zyss, J. *Chem. Mater.* **1997**, *9*, 406–408.
- Hu, X.; Beratan, D. N.; Yang, W. *J. Chem. Phys.* **2008**, *129*, 064102.
- Mueller, T.; Ceder, B. *Phys. Rev. B* **2009**, *80*, 024103.

CT900325P