CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Computational design and selection of optimal organic photovoltaic materials
Authors	O'Boyle, Noel M.;Campbell, Casey M.;Hutchison, Geoffrey R.
Publication date	2011-08
Original Citation	O'Boyle, NM; Campbell, CM; Hutchison, GR; (2011) 'Computational Design and Selection of Optimal Organic Photovoltaic Materials'. Journal of Physical Chemistry C, 115 :16200-16210. doi: http://dx.doi.org/10.1021/jp202765c
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1021/jp202765c
Rights	Copyright © 2011 American Chemical Society
Download date	2026-11-06 12:51:20
Item downloaded from	https://hdl.handle.net/10468/748

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

THE JOURNAL OF PHYSICAL CHEMISTRY C

ARTICLE

pubs.acs.org/JPCC

Computational Design and Selection of Optimal Organic Photovoltaic Materials

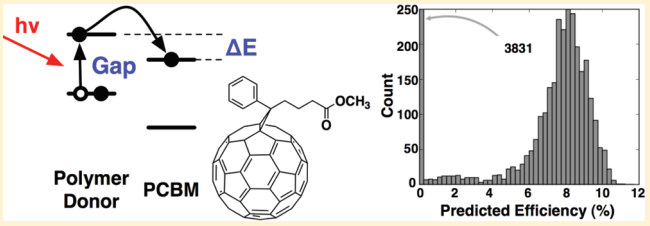
Noel M. O'Boyle,[†] Casey M. Campbell,[‡] and Geoffrey R. Hutchison*[,‡]

[†]Analytical and Biological Chemistry Research Facility, University College Cork, Western Road, Cork, Ireland

[‡]Department of Chemistry, University of Pittsburgh, 219 Parkman Avenue, Pittsburgh, Pennsylvania 15260, United States

Supporting Information

ABSTRACT: Conjugated organic polymers are key building blocks of low-cost photovoltaic materials. We have examined over 90 000 copolymers using computational predictions to solve the "inverse design" of molecular structures with optimum properties for highly efficient solar cells (specifically matching optical excitation energies and excited-state energies). Our approach, which uses a genetic algorithm to search the space of synthetically accessible copolymers of six or eight monomer units, yields hundreds of candidate copolymers with predicted efficiencies over 8% (the current experimental record), including many predicted to be over 10% efficient. We discuss trends in polymer sequences and motifs found in the most frequent monomers and dimers in these highly efficient targets and derive design rules for the selection of appropriate donor and acceptor molecules. We show how additional computationally intensive filtering steps can be used, for example, to eliminate targets likely to have poor hole mobilities. Our method effectively targets optimum electronic structure and optical properties far more efficiently than time-consuming serial experiments or computational studies and can be applied to similar problems in other areas of materials science.



INTRODUCTION

Organic photovoltaic materials offer the promise to significantly reduce the cost of solar electrical generation.[1−4] One particular design, the "bulk heterojunction" formed by intimate mixing of one electron donor and electron acceptor material,[5] has gained considerable attention due to ease of fabrication and synthetic chemical tailorability. Despite such promise, experimental improvements to the maximum energy conversion efficiency of bulk heterojunction photovoltaics have been slow over the last 4−5 years, with incremental improvements from ∼5 to 6%,[1,3,6−9] which have only recently been surpassed by ∼7−8%.[10] Several studies have focused on perceived inefficiencies of device engineering and basic physical and chemical materials limitations, and now, the prospect of low-cost, 10% efficient organic solar cells is renewed.[6,11−19]

On the basis of criteria such as optimizing the molecular optical absorption spectra with respect to the solar spectrum and minimizing voltage loss due to energy misalignment, suggested theoretical maximum efficiencies from a single cell design are around 11−12%.[11−13] These analyses, however, do not indicate which materials to target. Instead, they offer an "inverse design" problem that requires finding small molecules, oligomers, and polymers that match the electronic and optical properties required to produce a high-efficiency solar cell. Moreover, besides the optical absorption and energy level criteria, one may wish to incorporate other design criteria such as optical absorption intensity, electrical conductivity, loss mechanisms such as recombination, solubility, synthetic accessibility, processability, and solid-state structure.

A variety of work, particularly in the realm of computational drug design, has attempted to address the question of inverse design of molecular structure.[20−31] Chemical space (that is, the set of all stable molecules) is predicted to be vast, containing at least 10^{60} standard "organic-like" molecules.[32−35] Several approaches have been used for computational design of materials, including quantum mechanical operators to optimize atomic number.[21−27] These methods work best when refining an existing "scaffold" structure by mutating particular atoms or sets of atoms; unfortunately, in the realm of π-conjugated polymers, many such scaffolds exist. Furthermore, small changes to substitution patterns in polymers can drastically change the conformation and thus the electronic structure.[36,37]

In this work, we have studied over 90 000 π-conjugated copolymers to identify targets for next-generation efficient organic photovoltaics. In particular, we have concentrated on computationally efficient methods that serve as an initial filtering step in a pipeline model for materials discovery. Subsequent steps can refine candidates based on synthetic accessibility, crystal structure, charge mobility, and other properties. The approach is quite general and can be applied to similar problems in other fields of materials science, including finding chromophores with particular optical absorption or emission properties, electrochemical materials with specific oxidation or reduction profiles, or other similar electronic structure issues.

Received: March 24, 2011
Revised: May 9, 2011
Published: July 06, 2011

ACS Publications © 2011 American Chemical Society

Table 1. Monomer Index Number and IUPAC Name for All Monomers Used in This Study[a]

index	IUPAC name	index	IUPAC name	index	IUPAC name
0	thieno[3,4-*b*]thiophene 1,1-dioxide	44	cyclopenta[*c*]thiophen-4-one	88	2-benzothiophene-5,6-dicarbonitrile
1	3,4-dihydro-2H-thieno[3,4-*b*][1,4]dioxepine	45	furo[3,2-*b*]furan	89	thieno[3,4-d]pyridazine
2	1,3,4-oxadiazole	46	4,5,6,7-tetrahydro-2-benzothiophene	90	1,4-dihydropentalene
3	2,1,3-benzothiadiazole	47	buta-1,3-diyne	91	styrene
4	thieno[3,4-*b*][1,4]dithiine	48	3-methoxythiophene	92	5,6-dimethoxy-2-benzothiophene
5	4-nitrothiophen-3-amine	49	thieno[3,4-*b*]thiophene	93	thiophene-3-carboxylic acid
6	3-nitrothiophene	50	thieno[3,2-*b*]thiophene	94	4-methoxythiophen-3-amine
7	3-(4-fluorophenyl)thiophene	51	2-ethynyl-1H-pyrrole	95	2,3-dihydrothieno[3,4-*b*][1,4]dioxine
8	fulvene	52	1,3-dihydrothieno[3,4-*d*]imidazol-2-one	96	3,4-dinitrothiophene
9	1,3-dihydrothieno[3,4-*d*]imidazole-2-thione	53	2-methylideneindene	97	1H-thieno[3,4-*b*]pyrrole
10	thieno[3,4-*d*][1,3]dithiole	54	4-methylidenecyclopenta[*c*]thiophene	98	thieno[3,4-*b*]thiophene 1-oxide
11	pyridine	55	4-methoxythiophene-3-carbonitrile	99	3-thiabicyclo[3.2.0]hepta-1,4,6-triene
12	9H-carbazole	56	5,6-difluoro-2-benzothiophene	100	9H-carbazole
13	furan	57	4-hydroxythiophene-3-carboxylic acid	101	2,3-dihydrothieno[3,4-*b*][1,4]dithiine
14	1H-pyrrole	58	aniline	102	thieno[3,4-*d*]oxadiazole
15	2-benzothiophene	59	2,1,3-benzoxadiazole	103	2-ethenyl-1H-pyrrole
16	3,4-dimethylthiophene	60	thieno[3,4-*b*][1,4]dioxine	104	dibenzofuran
17	4H-dithieno[3,2-*b*:2′,3′-*d*]pyrrole	61	3,4-difluorothiophene	105	4-sulfanylthiophen-3-ol
18	benzene	62	thieno[3,4-*c*]furan-1,3-dione	106	1,3,4-thiadiazole
19	4H-cyclopenta[2,1-*b*:3,4-*b*′]dithiophene	63	thieno[3,4-*b*]furan	107	3-(4-methoxyphenyl)thiophene
20	5,6-dihydro-4H-cyclopenta[*c*]thiophene	64	3-(trifluoromethyl)thiophene	108	thieno[3,4-*d*][1,3]dioxole-2-thione
21	4H-cyclopenta[1,2-*b*:5,4-*b*′]bisthiophen-4-one	65	1,3-thiazole	109	thieno[3,2-b]thiophene
22	cyclopenta-2,4-dien-1-one	66	1H-thieno[3,4-*d*]imidazole	110	thieno[3,4-*c*]pyrrole-4,6-dione
23	pyrrolo[3,4-*f*]isoindole-1,3,5,7-tetrone	67	ethenyldiazene	111	thieno[3,4-*d*][1,3]dioxole
24	3-methylthiophene	68	1,2,3,4-tetrahydrothieno[3,4-*b*]pyrazine	112	3-(4-nitrophenyl)thiophene
25	cyclopenta[*c*]thiophene-4,6-dione	69	1,4-dihydrothieno[3,4-*b*]pyrazine-2,3-dione	113	1-thiophen-3-ylethanone
26	2,5-dihydropyrrolo[3,4-*c*]pyrrole-3,6-dione	70	thieno[3,4-*d*][1,3]thiazole	114	3-ethenylthiophene
27	thiophene-3-carbonitrile	71	furo[3,4-*f*][2]benzofuran-1,3,5,7-tetrone	115	4,5,6,7-tetrafluoro-2-benzothiophene
28	thieno[3,4-*b*][1,4]dithiine-2,3-dione	72	1,3-oxazole	116	1,4-dihydrothieno[3,4-*b*]pyrazine
29	but-1-en-3-yne	73	1H-imidazole	117	6H-cyclopenta[*c*]thiophene
30	cyclopenta[*b*]thiophen-4-one	74	3-ethenyl-4-methylthiophene	118	2-nitrothieno[3,4-*b*]thiophene
31	1,4-dihydropyrrolo[3,2-*b*]pyrrole	75	thiophen-3-ol	119	3,4-dimethoxythiophene
32	thieno[2,3-*f*][1]benzothiole	76	3-methoxy-4-(trifluoromethyl)thiophene	120	thiophene-3-thiol
33	2,2,2-trifluoro-1-thiophen-3-ylethanone	77	thieno[3,4-*d*][1,3]oxazole	121	5,6-dinitro-2-benzothiophene
34	thieno[3,4-*d*]thiadiazole	78	thieno[3,4-*b*]pyrazine	122	thiophene-3,4-dicarbonitrile
35	thieno[3,4-*b*][1,4]dioxine-2,3-dione	79	buta-1,3-diene	123	inden-2-one
36	2,3-dihydrothieno[3,4-*b*][1,4]oxathiine	80	thiophene-3-carbaldehyde	124	thieno[3,4-*d*]pyrimidine
37	4,7-dihydro-2-benzothiophene	81	thieno[3,4-*b*]thiophene-2-carbonitrile	125	thieno[3,4-*c*]thiophene-4,6-dione
38	thieno[3,4-*d*][1,3]dioxol-2-one	82	2-(4H-cyclopenta[2,1-*b*:3,4-*b*′]dithien-4-ylidene)propanedinitrile	126	methyl thiophene-3-carboxylate
39	4-(trifluoromethyl)thiophene-3-carbonitrile	83	fluoren-9-one	127	2-methylidenecyclobutane-1,3-dione
40	9H-fluorene	84	thieno[3,2-*b*]furan	128	3-phenylthiophene
41	fluoren-9-one	85	thieno[3,4-*d*][1,3]dithiole-2-thione	129	2-ethenylthiophene
42	thiophen-3-amine	86	2,3-dihydro-1H-thieno[3,4-*d*]imidazole	130	4-thiophen-3-ylaniline
43	2-ethenylpyridine	87	4H-thieno[3,2-*b*]pyrrole	131	thieno[3,4-*d*][1,3]dithiol-2-one

[a] More information is available in Table S1, Supporting Information.

We describe the computational methodology, which involves the use of a genetic algorithm (GA) to select the most promising candidates from this vast chemical structure space, including hundreds predicted to exhibit energy conversion efficiencies above 8% and many predicted to have efficiencies over 10%. We also analyze trends in copolymer sequences and key motifs in monomers and dimers.

COMPUTATIONAL METHODS

Monomer Data Set. The 132 monomers used in this study were selected from literature reports or obvious synthetic modification of other conjugated monomers (for example, substituted polythiophenes and phenylene vinylenes) and span a range of aromatic and conjugated species. For this study, monomers were

limited to species containing C, H, N, O, S, and F, and we restricted polymerization sites to those considered most synthetically likely. A range of electron-donating and electron-withdrawing substituents were considered. Table 1 lists the systematic names of the monomers, and all structures, systematic names, computed frontier orbital eigenvalues, excitation energies, and SMILES are given in Table S1, Supporting Information. The monomers span a wide range of electronic properties, with highest occupied molecular orbital (HOMO) eigenvalues from −10.72 to −6.94 eV and lowest unoccupied molecular orbital (LUMO) eigenvalues from −6.60 to −0.39 eV. For comparison, thiophene is computed to have a HOMO eigenvalue of −8.54 eV and a LUMO at −2.03 eV.

Generation of Optimized 3D Structures. The 3D structure of a polymer was generated using a multistep process starting with the SMILES string for the polymer.[38] An initial 3D structure was generated using Open Babel 2.2.3[39] (accessed through its Python interface Pybel)[40] and minimized using the MMFF94 force field[41−45] (500 steps using steepest descent minimization, convergence at 1.0^{-4} kcal/mol). Next a weighted-rotor search (MMFF94, 100 iterations, 20 geometry optimization steps) was carried out to find a low-energy conformer. This was then further optimized using MMFF94 (500 steps). Finally, Gaussian09 was used to optimize the structure using the PM6 semiempirical method.[46] The entire procedure required ∼8 min per oligomer on one CPU core.

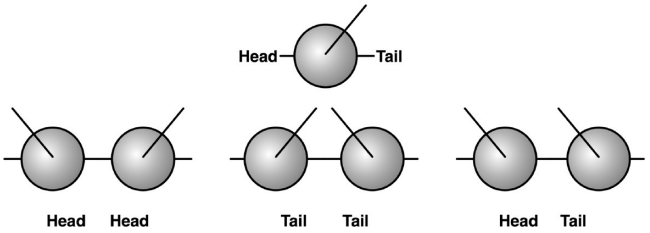
Prediction of Electronic Structure and Optical Excitation Energies. The energies and oscillator strengths of the 15 lowest-energy electronic transitions were calculated using the PM6-optimized geometry[46] using the ZINDO/S method[47] as implemented in Gaussian09.[48] The Python library cclib[49] was used to extract the molecular orbital eigenvalues as well as the energies and oscillator strengths of the electronic transitions. To determine the accuracy of this method, we computed the primary excitation energies across a 60-compound test set[50] and found a rms error of ±0.28 eV compared to experimental UV/vis peaks (Figure S1, Supporting Information). This is comparable to those obtained previously with AM1/ZINDO.[50] To test the accuracy of calculated orbital eigenvalues to predict ionization potentials using Koopmans' theorem, we compared the ZINDO orbital eigenvalue for the HOMO to experimental gas-phase vertical ionization energies for 100 compounds.[51] Results are included in the Supporting Information (Figure S2 and Table S2) and suggest that the ionization potentials show a systematic shift of ∼−0.4 eV relative to experiment; however, because our results compare the difference in two ionization potentials, these shifts will cancel.[52]

Synthetic Accessibility. To limit the possible search space and to concentrate on the most synthetically relevant species, we considered copolymers formed by preparing a dimer of two different monomers, followed by polymerization to make tetramers, hexamers, and octamers. Because some monomers are asymmetric (66 out of 132 monomers), head-to-head, tail-to-tail, and head-to-tail dimers were considered, giving a total of 19 701 distinct dimers. To form polymers of length 4, the base dimer was joined with itself either head-to-head, tail-to-tail, or head-to-tail (Scheme 1). This gave a total of 58 707 synthetically accessible tetramers, a value much smaller than the total space of tetramers of just over 768 million (including species with up to four different monomers).

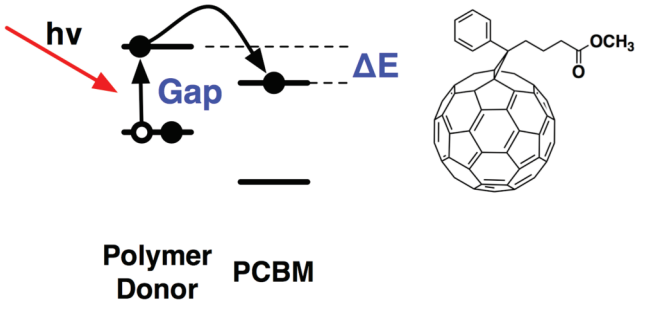
Calculation of Energy Conversion Efficiency. The energy conversion efficiency was calculated as described by Scharber et al.[12] This model predicts the energy conversion efficiency based on the properties of the polymer donor material, given the electronic structure of the electron acceptor (Scheme 2); we used the common acceptor material phenyl-C61-butyric acid methyl ester (PCBM). Assuming constant values for other parameters of the photovoltaic device, the efficiency may be improved by optimizing the HOMO−LUMO gap of the polymer donor and minimizing the energy difference ΔE between the LUMO of the polymer and PCBM. Some minimal ΔE is required to separate the electron−hole pair (the exciton binding energy), and here, we assume this value to be 0.3 eV, as suggested previously.[12] While more recent results suggest that a larger ΔE may be required due to vibrational relaxation and electrostatic effects at the polymer/acceptor interface,[53−55] the predictions of device efficiencies using existing models will have variability due to many factors, including the accuracy of the electronic structure methods, the solid-state packing and film morphology in the device, and several others. As discussed below, this work aims to provide an efficient mechanism to generate potential synthetic targets rather than a quantitative predictive model of solar cell efficiencies.

Where the model proposed by Scharber used the HOMO−LUMO gap to determine the lowest-energy electronic transition, we use the lowest energy of a ZINDO/S electronic transition. For a particular polymer, we chose the lowest-energy singlet transition with an oscillator strength greater than 1.0; in most cases, this corresponded to the first singlet excited state. Where none existed (among the 15 calculated), we used the transition with the maximum oscillator strength but scaled the resulting efficiency by the value of the oscillator strength to approximate the effect of a decreased optical extinction coefficient. The value used for the PCBM LUMO in this scheme is −4.61 eV, derived from a ZINDO/S calculation on a PM6-optimized PCBM structure by adding the energy of the lowest-energy transition (2.22 eV) to the HOMO (−6.83 eV). Under

Scheme 1



Scheme 2



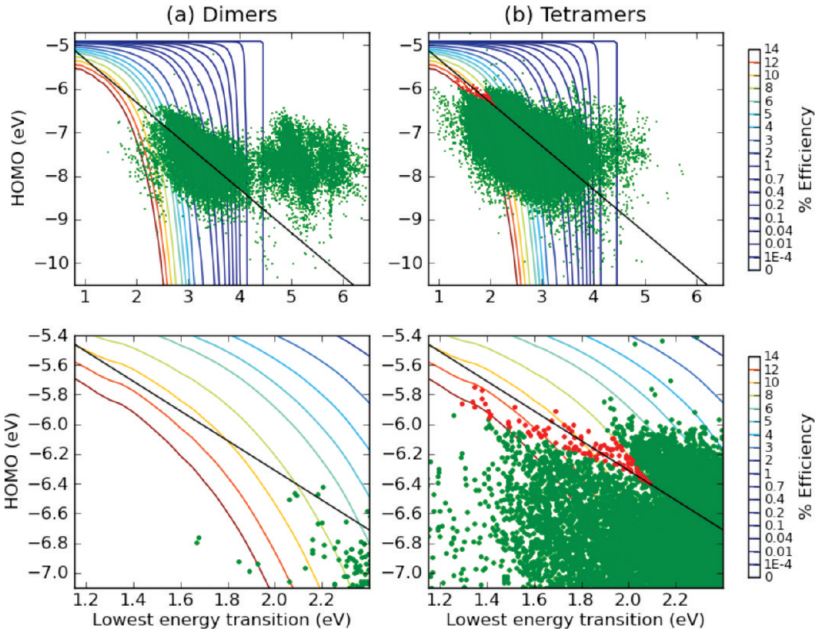


Figure 1. (a) The set of all possible dimers formed from the 132 monomers studied in terms of the calculated lowest-energy transition and HOMO and (b) the set of all synthetically accessible tetramers. The contour lines indicate the calculated value of the energy conversion efficiency. The bottom row shows the same data but focuses on the area of highest efficiency. Highlighted in red is the set of 101 "promising" tetramers (see text).

the assumption (see above) that a 0.3 eV energy difference between the energies of the excited state of the donor (polymer) and the acceptor (PCBM) is the minimum required to separate the charges, for a particular polymer, the sum of the HOMO plus the electronic transition must be greater than or equal to −4.31 eV to have any efficiency. Where this was not true, the efficiency was set to 0 as the polymer would not serve as an electron donor to PCBM. The efficiency "landscape" in terms of the HOMO and electronic transition is show as contours in Figure 1. The maximum value for the efficiency is 11.1%, which occurs at a HOMO of −5.7 eV and an electronic transition of 1.39 eV. This is the same maximum value found previously by Scharber et al.,[12] but the optimal value of the HOMO was instead −5.39 eV, consistent with the ∼0.4 eV shift between computed and experimental orbital energies described above.

For comparison between experiment and prediction, we calculated the efficiencies of the commonly used poly-3-hexylthiophene (P3HT) and poly[2-methoxy-5-(3′,7′-dimethyl-octyloxy)]-1,4-phenylene vinylene (MDMO-PPV) polymers using octamers as models for the longer polymers. The predicted efficiencies are 3.36 and 1.84% for P3HT and MDMO-PPV, respectively, compared to ∼5 and ∼3.5% in optimal experimental devices.[4,56,57] One main reason for the lower predicted efficiency was that the most stable conformations predicted by MMFF94 followed by semiempirical PM6 optimization are nonplanar. Semiempirical methods are known to poorly predict dihedral angles in conjugated species, decreasing the effective conjugation length, and thus overestimate excitation energies.[58−60] Consequently, our methodology would be improved by more accurate semiempirical parametrizations. Nevertheless, the calibration suggests that our computational method predicts energy conversion efficiencies of bulk photovoltaic devices from molecular properties to within ∼2%.

Genetic Algorithm. A GA is a stochastic method for global optimization. It is based on concepts from evolutionary biology, where a population of solutions (or chromosomes) is optimized in successive iterations (or generations) by applying the evolutionary operators of crossover, mutation, and selection. In this case, the chromosomes were the candidate polymers, and the objective function being minimized was the deviation from the desired HOMO and electronic transition energy, −5.7 and 1.39 eV, respectively, the values at the point of maximum efficiency (see above).

A key feature of our GA implementation was that the mutation operator mutated between monomers with similar electronic properties. To define similar, for each of the 132 monomers, we generated 3D structures (as described above) of the corresponding homopolymer of length 4 and carried out a ZINDO/S single-point calculation. Similar monomers were defined as those whose homopolymers had similar LUMOs and similar HOMO−LUMO gaps (measured by Euclidean distance). Full details of the GA are available in the Supporting Information.

RESULTS AND DISCUSSION

Dimers and Tetramers: Exhaustive Search and Analysis. The entire set of 19 701 synthetically accessible dimers and 58 707 tetramers were generated and their electronic properties calculated. The results are shown in Figure 1 along with contour lines showing their calculated efficiencies. The diagonal line separates molecules that have an excited state with sufficient energy to inject an electron into the acceptor, PCBM (top right), from those that do not (bottom left). The latter thus have a predicted efficiency value of 0.

For the dimers, the lowest-energy significant transition follows a bimodal distribution with a peak at ∼3.3 eV and another smaller peak at ∼5.0 eV. These transitions are shifted to lower energy for the tetramers due to the greater degree of conjugation and occur at a mean value of 2.8 eV. The mean value of the HOMO increases from −7.7 to −7.2 eV upon going from the

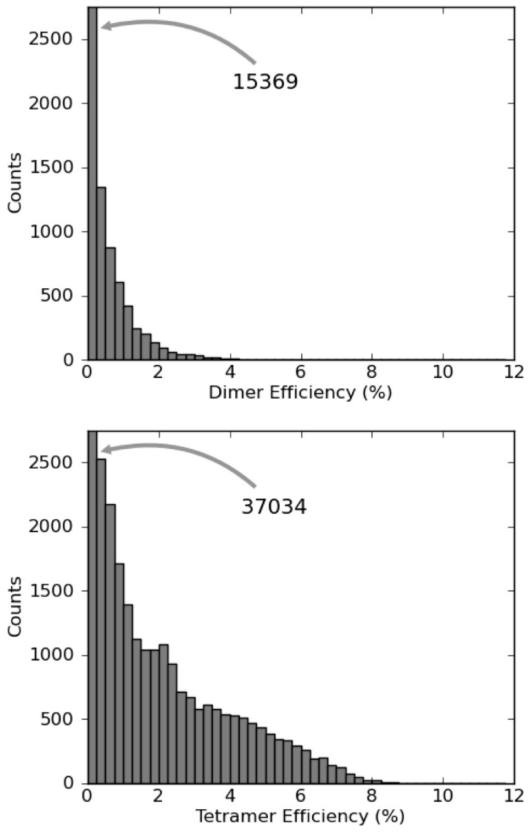


Figure 2. Histograms of predicted efficiencies for all dimers (top) and all tetramers (bottom). Note that almost all species are predicted to have low efficiency (<3%), and only 0.1% of all tetramers yield a predicted efficiency above 8%.

dimers to the tetramers. Taken together, both of these shifts move the electronic properties of the tetramers toward the region of highest efficiency at a HOMO of −5.7 eV and a transition energy of 1.4 eV. Figure 2 shows the distribution of efficiency values; there are 53 tetramers with efficiencies greater than 8% (12 greater than 9%, 4 greater than 10%) but only two dimers with efficiencies greater than 5% (with efficiencies of 5.7 and 6.1%).

A set of 101 "promising" tetramers was identified for training the GA (see below). These tetramers (shown in red in Figure 1) either have calculated efficiencies greater than 8% or lie just to the wrong side of the cutoff line for electron injection to PCBM, but close to the point of maximum efficiency. Given that the precise cutoff is not accurately known, these species should not be removed from future experimental consideration, especially as they may also be useful for photovoltaics with an acceptor material other than PCBM.

GA Development and Testing. It is clear that increasing the polymer length again beyond tetramers will yield a higher chance of finding molecular wires with the desired properties. However, an exhaustive search of the space would incur an excessive computational cost due to the increased size of the search space as well as the increased size of the molecules themselves. For example, the search space of hexamers is ∼78 000, and the space of octamers is ∼200 000, even with our restrictions to two-component copolymers and synthetically accessible sequences. To handle this problem, we implemented a GA (see the Computational Methods section). This is a stochastic algorithm for global optimization, and as such, it cannot guarantee to find the best solution, but it should find a good solution within a reasonable length of time.

The GA was parametrized based on its performance on the set of tetramers. Initially, the calculated efficiency was used as an objective function. It was found, however, that better results were obtained by minimizing the Euclidean distance to the point of maximum efficiency because the most efficient polymers lay close to the cutoff line for electron injection, and the polymers to which they were most structurally similar lay on the wrong side of the cutoff line. If efficiency was used as the objective function, such polymers were assigned an efficiency of 0 and thus would not have a chance to mutate further to the desired efficient polymers.

One of the key parameters of a GA is the number of chromosomes used. This is the number of candidate polymers retained for each iteration of the algorithm. The more chromosomes included, the better the coverage of the search space, but the longer the run time. On the basis of multiple runs using the tetramer data, a value of 64 chromosomes was chosen as the best compromise between performance and computational effort. Using these parameters, on average, based on 10 repetitions with different initial populations, the GA sampled 4.0% of the total space, but despite this, on average, it was able to find 7.2 of the top 10 and 58.7 of the 101 most promising tetramers.

GA and Local Search: Results on Hexamers and Octamers. After training on the tetramers, the GA was used to search for hexamers and octamers with suitable electronic properties. The results are depicted in Figure 3. To improve the likelihood of finding top candidates, the run of the GA was followed by a local search across all possible copolymers composed of the monomers that appeared most frequently during the GA. We believe this is a more efficient and effective way of finding oligomers missed by the GA than to run the GA multiple times because the GA will find the "top monomers" and the local search can ensure this smaller set of monomers is then explored exhaustively.

In the case of the hexamers, 3655 polymers were generated during the GA. For the local search, the top 29 most frequently occurring monomers were combined to give 2904 polymers, of which 1355 were novel, for a total of 5010. Similarly for the octamers, there were 4681 polymers generated during the GA. Combining the 21 most frequently occurring monomers for the local search gave 4259 polymers, of which 2504 were novel, for a total of 7185. The values of 29 and 21 were chosen so that the number of novel polymers to be calculated was around half of the number generated during the GA.

A total of 85 hexamers were found with efficiencies greater than 9%, 10 greater than 10%, and 1 greater than 11%. For the octamers, the corresponding figures were 524, 79, and 1. Electronic properties and component dimers of the top 10 hexamers and octamers are given in Tables S4 and S5 (Supporting Information).

From Figure 2, it is clear that despite the exhaustive calculation of all dimers and tetramers, the vast majority of those oligomers are predicted to have low efficiency (<3%). In almost all cases (>99%), trial-and-error or exhaustive searching, particularly of smaller oligomers, will largely focus on species with poor power efficiencies. In comparison, the distributions for the hexamers and octamers in Figure 4, with their peaks in the region of high efficiency, show that the use of a GA is a much more effective way to identify high-efficiency polymers.

Analysis of Oligomer Composition and Sequence. The top 25 hexamers and octamers are given in Tables 2 and 3, respectively, ordered by predicted energy conversion efficiency. All of

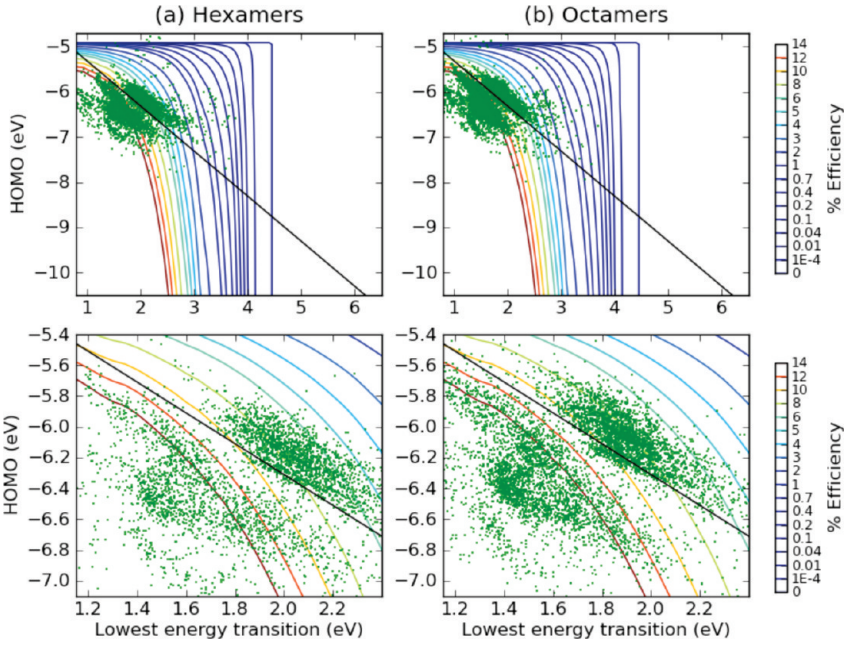


Figure 3. The electronic features and calculated energy conversion efficiency of (a) the hexamers and (b) the octamers generated during the GA and subsequent local search. The bottom row illustrates the same data but focuses on the area of highest efficiency.

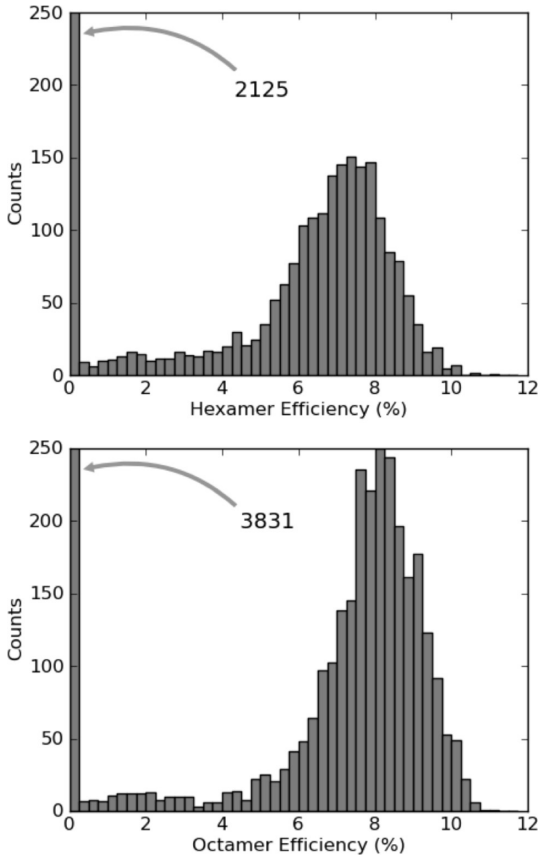


Figure 4. Histogram of predicted efficiencies for hexamers (top) and octamers (bottom) derived from the GA and subsequent local search. Note that while the full distributions of dimers and tetramers include vast numbers of low-efficiency species, the GA results for hexamers and octamers produce a distribution that has a peak at high efficiency.

the top 25 hexamers have predicted efficiencies above 9.5%, and the top 25 octamers, with their increased conjugation length, have predicted efficiencies above 10.3%.

To analyze the composition of donor and acceptor components in these top oligomers and consider effects of sequence on predicted efficiency, a novel visualization of "molecular barcodes" is used (Tables 2 and 3). A color is assigned based on the HOMO or LUMO eigenvalue of each monomer. For HOMO barcodes, the color scale ranges from red (electron-deficient monomers with high ionization potential) to blue (electron-donating monomers with low ionization potentials), with the midpoint set by thiophene. For LUMO barcodes, the color scale ranges from yellow to green based on the excited-state energy, again with white as the midpoint set by thiophene.

One general "design rule" for tailoring the band gap of π-conjugated polymers is to use a copolymer composed of an electron-donor monomer (i.e., easy to oxidize) and an electron-acceptor monomer (i.e., easy to reduce). A misconception about this approach is that the resulting HOMO of the copolymer will coincide with the HOMO of the donor, and the LUMO of the copolymer will coincide with the LUMO of the acceptor, as illustrated in Scheme 3a. This behavior, however, will only occur in the limit of no electronic interaction between the donor and the acceptor. Consequently, one expects the top oligomers to reflect a combination of the electronic structures of donors and acceptors, as illustrated in Scheme 3b.

Instead, the molecular barcodes in Tables 2 and 3 illustrate that the top oligomers exhibit only small differences in electronic structure between donor and acceptor monomers. In fact, relative to thiophene, almost all monomers present in the top oligomers are electron-donating (some shade of blue for the HOMO; some shade of green for the LUMO). This tendency toward low ionization potentials is more obvious in the barcodes than the trend in LUMO/excited state energies, which tend toward pale color shades near the thiophene midpoint.

Table 2. Predicted Efficiencies, HOMO and LUMO Eigenvalues (in eV), and Colored Molecular Barcodes for the Top 25 Hexamers and Two Reference Compounds in Scheme 4, Using ZINDO-Computed Monomer Orbital Eigenvalues (in eV)[a]

Global Color Scales	Minimum	Mid-Point	Maximum
HOMO Scale (in eV)	-10.72	-8.54	-6.94
LUMO Scale (in eV)	-6.60	-2.03	-0.39

Entry	Efficiency	HOMO	LUMO	HOMO Barcode (energies in eV)						LUMO Barcode (energies in eV)					
1	11.02	-5.75	-4.31	-7.89	-7.98	-7.89	-7.98	-7.98	-7.89	-1.81	-2.68	-1.81	-2.68	-2.68	-1.81
2	10.59	-5.88	-4.29	-7.98	-8.35	-7.98	-8.35	-8.35	-7.98	-2.68	-2.15	-2.68	-2.15	-2.15	-2.68
3	10.58	-5.83	-4.28	-7.98	-8.24	-8.24	-7.98	-8.24	-7.98	-2.68	-2.24	-2.24	-2.68	-2.24	-2.68
4	10.18	-6.03	-4.29	-7.98	-8.32	-8.32	-7.98	-8.32	-7.98	-2.68	-2.24	-2.24	-2.68	-2.24	-2.68
5	10.10	-6.06	-4.30	-7.29	-7.60	-7.60	-7.29	-7.29	-7.60	-2.18	-1.47	-1.47	-2.18	-2.18	-1.47
6	10.10	-5.98	-4.27	-7.98	-7.34	-7.34	-7.98	-7.34	-7.98	-2.68	-0.39	-0.39	-2.68	-0.39	-2.68
7	10.09	-6.01	-4.28	-7.98	-8.27	-8.27	-7.98	-8.27	-7.98	-2.68	-2.24	-2.24	-2.68	-2.24	-2.68
8	10.02	-5.74	-4.23	-8.35	-7.98	-7.98	-8.35	-7.98	-8.35	-2.15	-2.68	-2.68	-2.15	-2.68	-2.15
9	10.01	-6.03	-4.28	-7.98	-7.48	-7.48	-7.98	-7.48	-7.98	-2.68	-2.43	-2.43	-2.68	-2.43	-2.68
10	10.01	-6.06	-4.29	-7.03	-7.54	-7.03	-7.54	-7.54	-7.03	-3.96	-1.37	-3.96	-1.37	-1.37	-3.96
11	9.97	-6.05	-4.29	-7.03	-7.54	-7.03	-7.54	-7.03	-7.54	-3.96	-1.37	-3.96	-1.37	-3.96	-1.37
12	9.94	-6.05	-4.28	-7.60	-7.29	-7.29	-7.60	-7.29	-7.60	-1.47	-2.18	-2.18	-1.47	-2.18	-1.47
13	9.82	-5.89	-4.23	-8.27	-7.35	-8.27	-7.35	-8.27	-7.35	-2.24	-2.05	-2.24	-2.05	-2.24	-2.05
14	9.78	-5.98	-4.25	-7.98	-6.93	-7.98	-6.93	-6.93	-7.98	-2.68	-3.74	-2.68	-3.74	-3.74	-2.68
15	9.75	-6.03	-4.26	-7.29	-8.39	-7.29	-8.39	-7.29	-8.39	-2.18	-2.18	-2.18	-2.18	-2.18	-2.18
16	9.74	-5.99	-4.25	-8.01	-7.89	-7.89	-8.01	-7.89	-8.01	-2.85	-1.81	-1.81	-2.85	-1.81	-2.85
17	9.72	-6.10	-4.29	-8.21	-7.35	-8.21	-7.35	-7.35	-8.21	-1.93	-2.05	-1.93	-2.05	-2.05	-1.93
18	9.71	-6.11	-4.29	-7.89	-7.54	-7.89	-7.54	-7.54	-7.89	-1.81	-1.37	-1.81	-1.37	-1.37	-1.81
19	9.70	-6.01	-4.25	-8.01	-7.48	-7.48	-8.01	-7.48	-8.01	-2.85	-2.43	-2.43	-2.85	-2.43	-2.85
20	9.69	-6.04	-4.26	-8.32	-7.29	-8.32	-7.29	-7.29	-8.32	-2.04	-2.18	-2.04	-2.18	-2.18	-2.04
21	9.69	-6.12	-4.29	-7.35	-8.39	-7.35	-8.39	-7.35	-8.39	-2.05	-2.18	-2.05	-2.18	-2.05	-2.18
22	9.69	-6.14	-4.30	-7.29	-7.60	-7.29	-7.60	-7.60	-7.29	-2.18	-1.47	-2.18	-1.47	-1.47	-2.18
23	9.69	-6.09	-4.28	-7.03	-7.54	-7.54	-7.03	-7.03	-7.54	-3.96	-1.37	-1.37	-3.96	-3.96	-1.37
24	9.66	-6.08	-4.27	-7.54	-7.03	-7.03	-7.54	-7.03	-7.54	-1.37	-3.96	-3.96	-1.37	-3.96	-1.37
25	9.65	-5.95	-4.23	-7.54	-7.48	-7.48	-7.54	-7.48	-7.54	-1.37	-2.43	-2.43	-1.37	-2.43	-1.37
26	0	-7.57	-5.25	-10.32	-8.32	-10.32	-8.32	-10.32	-8.32	-4.97	-2.24	-4.97	-2.24	-4.97	-2.24
27	0	-6.97	-5.15	-9.38	-8.32	-9.38	-8.32	-9.38	-8.32	-4.00	-2.24	-4.00	-2.24	-4.00	-2.24

[a] The midpoint value is established from thiophene. Note that top candidates show very little donor−acceptor alternation in monomer energies. Diagrams of component dimers are available in Table S4, Supporting Information.

Furthermore, considering the sequences illustrated in the molecular barcodes of the top 25 hexamers and top 25 octamers, almost none exhibit the frequently used $(ABAB)_n$ sequence of simply alternating copolymers. Because the GA was followed by a local search to determine these species, the unusual sequences including AA and BB patterns in these top oligomers must yield improved efficiency relative to their simply alternating forms. Indeed, across the top 25 hexamers and octamers in Tables 2 and 3 for the alternating sequences, there is an average *decrease* in predicted efficiency of 0.4%, and for a few species the decrease is as high as 1.5−2%. In many cases, the alternating sequence was predicted to have an excited state below that of PCBM and thus a predicted efficiency of 0. For example, two more conventional donor−acceptor alternating copolymers are illustrated in Scheme 4 and are predicted to have 0 efficiencies because the polymer LUMO energy will fall below PCBM.

Examining the electronic structures for the alternating (ABAB) sequence analogues of our top hexamers and octamers, there is, on average, a decrease of ∼0.6−0.8 eV in the ionization potential, resulting in a decrease of ∼0.7 eV in the excitation energy compared with the tailored sequences. Because the optimal band gap is 1.4 eV, decreasing the excitation energy further cannot improve efficiency. In short, the GA and local search yield complex sequences as a mechanism to optimize the electronic structure given a fixed set of monomers.

One possible concern is that given such complex sequences as ABBABA, the hexamers and octamers may not be suitable models for the electronic structure of a longer polymer (of course, the oligomers themselves may be potential targets for technological applications). However, if we take the top 25 hexamers in Table 2, we find that the corresponding 12-mers and 18-mers show changes in HOMO energies by only 0.1 eV on average and changes in excitation energies by ∼0.2 eV on average. It is likely that such small changes occur because the top hexamers already demonstrate a high degree of delocalization.

Another contribution to the overall efficiency of a photovoltaic is the optical absorption intensity of the material. All of the top hexamers and octamers have low-energy transitions with oscillator strengths > 1.0 because the objective used by the GA penalizes copolymers that do not have such an electronic transition (see Supporting Information). Comparing the top 25 hexamers and octamers to their alternating sequences, the

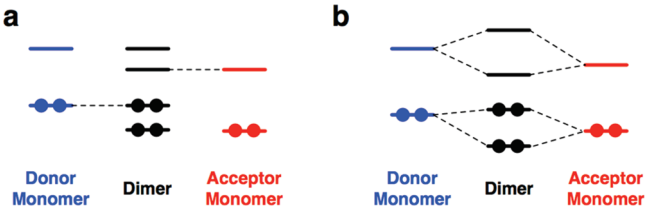
Table 3. Predicted Efficiencies, HOMO and LUMO Eigenvalues (in eV), and Colored Molecular Barcodes for the Top 25 Octamers and Two Reference Compounds in Scheme 4, Using ZINDO-Computed Monomer Orbital Eigenvalues (in eV).[a]

Global Color Scales	Minimum	Mid-Point	Maximum
HOMO Scale (in eV)	-10.72	-8.54	-6.94
LUMO Scale (in eV)	-6.60	-2.03	-0.39

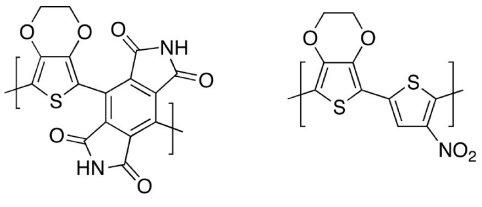
Entry	Eff.	HOMO	LUMO	HOMO Barcode								LUMO Barcode							
1	11.08	-5.69	-4.31	-7.98	-7.89	-7.89	-7.98	-7.89	-7.98	-7.89	-7.98	-2.68	-1.81	-1.81	-2.68	-1.81	-2.68	-1.81	-2.68
2	10.80	-5.73	-4.29	-7.48	-7.98	-7.98	-7.48	-7.98	-7.48	-7.48	-7.98	-2.43	-2.68	-2.68	-2.43	-2.68	-2.43	-2.43	-2.68
3	10.64	-5.94	-4.30	-7.29	-8.32	-8.32	-7.29	-8.32	-7.29	-8.32	-7.29	-2.18	-2.04	-2.04	-2.18	-2.04	-2.18	-2.04	-2.18
4	10.64	-5.85	-4.29	-8.27	-7.98	-8.27	-7.98	-8.27	-7.98	-7.98	-8.27	-2.24	-2.68	-2.24	-2.68	-2.24	-2.68	-2.68	-2.24
5	10.59	-5.97	-4.31	-8.01	-7.48	-8.01	-7.48	-7.48	-8.01	-7.48	-8.01	-2.85	-2.43	-2.85	-2.43	-2.43	-2.85	-2.43	-2.85
6	10.53	-5.81	-4.27	-7.98	-8.32	-7.98	-8.32	-8.32	-7.98	-8.32	-7.98	-2.68	-2.24	-2.68	-2.24	-2.24	-2.68	-2.24	-2.68
7	10.50	-5.99	-4.31	-7.60	-7.29	-7.29	-7.60	-7.60	-7.29	-7.29	-7.60	-1.47	-2.18	-2.18	-1.47	-1.47	-2.18	-2.18	-1.47
8	10.50	-5.92	-4.29	-7.98	-7.03	-7.98	-7.03	-7.03	-7.98	-7.03	-7.98	-2.68	-3.96	-2.68	-3.96	-3.96	-2.68	-3.96	-2.68
9	10.50	-5.72	-4.27	-7.35	-8.27	-8.27	-7.35	-8.27	-7.35	-8.27	-7.35	-2.05	-2.24	-2.24	-2.05	-2.24	-2.05	-2.24	-2.05
10	10.49	-5.97	-4.30	-7.89	-7.54	-7.54	-7.89	-7.89	-7.54	-7.54	-7.89	-1.81	-1.37	-1.37	-1.81	-1.81	-1.37	-1.37	-1.81
11	10.44	-5.94	-4.29	-7.98	-6.93	-7.98	-6.93	-7.98	-6.93	-6.93	-7.98	-2.68	-3.74	-2.68	-3.74	-2.68	-3.74	-3.74	-2.68
12	10.44	-5.87	-4.28	-8.01	-7.89	-8.01	-7.89	-7.89	-8.01	-7.89	-8.01	-2.85	-1.81	-2.85	-1.81	-1.81	-2.85	-1.81	-2.85
13	10.43	-5.82	-4.26	-8.32	-7.48	-8.32	-7.48	-7.48	-8.32	-7.48	-8.32	-2.04	-2.43	-2.04	-2.43	-2.43	-2.04	-2.43	-2.04
14	10.42	-5.91	-4.28	-8.32	-7.29	-8.32	-7.29	-8.32	-7.29	-8.32	-7.29	-2.04	-2.18	-2.04	-2.18	-2.04	-2.18	-2.04	-2.18
15	10.42	-6.00	-4.30	-7.89	-7.54	-7.89	-7.54	-7.89	-7.54	-7.54	-7.89	-1.81	-1.37	-1.81	-1.37	-1.81	-1.37	-1.37	-1.81
16	10.40	-5.71	-4.26	-8.27	-7.35	-8.27	-7.35	-8.27	-7.35	-8.27	-7.35	-2.24	-2.05	-2.24	-2.05	-2.24	-2.05	-2.24	-2.05
17	10.39	-6.03	-4.31	-7.03	-7.54	-7.54	-7.03	-7.54	-7.03	-7.03	-7.54	-3.96	-1.37	-1.37	-3.96	-1.37	-3.96	-3.96	-1.37
18	10.38	-6.02	-4.31	-7.29	-7.60	-7.60	-7.29	-7.60	-7.29	-7.29	-7.60	-2.18	-1.47	-1.47	-2.18	-1.47	-2.18	-2.18	-1.47
19	10.38	-5.62	-4.27	-7.98	-7.89	-7.98	-7.89	-7.89	-7.98	-7.89	-7.98	-2.68	-1.81	-2.68	-1.81	-1.81	-2.68	-1.81	-2.68
20	10.36	-5.79	-4.25	-7.98	-8.27	-7.98	-8.27	-8.27	-7.98	-8.27	-7.98	-2.68	-2.24	-2.68	-2.24	-2.24	-2.68	-2.24	-2.68
21	10.35	-5.90	-4.28	-8.32	-7.29	-8.32	-7.29	-8.32	-7.29	-7.29	-8.32	-2.04	-2.18	-2.04	-2.18	-2.04	-2.18	-2.18	-2.04
22	10.34	-5.62	-4.27	-7.89	-7.98	-7.89	-7.98	-7.89	-7.98	-7.98	-7.89	-1.81	-2.68	-1.81	-2.68	-1.81	-2.68	-2.68	-1.81
23	10.34	-5.99	-4.30	-7.60	-7.29	-7.60	-7.29	-7.29	-7.60	-7.29	-7.60	-1.47	-2.18	-1.47	-2.18	-2.18	-1.47	-2.18	-1.47
24	10.31	-6.02	-4.30	-8.32	-7.29	-7.29	-8.32	-7.29	-8.32	-8.32	-7.29	-2.04	-2.18	-2.18	-2.04	-2.18	-2.04	-2.04	-2.18
25	10.30	-6.02	-4.30	-7.54	-7.03	-7.03	-7.54	-7.03	-7.54	-7.03	-7.54	-1.37	-3.96	-3.96	-1.37	-3.96	-1.37	-3.96	-1.37
26	0.00	-7.38	-4.75	-10.32	-8.32	-10.32	-8.32	-10.32	-8.32	-10.32	-8.32	-4.97	-2.24	-4.97	-2.24	-4.97	-2.24	-4.97	-2.24
27	0.00	-6.89	-5.15	-9.38	-8.32	-9.38	-8.32	-9.38	-8.32	-9.38	-8.32	-4.00	-2.24	-4.00	-2.24	-4.00	-2.24	-4.00	-2.24

[a] The midpoint value is established from thiophene. Note that top candidates show very little donor−acceptor alternation in monomer energies. Diagrams of component dimers are available in Table S5, Supporting Information.

Scheme 3



Scheme 4



26 27

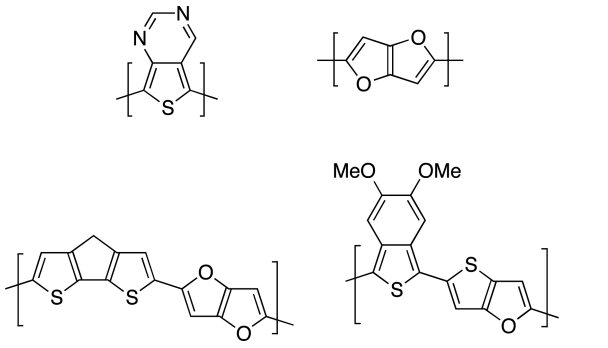
computed oscillator strength decreases on average by ~0.6, suggesting that using AA and BB patterns can greatly improve the optical density of the material and boost overall efficiency.

Analysis of Top Monomers and Dimers. As discussed above, monomers found in the top hexamers and octamers are computed to have low ionization potentials. More surprisingly, across the 101 promising tetramers, the top 25 hexamers, and the top 25 octamers (151 species in total), only 36 different monomers out of 132 appear. The two most frequent are shown in Scheme 5 and appear in over 30 of these top candidates. Neither has previously been synthesized as either a monomer or polymer.

Similarly, only 64 dimers in different oligomer lengths and sequences are found across these top 151 candidates, out of a possible 8778 combinations of monomers in the entire data set. The top two by frequency are illustrated in Scheme 5 and appear nine times each in different oligomer lengths and sequences. While the results from the GA are stochastic, and

dx.doi.org/10.1021/jp202765c |J. Phys. Chem. C 2011, 115, 16200–16210

Scheme 5



the exact number of "top dimers" may vary from run to run, the fact that such a small fraction of possible dimers (0.7%) represents all top candidates suggests that finding optimum dimer pairs is critical.

An important question is what makes these particular monomers and dimer combinations optimal? In other words, what are the design rules to create other, more synthetically accessible species for high-efficiency organic photovoltaics? To answer this, one can consider the dimers under a two-site Hückel model, comparing the occupied orbitals of the dimer to the HOMO orbitals of the donor and acceptor monomers, as illustrated in Scheme 3b. This analysis explicitly considers the interaction (H_{ab}) between each monomer, unlike the no-interaction model illustrated in Scheme 3a. Using this two-site Hückel model, the difference in energy between the HOMO and HOMO−1 orbitals in the dimer (ΔE) is derived from the HOMO energies of donor (E_d) and acceptor (E_a) monomers as follows

$$\Delta E = \sqrt{(E_d - E_a)^2 + (2H_{ab})^2} \quad (1)$$

Using eq 1, the average H_{ab} value across the top 64 dimers is 0.81, slightly lower than that of thiophene/bithiophene (0.9). These top dimers, however, have *significantly* higher electronic coupling than the dimer set as a whole, which yields an average H_{ab} of 0.51. From this, we can conclude that the top dimers emerging from the GA search do not simply have the "perfect" orbital energies but also exhibit a very high degree of delocalization and electronic coupling between monomers, almost as high as that of polythiophene. In short, compared to Scheme 3a, the picture in Scheme 3b is a more accurate representation of donor−acceptor copolymers, where the resulting orbitals derive as much from the electronic coupling between monomers (H_{ab}) as the monomer orbital energies.

Additional Filtering Criteria. Beyond the application of electronic and optical properties discussed above, additional filtering criteria can be applied to candidates. In this way, a computational prediction pipeline may continue to refine the list of top candidates for synthesis and experimental investigation.

Another important criterion in overall photovoltaic device performance is charge mobility. Because crystal structure prediction is an extremely difficult problem, it is consequently hard to accurately predict charge mobility in these novel species. However, charge transport in organic materials primarily occurs via hopping transport, based fundamentally on a bimolecular charge-transfer step. Thus, the reaction rate constant can be computed via Marcus−Hush theory[37,61−65]

$$k_{et} = \frac{2\pi\langle H_{ab}\rangle^2}{\hbar}\frac{1}{\sqrt{4\pi\lambda k_B T}} e^{-(\lambda + \Delta G^0)^2/4\lambda k_B T} \quad (2)$$

where ΔG^0, H_{ab}, and λ are the difference in energy between molecular sites, the electronic coupling matrix element between the orbitals involved, and the total reorganization energy of the reaction. The last (λ) represents the geometric change that occurs when accepting a hole or electron carrier. Although both the internal geometry of the molecule and the geometry of the surrounding medium will change, the internal reorganization energy is the more important factor. A recent study of the external reorganization energy in organic semiconductors computes that its value is extremely small,[66] and several works have established that the computed internal reorganization energy is correlated with the experimental hole mobility in organic materials.[67]

Table 4 gives the internal reorganization energies for the top 25 hexamers and octamers, calculated using density functional theory and the B3LYP functional,[68,69] as described previously.[37] Also in Table 4 are the computed relative rate constants for the

Table 4. Computed Internal Reorganization Energies (in eV) and Relative Intrinsic Charge-Transfer Rates for Hole Transfer, Assuming that All Other Contributions Remain Consistent[a]

entry	reorganization energy (eV)	relative rate	entry	reorganization energy (eV)	relative rate
	hexamers			octamers	
1	0.382	0.125	1	0.155	0.645
2	0.294	0.292	2	0.202	0.411
3	0.330	0.206	3	0.156	0.639
4	0.178	0.902	4	0.121	0.900
5	0.238	0.503	5	0.213	0.369
6	0.179	0.884	6	0.132	0.805
7	0.168	0.987	7	0.213	0.369
8	0.205	0.692	8	0.110	1.000
9	0.268	0.376	9	0.173	0.546
10	0.167	1.000	10	0.341	0.107
11	0.335	0.196	11	0.334	0.115
12	0.308	0.254	12	0.372	0.080
13	0.236	0.513	13	0.349	0.099
14	0.439	0.072	14	0.168	0.570
15	0.204	0.701	15	0.348	0.100
16	0.205	0.691	16	0.167	0.576
17	0.356	0.161	17	0.230	0.312
18	0.415	0.091	18	0.225	0.327
19	0.212	0.644	19	0.346	0.102
20	0.212	0.644	20	0.286	0.181
21	0.233	0.529	21	0.355	0.093
22	0.324	0.219	22	0.322	0.128
23	0.344	0.180	23	0.344	0.104
24	0.366	0.146	24	0.322	0.128
25	0.342	0.184	25	0.397	0.062

[a] The results suggest that some candidates are likely to have poor hole mobilities.

hexamers or octamers using computed reorganization energies and eq 2, assuming that all other contributions remain constant. The results are not intended to be accurate predictions but rather a method to screen out species with high reorganization energies that are likely to exhibit poor hole mobility. For example, hexamer entry **1**, while it is predicted to have the highest efficiency (11%) on the basis of the electronic structure and optical excitation energy, also has a high computed internal reorganization energy. Octamer entries **10**, **11**, **12**, **13**, **15**, **19**, **21**, **23**, and **25** all have high computed internal reorganization energies.

Because the degree of conjugation is affected by the dihedral angle between monomers in the oligomers, we also compared the ZINDO orbital eigenvalues and excitations energies for the original PM6-predicted structures to those obtained for B3LYP-optimized structures. We find that the effect is small, altering the ZINDO HOMO energies of the top 25 hexamers by 0.15 eV on average and the excitation energies by 0.19 eV on average. Similar changes occur in the octamers. We consider this as another step in the filtering process, using the PM6/ZINDO combination as a computationally efficient screen, followed by computation of more accurate geometries and electronic structure properties on the top species found by the methods described above.

CONCLUSIONS

We have surveyed over 90 000 π-conjugated oligomers to predict their potential efficiency as components in organic bulk heterojunction photovoltaics. We believe this to be the largest available database of conjugated polymers for photovoltaic applications. Of these, several hundred tetramers, hexamers, and octamers are predicted to yield energy conversion efficiencies > 8%, the current validated efficiency record, with many hexamers and octamers at or above 10%. While the semiempirical quantum chemical methods are computationally efficient (~8−10 min per compound), given the size of the search space, a GA has been used to target the most efficient oligomers, with optimal electronic structure and optical properties. In the case of the hexamers and octamers, histograms demonstrate that the GA and subsequent local search clearly find a large fraction of high-efficiency targets.

Our analysis of component monomers, dimers, and the copolymer sequence demonstrates important design rules for copolymer photovoltaics. Most importantly, the conventional picture of combining a strong donor and strong acceptor into an alternating copolymer is found to frequently yield poor energy-level alignment. Instead, our top hexamers and octamers reflect a decreased optical band gap due to high coupling between the two-component monomers, not solely due to particular HOMO or LUMO energies of the monomers themselves. Tailored sequences of these monomers yield a near-optimal electronic structure of the oligomers and increased optical oscillator strengths over purely alternating sequences.

These procedures can be used as the start of a computational pipeline for targeted organic materials, for example, by subsequent filtering of compounds with high internal reorganization energies that will likely yield poor hole mobilities. Such a pipeline strategy for rational synthesis is significantly more efficient than time-consuming serial experiments and draws on the success of computational drug design. It will be of paramount importance in multilayer solar cells, where the band gaps and energetics of each layer must be carefully matched for optimum performance, and for combined improvement of both polymer donor and acceptor materials.[70]

Future work in this area must focus on careful calibration of predicted and experimental efficiencies, particularly noting issues such as poor material conductivity, optical absorption, charge recombination rates, or other issues that may limit device efficiency that are not considered by this model. The use of broader monomer libraries and alternate polymerization sites will enable a much more thorough sampling of chemical space for worthwhile materials, and a computational investigation of the effects of monomer sequence on predicted electronic properties will suggest whether designed repeating-sequence copolymers can significantly enhance photovoltaic efficiency, as suggested by these preliminary results. Finally, beyond the simple electronic structure descriptors used here, before synthesis can begin on promising candidates, additional filtering is needed to consider potential solubility, crystal structure packing, and other properties. Each of these components is likely to remove a large fraction of candidate structures but will assist in finding the best possible candidates for eventual manufacturing of working high-efficiency devices. Using the GA and diversity library method described here will easily provide many initial candidates for such a materials discovery pipeline.

ASSOCIATED CONTENT

Supporting Information. Complete text of ref 48, table of monomer structures, full details of the genetic algorithm implementation, figures and data for calibration of PM6/ZINDO excitation energies and ionization potentials with experiment, and tables of top tetramers, hexamers, and octamers. Details can be found at http://hutchison.chem.pitt.edu/. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: geoffh@pitt.edu.

ACKNOWLEDGMENT

N.M.O'B. is supported by a Career Development Fellowship from the Health Research Board (Grant PD/2009/13). G.R.H. thanks the Center for Energy at the University of Pittsburgh for support. The authors wish to acknowledge the SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities and support. We thank Dr. Markus Scharber for help in computing efficiency values and Prof. Tara Meyer for helpful discussions.

REFERENCES

(1) Mayer, A. C.; Scully, S. R.; Hardin, B. E.; Rowell, M. W.; McGehee, M. D. *Mater. Today* **2007**, *10* (11), 28.
(2) Coakley, K.; McGehee, M. *Chem. Mater.* **2004**, *16* (23), 4533.
(3) Lloyd, M. T.; Anthony, J. E.; Malliaras, G. G. *Mater. Today* **2007**, *10* (11), 34.
(4) Gunes, S.; Neugebauer, H.; Sariciftci, N. S. *Chem. Rev.* **2007**, *107* (4), 1324.
(5) Yu, G.; Gao, J.; Hummelen, J. C.; Wudl, F.; Heeger, A. J. *Science* **1995**, *270* (5243), 1789.
(6) Heremans, P.; Cheyns, D.; Rand, B. P. *Acc. Chem. Res.* **2009**, *42* (11), 1740.
(7) Roncali, J. *Acc. Chem. Res.* **2009**, *42* (11), 1719.
(8) Dennler, G.; Scharber, M. C.; Brabec, C. J. *Adv. Mater.* **2009**, *21* (13), 1323.

(9) Thompson, B. C.; Frechet, J. M. J. *Angew. Chem., Int. Ed.* **2008**, *47* (1), 58.
(10) Liang, Y.; Xu, Z.; Xia, J.; Tsai, S.-T.; Wu, Y.; Li, G.; Ray, C.; Yu, L. *Adv. Mater.* **2010**, *22* (20), E135.
(11) Rand, B.; Burk, D.; Forrest, S. *Phys. Rev. B* **2007**, *75* (11), 115327.
(12) Scharber, M.; Wuhlbacher, D.; Koppe, M.; Denk, P.; Waldauf, C.; Heeger, A.; Brabec, C. *Adv. Mater.* **2006**, *18* (6), 789.
(13) Koster, L.; Mihailetchi, V.; Blom, P. *Appl. Phys. Lett.* **2006**, *88*, 093511.
(14) Potscavage, W. J.; Sharma, A.; Kippelen, B. *Acc. Chem. Res.* **2009**, *42* (11), 1758.
(15) Kippelen, B.; Brédas, J. L. *Energy Environ. Sci.* **2009**, *2* (3), 251.
(16) Brédas, J. L.; Norton, J. E.; Cornil, J.; Coropceanu, V. *Acc. Chem. Res.* **2009**, *42* (11), 1691.
(17) Peet, J.; Heeger, A. J.; Bazan, G. C. *Acc. Chem. Res.* **2009**, *42* (11), 1700.
(18) Peet, J.; Senatore, M. L.; Heeger, A. J.; Bazan, G. C. *Adv. Mater.* **2009**, *21* (14−15), 1521.
(19) Kroon, R.; Lenes, M.; Hummelen, J. C.; Blom, P. W. M.; De Boer, B. *Polym. Rev.* **2008**, *48* (3), 531.
(20) Durrant, M. C. *Chem.—Eur. J.* **2007**, *13* (12), 3406.
(21) Xiao, D.; Yang, W.; Beratan, D. N. *J. Chem. Phys.* **2008**, *129* (4), 044106.
(22) Hu, X.; Beratan, D. N.; Yang, W. *J. Chem. Phys.* **2008**, *129* (6), 064102.
(23) Keinan, S.; Hu, X.; Beratan, D. N.; Yang, W. *J. Phys. Chem. A* **2007**, *111* (1), 176.
(24) Wang, M.; Hu, X.; Beratan, D.; Yang, W. *J. Am. Chem. Soc.* **2006**, *128* (10), 3228.
(25) von Lilienfeld, O. A.; Lins, R. D.; Rothlisberger, U. *Phys. Rev. Lett.* **2005**, *95* (15), 153002.
(26) von Lilienfeld, O. A.; Tuckerman, M. E. *J. Chem. Phys.* **2006**, *125* (15), 154104.
(27) von Lilienfeld, O. A.; Tuckerman, M. E. *J. Chem. Theory Comput.* **2007**, *3* (3), 1083.
(28) Rinderspacher, B. C.; Andzelm, J.; Rawlett, A.; Dougherty, J.; Beratan, D. N.; Yang, W. T. *J. Chem. Theory Comput.* **2009**, *5* (12), 3321.
(29) Glen, R. C.; Payne, A. W. R. *J. Comput.-Aided Mol. Des.* **1995**, *9* (2), 181.
(30) Kamphausen, S.; Holtge, N.; Wirsching, F.; Morys-Wortmann, C.; Riester, D.; Goetz, R.; Thurk, M.; Schwienhorst, A. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 551.
(31) Venkatasubramanian, V.; Chan, K.; Caruthers, J. *J. Chem. Inf. Comput. Sci.* **1995**, *35* (2), 188.
(32) Bohacek, R.; McMartin, C.; Guida, W. *Med. Res. Rev.* **1996**, *16* (1), 3.
(33) Fink, T.; Bruggesser, H.; Reymond, J. *Angew. Chem., Int. Ed.* **2005**, *44* (10), 1504.
(34) Fink, T.; Reymond, J.-L. *J. Chem. Inf. Model.* **2007**, *47* (2), 342.
(35) van Deursen, R.; Reymond, J. *ChemMedChem* **2007**, *2* (5), 636.
(36) Hutchison, G. R.; Ratner, M.; Marks, T. *J. Phys. Chem. B* **2005**, *109* (8), 3126.
(37) Hutchison, G. R.; Ratner, M.; Marks, T. *J. Am. Chem. Soc.* **2005**, *127* (7), 2339.
(38) Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31.
(39) *Open Babel* v. 2.2.3. http://openbabel.org/ (2011).
(40) O'Boyle, N. M.; Morley, C.; Hutchison, G. R. *Chem. Cent. J.* **2008**, *2*, 5.
(41) Halgren, T. *J. Comput. Chem.* **1996**, *17* (5−6), 490.
(42) Halgren, T. *J. Comput. Chem.* **1996**, *17* (5−6), 520.
(43) Halgren, T. *J. Comput. Chem.* **1996**, *17* (5−6), 553.
(44) Halgren, T. *J. Comput. Chem.* **1996**, *17* (5−6), 616.
(45) Halgren, T.; Nachbar, R. *J. Comput. Chem.* **1996**, *17* (5−6), 587.
(46) Stewart, J. J. P. *J. Mol. Model.* **2007**, *13* (12), 1173.
(47) Ridley, J.; Zerner, M. *Theor. Chim. Acta* **1973**, *32*, 111.
(48) Frisch, M. J., et al. *Gaussian 09*, revision A.2; Gaussian, Inc.: Wallingford CT, 2009.
(49) O'Boyle, N. M.; Tenderholt, A. L.; Langner, K. M. *J. Comput. Chem.* **2008**, *29* (5), 839.
(50) Hutchison, G. R.; Ratner, M. A.; Marks, T. J. *J. Phys. Chem. A* **2002**, *106* (44), 10596.
(51) Lias, S. G. Ionization Energy Evaluation. In *NIST Chemistry WebBook, NIST Standard Reference Database*, Mallard, W. G., Linstrom, P. J., Liebman, J. F., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 2010; Vol. 69.
(52) Because the ZINDO eigenvalues are intrinsically gas-phase quantities and there may be electrostatic and dielectric effects in a solid-state film similar to solvation, we also compared HOMO eigenvalues using a PCM solvation model for both benzene and pyridine solvents and found systematic shifts, with an average deviation between the gas phase and either solvent to be ∼0.1 eV, well within the error of the method itself.
(53) Shuttle, C.; Hamilton, R.; O'Regan, B.; Nelson, J.; Durrant, J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (38), 16448.
(54) Clarke, T. M.; Ballantyne, A. M.; Tierney, S.; Heeney, M.; Duffy, W.; McCulloch, I.; Nelson, J.; Durrant, J. R. *J. Phys. Chem. C* **2010**, *114* (17), 8068.
(55) Shoaee, S.; An, Z.; Zhang, X.; Barlow, S.; Marder, S. R.; Duffy, W.; Heeney, M.; Mcculloch, I.; Durrant, J. R. *Chem. Commun.* **2009**, No. 36, 5445.
(56) Brabec, C. J.; Shaheen, S. E.; Winder, C.; Sariciftci, N. S.; Denk, P. *Appl. Phys. Lett.* **2002**, *80* (7), 1288.
(57) Kim, J. Y.; Kim, S. H.; Lee, H. H.; Lee, K.; Ma, W. L.; Gong, X.; Heeger, A. J. *Adv. Mater.* **2006**, *18* (5), 572.
(58) Quattrocchi, C.; Lazzaroni, R.; Brédas, J.-L. *Chem. Phys. Lett.* **1993**, *208* (1−2), 120.
(59) Karpfen, A.; Choi, C.; Kertesz, M. *J. Phys. Chem. A* **1997**, *101* (40), 7426.
(60) Viruela, P. M.; Viruela, R.; Ortí, E.; Brédas, J.-L. *J. Am. Chem. Soc.* **1997**, *119* (6), 1360.
(61) Brédas, J.-L.; Calbert, J.; da Silva, D.; Cornil, J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (9), 5804.
(62) da Silva, D.; Kim, E.; Brédas, J. *Adv. Mater.* **2005**, *17* (8), 1072.
(63) Kim, E.-G.; Coropceanu, V.; Gruhn, N. E.; Sanchez-Carrera, R. S.; Snoeberger, R.; Matzger, A. J.; Brédas, J.-L. *J. Am. Chem. Soc.* **2007**, *129* (43), 13072.
(64) Kjelstrup-Hansen, J.; Norton, J. E.; da Silva Filho, D. A.; Brédas, J.-L.; Rubahn, H.-G. *Org. Electron.* **2009**, *10* (7), 1228.
(65) Li, H.; Brédas, J.-L.; Lennartz, C. *J. Chem. Phys.* **2007**, *126* (16), 164704.
(66) McMahon, D. P.; Troisi, A. *J. Phys. Chem. Lett.* **2010**, *1* (6), 941.
(67) Malagoli, M.; Brédas, J.-L. *Chem. Phys. Lett.* **2000**, *327* (1−2), 13.
(68) Becke, A. *J. Chem. Phys.* **1993**, *98* (7), 5648.
(69) Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, *37* (2), 785.
(70) Dennler, G.; Scharber, M. C.; Ameri, T.; Denk, P.; Forberich, K.; Waldauf, C.; Brabec, C. J. *Adv. Mater.* **2008**, *20* (3), 579.