



# The Harvard Clean Energy Project: Large-Scale Computational Screening and Design of Organic Photovoltaics on the World Community Grid

## Citation

Hachmann, Johannes, Roberto Olivares-Amaya, Sule Atahan-Evrenk, Carlos Amador-Bedolla, Roel S. Sánchez-Carrera, Aryeh Gold-Parker, Leslie Vogt, Anna M. Brockway, and Alán Aspuru-Guzik. 2011. The Harvard Clean Energy Project: Large-Scale computational screening and design of organic photovoltaics on the world community grid. *Journal of Physical Chemistry Letters* 2(17): 2241–2251.

## Published Version

doi:10.1021/jz200866s

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:8364968>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

## Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# The Harvard Clean Energy Project: Large-Scale Computational Screening and Design of Organic Photovoltaics on the World Community Grid

Johannes Hachmann,<sup>1,\*</sup> Roberto Olivares-Amaya,<sup>1</sup> Sule Atahan-Evrenk,<sup>1</sup> Carlos Amador-Bedolla,<sup>2,1</sup> Roel S. Sánchez-Carrera,<sup>†,1</sup> Aryeh Gold-Parker,<sup>1</sup> Leslie Vogt,<sup>1</sup> Anna M. Brockway,<sup>3</sup> and Alán Aspuru-Guzik<sup>1,‡</sup>

<sup>1</sup>*Department of Chemistry and Chemical Biology, Harvard University,  
12 Oxford St, Cambridge, MA 02138, USA*

<sup>2</sup>*Facultad de Química, Universidad Nacional Autónoma de México, México, DF 04510, Mexico*

<sup>3</sup>*Department of Chemistry, Haverford College, 370 Lancaster Ave, Haverford, PA 19041, USA*

(Dated: August 10, 2011)

This perspective article introduces the *Harvard Clean Energy Project* (CEP), a theory-driven search for the next generation of organic solar cell materials. We give a broad overview of its setup and infrastructure, present first results, and outline upcoming developments.

CEP has established an automated, high-throughput, *in silico* framework to study potential candidate structures for organic photovoltaics. The current project phase is concerned with the characterization of millions of molecular motifs using *first-principles* quantum chemistry. The scale of this study requires a correspondingly large computational resource which is provided by distributed volunteer computing on IBM’s World Community Grid. The results are compiled and analyzed in a reference database and will be made available for public use. In addition to finding specific candidates with certain properties, it is the goal of CEP to illuminate and understand the structure-property relations in the domain of organic electronics. Such insights can open the door to a rational and systematic design of future high-performance materials. The computational work in CEP is tightly embedded in a collaboration with experimentalists, who provide valuable input and feedback to the project.

## I. INTRODUCTION

The sun is an abundant source of energy and its input on earth exceeds the global consumption by 4 orders of magnitude. It is hence an obvious alternative to fossil or nuclear energy supplies and will play an important role in safely and sustainably covering the rising demands of the future [1–4]. The current cost of electricity from commercial silicon-based solar cells is unfortunately still around 10 times higher than that of utility-scale electrical power generation [5, 6]. These traditional inorganic photovoltaics come with further shortcomings, such as a complicated and energy-intensive manufacturing process which leads to high production costs. They can also contain rare or hazardous elements, and the devices tend to be heavy, bulky, rigid, and fragile.

Carbon-based solar cells have emerged as one of the interesting alternatives to this conventional technology [7–9]. Organic photovoltaics (OPVs) range from crystalline small molecule approaches [10–12] and certain dye-sensitized Grätzel cells [13] to amorphous polymers (plastics) [14, 15]. OPVs have great potential in two important areas: they promise simple, low-cost, and high-volume production [16] as well as the prospect of merging the unique flexibility and versatility of plastics with electronic features. OPVs can be processed via roll-to-roll

printing [17, 18]; there is active research in sprayable and paintable materials [19–22]; OPVs can be semi-transparent [23], variously colored [24], they are lightweight [14], and can essentially be molded into any shape [25]. These properties make OPVs a promising candidate to achieve the ubiquitous harvesting of solar energy [26, 27], with building-integrated [28–30] and ultra-portable applications [31–33] as primary targets. The Equinox Summit international committee, for example, recently suggested the use of OPVs for the basic electrification of 2.5 billion people in rural areas without access to the power grid [34].

There are, however, still significant issues to overcome in order to make plastic solar cells a viable technology for the future. The cardinal problems are their relatively low efficiency and limited lifetime [35, 36]: the power conversion record has only reached 9.2% [37] and current materials still degrade when exposed to the environment [38–41]. An increase in efficiency to about 10–15% in combination with lifetimes of over 10 years (for production materials) could push the power generation costs of OPVs below that of other currently available energy sources [42, 43].

In 2008, we started the Harvard Clean Energy Project (CEP) [44] to help find such high-efficiency OPV materials. This perspective article gives a general overview of CEP and provides the context for a series of detailed technical and result-oriented papers to follow. In Sec. II we introduce the motivation and overall setup of the project, followed by a presentation of its different components: Sec. II A discusses our OPV candidates library, Sec. II B the use of cheminformatics descriptors to rapidly assess the potential of these candidates, and Sec. II C is

---

<sup>†</sup>Present address: Research and Technology Center North America, Robert Bosch LLC, Cambridge, MA 02142

\*Electronic address: [jh@chemistry.harvard.edu](mailto:jh@chemistry.harvard.edu)

<sup>‡</sup>Electronic address: [aspuru@chemistry.harvard.edu](mailto:aspuru@chemistry.harvard.edu)

focused on the high-level calculation hierarchy in CEP. Sec. IID is concerned with the calibration of the obtained results. The CEP database is introduced in Sec. IIE and the World Community Grid (WCG) – our primary computational resource – in Sec. IIF. Throughout Sec. II we also indicate the next stages and extensions to the current setup. We summarize our discussion in Sec. III.

## II. THE HARVARD CLEAN ENERGY PROJECT

The key parameters for the improvement of OPVs are essentially known, however, engineering materials which combine all these features is a hard problem [45–49]. Traditional experimental development is largely based on empirical intuition or experience within a certain family of systems, and only a few examples can be studied per year due to long turnaround times of synthesis and characterization [50]. Theoretical work is usually also restricted to a small set of candidates for which different aspects of the photovoltaic process are modeled [51–53].

The Clean Energy Project stands out from other computational materials science approaches as it combines conventional modeling with strategies from modern drug discovery [54–61]: CEP features an automated, high-throughput infrastructure for a systematic screening of millions of OPV candidates at a *first-principles* electronic structure level [62]. It also adopts techniques from cheminformatics [63, 64] and heavily relies on data mining [65, 66]. Pioneering work on cheminformatics-type approaches and massive electronic structure calculations was, e.g., performed by Rajan *et al.* [67–69] and Ceder *et al.* [70], respectively, in the context of inorganic solids. An *in silico* study combining the scale and level of theory found in CEP is, however, unprecedented.

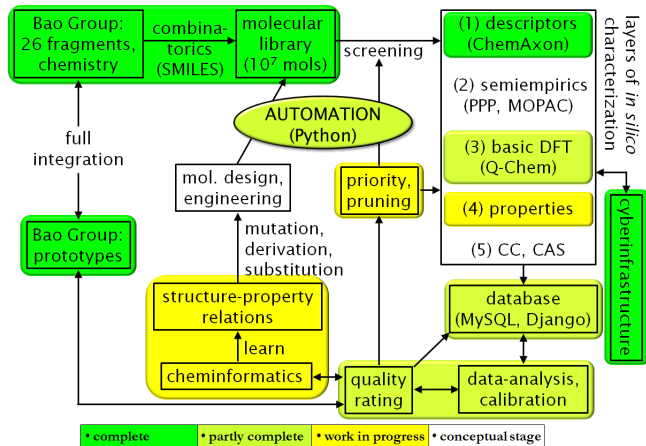


FIG. 1: Structure and workflow of CEP.

As the starting point for CEP we have chosen to investigate the molecular motifs at the heart of OPV materials [45, 71]. A suitable motif is a *necessary* condition for a successful OPV development. Only a limited

number of structural patterns have been explored so far, while the endless possibilities may well hold the key to overcoming the current material issues. We emphasize that a promising molecular structure is not a *sufficient* condition though, as condensed matter and device considerations have to be addressed as well. CEP is set up with a calculation hierarchy in which we will successively characterize relevant electronic structure aspects of our OPV candidates. Eventually, we will go beyond single molecule, gas-phase studies and consider intermolecular and bulk phase problems.

In addition to the search for specific structures with a desired set of properties, we also try to arrive at a systematic understanding of structure-property relationships [72–75]. Learning about underlying design principles is the key to moving from a screening effort towards an active engineering of novel organic electronics [76, 77].

While its centerpiece is the *in silico* study of OPV candidates, we point out that an overarching theme of CEP is also the tight integration of experimental and theoretical work. The project is in part guided by inputs from experimentalist collaborators (in particular the Bao Group at Stanford University), and our most promising candidates are subject to in-depth studies in their laboratories [78]. CEP is designed as a community tool and we invite and welcome joint ventures.

Fig. 1 summarizes the overall structure and workflow of CEP, and in the following sections we discuss its various components.

### A. Molecular candidate libraries

The molecular structure of essentially all organic electronic materials features a conjugated  $\pi$ -backbone [79]. Modern semiconducting compounds are often composed of linked or fused (hetero-)aromatic scaffolds [80].

We have developed a combinatorial molecule generator to build the primary CEP library. It contains  $\sim 10,000,000$  molecular motifs of potential interest ( $\sim 3,600,000$  distinct connectivities, each with a set of conformers) which cover small molecule OPVs and oligomer sequences for polymeric materials. They are based on 26 building blocks and bonding rules (see Fig. 2) which were chosen following advice from our experimental collaborators considering promise and feasibility. The fragments were linked and fused up to a length of 4-5 units according to the given rules. The generator is graph-based and employs a SMILES (simplified molecular input line entry specification) [81] string representation of the molecules as well as SMARTS (SMILES arbitrary target specification) in its engine. It is built around Marvin Reactor [82], and the Corina code [83] provides force-field optimized 3-dimensional structures. A detailed description of the library generator is in preparation.

The construction of the primary library was tailored towards OPV donor candidates, but the obtained struc-

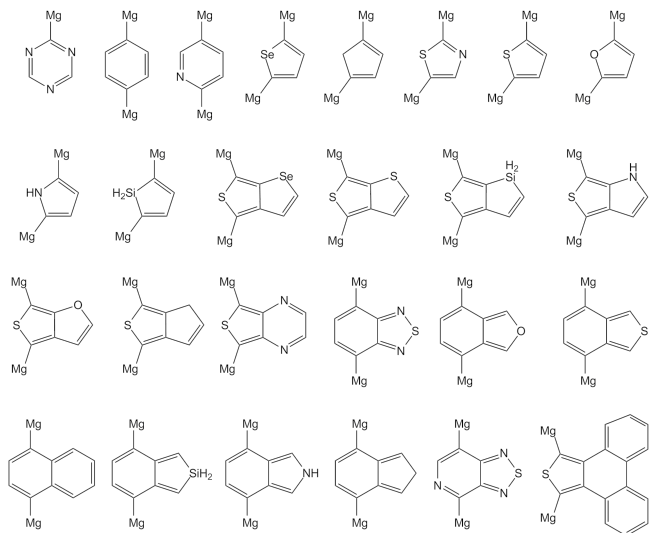


FIG. 2: The 26 building blocks used for generating the CEP molecular library. The Mg atoms represent chemical handles, i.e., the reactive sites in the generation process. We introduce simple links between two moieties (by means of substituting two Mg for a single C-C-bond) as well as the fusion of two rings.

tures are of interest for organic electronics in general [84, 85]. Our generator approach is flexible and can readily produce additional libraries (e.g., geared towards acceptor materials or Grätzel cell dyes) by using a different set of fragments and connection rules. Substituents can be incorporated in a similar fashion. Combinatorial libraries allow for an exhaustive and systematic exploration of well defined chemical spaces, but the number of generated molecules grows exponentially. At a later stage we plan to use a genetic algorithm [86–90] as a complement to the current approach [62]. The fitness function will be based on the information gathered from the screening of the primary library. Finally, CEP also provides the facility (e.g., to collaborators) to manually add structures to the screening pool.

## B. Cheminformatics descriptors

While the main objective of CEP is the *first-principles* characterization of OPV candidates, we also explore the use of cheminformatics descriptors [91, 92] and ideas from machine learning [93, 94], pattern recognition [95, 96], and drug discovery [54, 56, 64] to rapidly gauge their quality. We have devised descriptor models for parameters such as the short-circuit current density ( $J_{sc}$ ), the open-circuit voltage ( $V_{oc}$ ), and the power conversion efficiency (PCE). Our models are the first step in a successive ladder of approximations for these key quantities associated with photovoltaic performance.

The basic strategy behind this approach is to identify

and exploit correlations between certain physicochemical or topological descriptors and the properties of interest. Suitable descriptors are combined into models which are then empirically parametrized using a training set of experimentally well-characterized reference systems. As descriptors are easily computed, we can quickly (i.e., within a few days on a single workstation) obtain an initial assessment and preliminary ranking of the entire molecular library.

In Fig. 3 we present the linear regression model for  $V_{oc}$  along with the histogram of the calculated values for the molecular library. We note that our model for  $V_{oc}$  shows a very good correlation. This can be rationalized considering the principle dependence of  $V_{oc}$  on intramolecular properties which are apparently well reflected in the employed molecular descriptors. The  $J_{sc}$  model behaves similarly well although  $J_{sc}$  is also linked to bulk effects. For the fill factor – a quantity primarily determined by morphology and device characteristics – we could in contrast only obtain poor models.

The early stages of this work [97] utilized descriptors from the Marvin code by ChemAxon [82] and for the modeling we employed the R statistics package [98]. Recently, we started using the more comprehensive descriptor set from Dragon [99] and the specialized modeling code StarDrop [100]. A focus of our current work is to utilize the quantum chemical results discussed in the following section as a descriptor basis in our models.

As in biomedical applications of cheminformatics, we do not expect quantitative results, but this technique can yield valuable trends which we use to prioritize and prune the high-level screening and to uncover molecular motifs of particular interest. In Ref. [97] we give an introduction to this approach with a detailed discussion of the systematic construction and optimization of descriptor models.

## C. First-principles screening hierarchy

Electronic structure theory offers a way to probe the properties of OPV materials and the photophysical processes in organic solar cells [101, 102]. The complexity of these problems, however, poses severe methodological challenges. Multi-scale simulations have improved considerably in recent years [51, 103], but in practice we still commonly choose to model, approximate, or deduce the different aspects of the problem separately. CEP adopts such a divide-and-conquer approach and combines it with a calculation hierarchy to screen for promising material candidates. This multi-level setup is designed to successively address the relevant issues in OPVs and provide results at an increasing level of theory. At each stage, the candidates are rated with respect to the investigated parameters. The scoring is freely customizable to reflect different research priorities. The most promising candidates and related structures from the library receive priority in the CEP hierarchy and their characterization

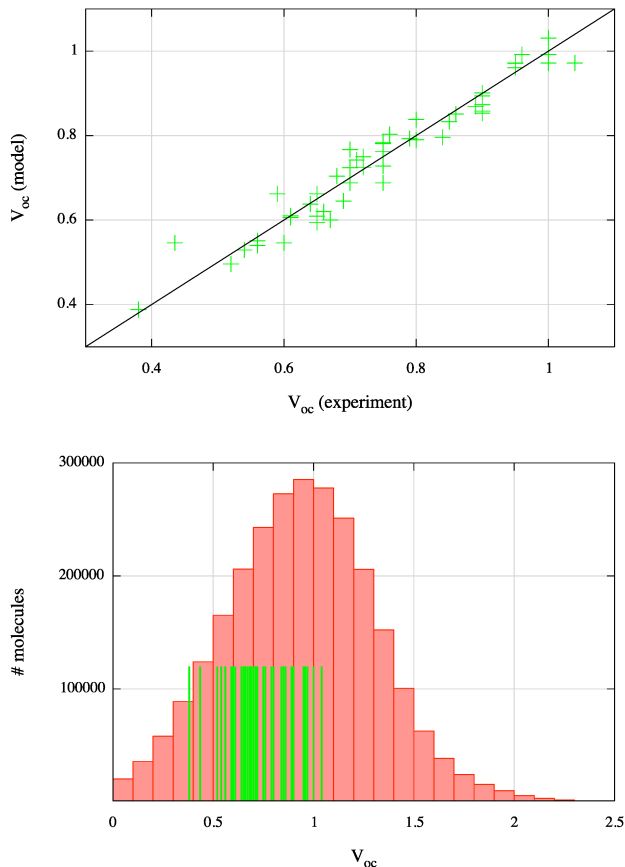


FIG. 3: Top: Linear regression model for  $V_{oc}$  with the training set data. Bottom: Histogram of the projected  $V_{oc}$  values in the molecular library with the position of the training set marked in green (with arbitrary height). See Ref. [97] for details.

is expedited. On the other hand, if a candidate is unfit with respect to a tested criterium and hence unlikely to be successful overall, it has lower priority and may be characterized in less detail.

The early CEP stages concentrate on various molecular properties of our material candidates and are hence most useful to assess macroscopic quantities which primarily depend on them, such as the  $V_{oc}$ . The later stages will shift the focus to intermolecular and condensed phase characteristics. The latter are central to, e.g., exciton and charge transport processes [46, 51, 52, 104–106], for which molecular properties alone are clearly of limited value. Bulk structure considerations impact quantities like the external quantum efficiency and thus  $J_{sc}$  as well as the overall PCE. Since the cost and complexity of such studies increase significantly, they can only be performed for a subset of highest rated candidates.

In the first CEP phase, we perform a set of density functional theory (DFT) calculations [108, 109] employing the BP86 [110, 111], B3LYP [110, 112, 113],

PBE0 [114–119], BH&HLYP [110, 112, 120], and M06-2X [121, 122] functionals as well as Hartree-Fock (HF) theory in combination with the single- $\zeta$  STO-6G [123, 124], double- $\zeta$  def2-SVP, and triple- $\zeta$  def2-TZVP [125] basis sets. We test and compare both restricted and spin-polarized settings. Our selection of functionals covers both generalized gradient approximation (GGA) and hybrid designs with a progression in the amount of exact exchange [113] (BP86 and HF can be seen as the limiting cases). The latter has a systematic influence on orbital localization and thus eigenvalues [126]. The GGA BP86 is a cost effective way to obtain good geometries, B3LYP is arguably the most widely used functional in molecular quantum mechanics, PBE0 has shown favorable performance in a variety of areas, as has the highly parameterized M06-2X. (We will further test meta-GGAs like TPSS [127] and double-hybrids like B2PLYP [128] when they become available for CEP.) This range of model chemistries was chosen to put our analysis on a broader footing, but also to assess the performance of the different theoretical methods [109]. In total, each molecule is characterized by a BP86/def2-SVP geometry optimization and 14 single-point ground state calculations. We obtain geometries, total energies (including their decomposition into different contributions), molecular orbitals (MOs) and their energy eigenvalues, electron and spin densities, electrostatic potentials, multipole moments, Mulliken [129, 130] and natural populations [131, 132], as well as natural atomic, localized molecular, and bond orbital analysis [133–135] for the different model chemistries [136].

These basic electronic quantities can be used as a first approximation to the following points. The MO energies and their differences can be related to ionization potentials, electron affinities, gaps, and partial density of states (we note that the application of Koopmans’ theorem is problematic in DFT) [126, 137–142]. The electronic levels have to be tuned for an efficient light absorption, for the necessary interplay between donor, acceptor, and lead materials, as well as for atmospheric stability [45]. The delocalization of the frontier orbitals can be associated with (intramolecular) exciton and charge carrier mobility [143–147]. Their spatial overlap indicates the transition character and probability of the corresponding excitation [148–150]. Excess spin densities reflect the inadequacy of simple closed-shell solutions and can also point to a complex and potentially unstable electronic situation [151]. Charge maps can identify chemically unstable sites in these highly unsaturated molecules as well as patterns which may have an impact on their packing in the condensed phase. The molecular multipole moments can correspondingly be correlated to the organization in the bulk structure and also to transport properties [152–155]. The wavefunction analysis techniques are of interest for the study of structure-property relations. The obtained data can furthermore be utilized as quantum chemical descriptors for the models described in Sec. II B. The results of a consecutive oligomer series can be used to

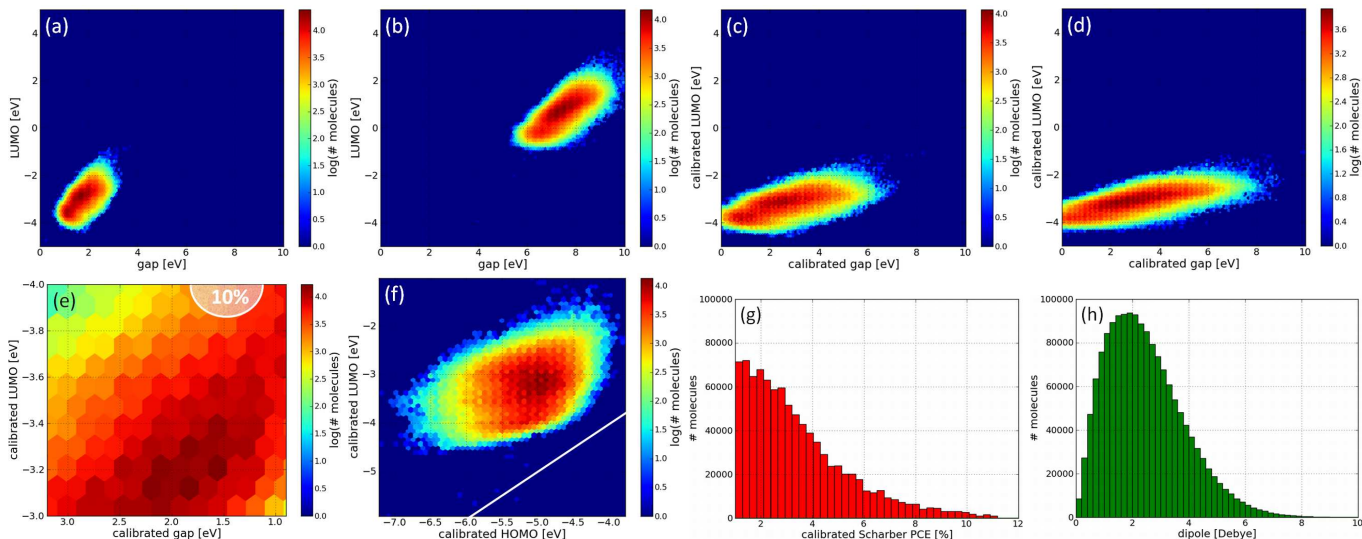


FIG. 4: (a-e) Distribution of CEP hits in the gap-LUMO plane suggested by Scharber *et al.* [107]. Note that the density plots are on a logarithmic scale, and the red entries correspond to  $\mathcal{O}(10,000)$  compounds and  $\mathcal{O}(150,000)$  individual electronic structure calculations; (a) and (b) show the raw data from BP86/def2-SVP and BP86/def2-SVP//HF/def2-SVP (i.e., calculations which incorporate 0% vs. 100% exact exchange), respectively; (c) and (d) display the corresponding data after preliminary calibration; (e) shows the OPV relevant parameter space with the 10% PCE region (with respect to a PCBm acceptor). About 0.3% of the screened compounds fall in this high-efficiency region; (f) indicates the dynamic gap range, i.e., the range of available LUMO energies given a particular HOMO and vice versa; (g) displays the PCE histogram according to the Scharber model, and (h) the dipole moment histogram; (e-h) are all at BP86/def2-SVP//PBE0/def2-TZVP level of theory (calibrated).

extrapolate to the polymer limit [156–159].

One example for how this basic information can be employed is the model developed by Scharber *et al.* in Ref. [107]. The authors assume (based on observations on actual OPV materials) that condensed matter aspects in such systems can be tuned to a certain viable level (i.e., a fill factor and external quantum efficiency of each 65% can be achieved). If that is the case, their PCE can be approximated using only gap and lowest unoccupied molecular orbital (LUMO) energy values. This model can readily be applied to the current CEP data as shown in Fig. 4. Our preliminary analysis reveals that only about 0.3% of the screened compounds (i.e., between 5,000-9,000 depending on the model chemistry) have the necessary energetic levels to realize OPVs with 10% or higher efficiency. This underscores the importance of carefully selecting the compounds to be synthesised and tested, and at the same time the value of the fast theoretical characterization and extensive search that CEP can provide towards this task. An unaided search has only a small chance of success, while CEP finds several thousand suitable structures. Fig. 4 also displays the dynamic gap range and dipole moment distribution of the screened candidates as examples for the wide range of electronic properties found in the CEP library. This versatility will be vital to a successful discovery of materials with specifically adjusted features (e.g., for different acceptors or tandem devices [160]).

We again emphasize that this first phase of CEP only

addresses a subset of the important material issues [161] and has the inherent accuracy of the employed model chemistries and calibrations, if taken at face value. There are, however, four factors that add considerable value to these calculations: i) we correlate the computed results to actual experimental quantities to provide insights into their relationship; ii) we use the electronic structure data as a source for new cheminformatics descriptors, which will put our modeling efforts on a more physical foundation; iii) the analysis of the aggregated data from millions of molecules in combination with structural similarity measures can reveal guiding trends, even if the absolute result for an individual candidate might be inaccurate due to a particular limitation of its electronic structure; iv) by employing a variety of different model chemistries, we compensate for the chance of such a failure in any particular method. We do not rely on any single result but use a composite scoring with many contributions.

The calculations allow for the elimination of unfit candidates and provide a first set of predictions for promising structural trends. For these we can carry out the subsequent levels of increasingly more sophisticated calculations which are planned as following: we will calculate i) vibrational spectra and partition functions to gauge phonon-scattering and trapping in vibrational modes [162, 163]; ii) anionic/cationic states for improved electron affinity/ionization potential values; a Dyson orbital analysis [164] can improve our insights into the charge transfer processes [165]; iii) optimized

geometries in the ionic states to deduce the reorganization during charge migration [163]; iv) triplet states and gaps to indicate potential for singlet fission processes [166]; v) linear response properties to assess the charge mobility [167, 168]; vi) excited states employing the maximum overlap method [169] and/or (range-separated) time-dependent DFT [170, 171] with an electron attachment/detachment density [172] and natural transition orbital analysis [173] for a more sound description of the absorption process. Each higher level result can be used to assess the interpretations at the lower levels. Further studies could involve the calculation of transfer integrals between oligomer pairs [163], packing and interactions in the bulk phase, as well as the use of high-level wavefunction theory for more reliable results in complicated bonding situations. We also want to consider the opposite approach, i.e., how well simple semiempirical and model Hamiltonian calculations (which are popular in other communities) perform compared to *first-principles* DFT [101, 174]. Possible stages of the CEP hierarchy were recently vetted in a successful proof-of-principle study: we predicted exceptional charge mobility in a novel organic semiconductor and this prediction could be confirmed experimentally [78]. We also used quantum chemical calculations in an earlier study to explain the observed mobility values in another system [175].

#### D. Empirical result calibration

To bridge the gap between theory and experiment, we have introduced an empirical calibration of the computational results. Such a calibration is a pragmatic way to approximately account and correct for differences in experimental and theoretical property definitions, as well as *in vacuo* vs. bulk, and oligomer vs. polymer results.

We have established the organic electronics 2011 (OE11) training set of experimentally well-characterized organic electronic materials for the calibration and aligned the theoretical findings with the corresponding data from experiment. The current calibration is largely based on data from bulk-heterojunction setups with PCBM as acceptor and introduces a corresponding bias. A different focus can be chosen provided appropriate reference data. The use of training sets is a common technique in other areas of quantum chemistry [121, 176]. The details of this work will be presented elsewhere. A preliminary calibration of the current CEP results was used in the analysis shown in Fig. 4, and panels (a-d) in particular demonstrate the success of this approach.

#### E. Database and data mining

The results of all calculations are used to build up a reference database – the Clean Energy Project Database (CEPDB). This data collection is comparable to the more

general but much smaller NIST Computational Chemistry Comparison and Benchmark Database (CCCBDB) [177]. As mentioned before, the information accumulated in CEPDB is not only relevant to OPVs but to organic electronics in general. It is also designed to provide benchmarks for the performance of various theoretical methods in this family of systems as well as a parameter repository for other calculations (e.g., for model Hamiltonians [178–180] or custom force fields [181]). It will be available to the public by 2012.

The primary purpose of CEPDB is to store and provide access to the CEP data. Candidates with a certain set of desired parameters can readily be identified. CEPDB also serves as the hub for all data mining, analysis, and scoring operations to facilitate the study of global trends, correlations, and OPV design rules. Finally, CEPDB is also responsible for bookkeeping, archiving the raw data (including the binary MO eigenvectors for subsequent calculations), keeping track of the project progress, and prioritizing the study, which are clearly important tasks considering the scope of the project and the volume of data.

#### F. The World Community Grid

The massive demand in computing time for CEP is largely provided by the World Community Grid [182, 183], a distributed volunteer computing platform for philanthropic research organized by IBM. Presently, ~560,000 users have signed up ~1,800,000 computers to the various WCG projects. Participants can donate computing time by running the supported science applications on their personal computers, either on low priority in the background or in screensaver mode during idle times (see Fig. 5). In CEP, we use a custom version of the Q-Chem 3.2 program package [184] which was ported to the Berkeley Open Infrastructure for Network Computing (BOINC) [185] environment. From the user perspective, the participation in a WCG project is fully automated and usually does not require any input or maintenance beyond the initial setup.

The WCG is a powerful resource and provides us with the means for our high-throughput investigation. It is, however, also unusual due to the non-specialized hardware and host demands. These limitations have to be taken into consideration in the design of suitable tasks. In addition to the WCG, CEP utilizes the Harvard FAS Odyssey cluster as well as accounts at NERSC [186] and TeraGrid [187] for problems which are outside the scope of volunteer grid computing (e.g., due to their size and computational complexity). CEP currently characterizes about 20,000 oligomer sequences per day, and so far we have studied ~2,600,000 structures in ~35,000,000 calculations, utilizing ~4,500 years of CPU time. Close to 100% of these calculations were performed on the grid, but the use of cluster resources will become more important in the more demanding stages of the project. Their



FIG. 5: Top: CEP project homepage and hub for project participants. Bottom: The CEP 'screensaver' as it is displayed on volunteer hosts.

contribution in terms of volume, however, will remain very limited.

One important aspect in a computational study at this scale is that all processes have to be automated to keep it feasible, efficient, and reduce human error. We created the necessary facilities using the Python scripting language [188, 189].

### III. CONCLUSIONS

This initial presentation of the Harvard Clean Energy Project outlined its overall architecture, the machinery we put in place, and upcoming extensions. We pointed to first results and more detailed reports on the various aspects of the project – tied together by this perspective – will be given in subsequent publications.

CEP applies a modern cyberinfrastructure paradigm to computational materials science and in particular to renewable energy research. Engineering successful OPV materials is a complex challenge as a number of optical and electronic requirements have to be met. We showed that CEP with its large-scale screening is well equipped for a knowledge-based search of systems with suitable features. It provides a unique access to data for a wide array of potential compounds with diverse electronic struc-

tures and it is ideally suited to identify highly promising donor or acceptor candidates in the infinite space of organic electronics. Our work can thus complement and accelerate traditional research approaches, and it can help develop an understanding of structure-property relationships to facilitate the rational design of new materials. We hope that joint efforts with experimental collaborators can contribute to overcoming the current limitations of OPV materials in order to provide a clean source of electricity which can compete with conventional power production.

### Acknowledgments

We wish to thank IBM for organizing the World Community Grid and the WCG members for their computing time donations. Additional information about the project as well as links to join the WCG and CEP can be found in Refs. [44] and [182].

We are grateful to our experimental collaborators Zhenan Bao, Rajib Mondal, and Anatoliy N. Sokolov at Stanford University; for support from the Q-Chem development team by Yihan Shao, Alex J. Sodt, and Zhengting Gan; from the IBM WCG team by Jonathan D. Armstrong, Kevin N. Reed, Keith J. Uplinger, Al Seipel, Viktors Berstis, and Bill Bovermann; from Harvard FAS Research Computing by Suvendra Dutta, Christopher Walker, Gerald I. Lotto, and James Cuff; to Aidan Daly, Erica Lin, and Lauren A. Kaye at Harvard.

The CEP is supported by the National Science Foundation through grants No. DMR-0820484 and DMR-0934480 as well as the Global Climate and Energy Project (Stanford grant No. 25591130-45282-A. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of Stanford University, the Sponsors of the Global Climate and Energy Project, or others involved with the Global Climate and Energy Project). ROA was supported by a Giorgio Ruffolo Fellowship in the Sustainability Science Program at Harvard University's Center for International Development, CAB acknowledges financial support from CONACyT México and computer access to DGSCA UNAM, AGP acknowledges funding by the Harvard College Research Program and the Harvard University Center for the Environment Summer Research Fund, RSSC thanks the Mary Fieser Postdoctoral Fellowship at Harvard University, LV was supported by an NSF Graduate Research Fellowship and AMB by the Harvard REU program.

We acknowledge software support from Q-Chem Inc., Optibrium Ltd., Dotmatics Ltd., Molecular Networks GmbH, and ChemAxon Ltd. Additional computing time was provided on the Odyssey cluster supported by the FAS Research Computing Group, NERSC, and the National Science Foundation TeraGrid.



### Biographies:

Johannes Hachmann obtained his Dipl.-Chem. degree (2004) after undergraduate studies at the University of Jena, Germany, and the University of Cambridge, UK. He received an M.Sc. (2007) and Ph.D. (2010) from Cornell University under the supervision of Prof. G.K.-L. Chan working on density matrix renormalization group theory and computational transition metal chemistry. In 2009 he joined the Aspuru-Guzik Group and the Clean Energy Project at Harvard University.

Roberto Olivares-Amaya is a Ph.D. Chemical Physics student in the Aspuru-Guzik Group at Harvard University and was a 2010-2011 Giorgio Ruffolo Sustainability Science Doctoral Fellow. He obtained his B.Sc. degree (2007) from the National Autonomous University of Mexico (UNAM). His research interests include the study of structure and response of OPV materials by means of *first-principles* electronic structure and machine learning methods.

Sule Atahan-Evrenk is a postdoctoral research fellow in the Aspuru-Guzik Group at Harvard University. She received a B.S. degree in chemistry from Bilkent University, Turkey, and a Ph.D. in physical chemistry from the University of Maryland College Park. Her research interests include the theoretical characterization of organic electronic materials by quantum chemistry and molecular dynamics methods.

Carlos Amador-Bedolla is a professor at the Facultad de Química of the National Autonomous University of Mexico in Mexico City. He obtained his doctorate in chemistry at UNAM (1989) and has worked with several leading groups at Case Western Reserve University, UC Berkeley, and Harvard University. His current research focuses on renewable energy materials (see [www.quimica.unam.mx/ficha\\_investigador.php?ID=77&tipo=2](http://www.quimica.unam.mx/ficha_investigador.php?ID=77&tipo=2)).

Roel S. Sánchez-Carrera is currently a research staff at the Robert Bosch Research and Technology Center North America. He obtained his Ph.D. (2008) from the Georgia Institute of Technology under the supervision of Prof. J.-L. Brédas. After completion, he moved to Harvard University, where he worked in the group of Prof. Aspuru-Guzik as a Mary-Fieser Postdoctoral Fellow. His present research focuses on *ab-initio* modeling of materials for energy applications.

Aryeh Gold-Parker is an undergraduate student at Harvard University and will complete his B.A. degree in chemistry and physics in 2012. His research interests are in physical chemistry, specifically in renewable energy materials. He joined the Aspuru-Guzik Group in 2009 to work on the Clean Energy Project.

Leslie Vogt completed a B.S. in physical chemistry at Juniata College in 2005 and spent 2006-2007 working with Prof. R. Zimmermann as a Fulbright Fellow in Germany. She is completing her Ph.D. at Harvard University in the Aspuru-Guzik Group studying the electronic properties of organic materials as an NSF Graduate Research Fellow.

Anna M. Brockway is a chemistry senior at Haverford

College where she works with Prof. J. Schrier in the field of computational theoretical chemistry. She joined the Clean Energy Project in 2010 during her time as an REU summer student at Harvard University.

Alán Aspuru-Guzik holds a Ph.D. in physical chemistry from UC Berkeley, and is currently an associate professor of Chemistry and Chemical Biology at Harvard University. His research lies at the intersection of quantum information/computation and theoretical chemistry. He is interested in energy transfer dynamics and renewable energy materials.

(See <http://aspuru.chem.harvard.edu/> and the project webpage <http://cleanenergy.harvard.edu/>)

### Quotes:

The Clean Energy Project stands out from other computational materials science approaches as it combines conventional modeling with strategies from modern drug discovery: CEP features an automated, high-throughput infrastructure for a systematic screening of millions of OPV candidates at a *first-principles* electronic structure level.

A suitable molecular motif is a *necessary* condition for a successful OPV development.

CEP applies a modern cyberinfrastructure paradigm to computational materials science and in particular to renewable energy research.



FIG. 6: ToC graphic: The Harvard Clean Energy Project utilizes idle computing time on the personal computers of project volunteers – organized by IBM’s World Community Grid – to perform a large-scale quantum chemical screening of material candidates for organic photovoltaics. (Cover art created by Lauren Aleza Kaye.)

### Keywords:

organic photovoltaics; distributed volunteer computing; quantum chemistry; high-throughput screening; computational materials science; World Community Grid; reference database

- 
- [1] D. Wöhrle and D. Meissner, *Adv. Mater.* **3**, 129 (1991).  
 [2] J. J. Conti et al., *Annual Energy Outlook 2011*, U.S. Energy Information Administration, Washington, DC, 2011.  
 [3] K. Zweibel, J. Mason, and V. Fthenakis, *Sci. Am.* **298**, 64 (2008).  
 [4] M. Z. Jacobson and M. A. Delucchi, *Sci. Am.* **301**, 58 (2009).  
 [5] N. S. Lewis, *Science* **315**, 798 (2007).  
 [6] M. Boxwell, *Solar Electricity Handbook – 2011 Edition*, Greenstream Publishing, Ryton on Dunsmore, UK, 4th edition, 2011.  
 [7] C. J. Brabec, V. Dyakonov, J. Parisi, and N. S. Sariciftci, *Organic Photovoltaics: Concepts and Realization*, Springer, Berlin, 2010.  
 [8] C. J. Brabec, V. Dyakonov, and U. Scherf, *Organic Photovoltaics: Materials, Device Physics, and Manufacturing Technologies*, Wiley-VCH, Weinheim, 2008.  
 [9] S.-S. Sun and N. S. Sariciftci, *Organic Photovoltaics: Mechanisms, Materials, and Devices*, CRC Press, Boca Raton, 2005.  
 [10] C. W. Tang, *Appl. Phys. Lett.* **48**, 183 (1986).  
 [11] N. Beaumont, I. Hancox, P. Sullivan, R. A. Hatton, and T. S. Jones, *Energy Environ. Sci.* **4**, 1708 (2011).  
 [12] M. Riede, T. Mueller, W. Tress, R. Schueppel, and K. Leo, *Nanotechnology* **19**, 424001 (2008).  
 [13] K. Kalyanasundaram, *Dye-sensitized Solar Cells*, CRC Press, Boca Raton, 2010.  
 [14] F. C. Krebs, *Polymer Photovoltaics: A Practical Approach*, SPIE Publications, Bellingham, WA, 2008.  
 [15] M. Pagliaro, G. Palmisano, and R. Ciriminna, *Flexible Solar Cells*, Wiley-VCH, Weinheim, 2008.  
 [16] M. A. Green, *Physica E* **14**, 65 (2002).  
 [17] Y. Galagan et al., *Chem. Eng. Process.* **50**, 454 (2011).  
 [18] F. C. Krebs, T. Tromholt, and M. Jørgensen, *Nanoscale* **2**, 873 (2010).  
 [19] D. Vak et al., *Appl. Phys. Lett.* **91**, 081102 (2007).  
 [20] R. Green et al., *Appl. Phys. Lett.* **92**, 33301 (2008).  
 [21] K. X. Steirer et al., *Sol. Energy Mater. Sol. Cells* **93**, 447 (2009).

- [22] M. O'Neill and S. M. Kelly, *Adv. Mater.* **23**, 566 (2011).
- [23] R. F. Bailey-Salzman, B. P. Rand, and S. R. Forrest, *Appl. Phys. Lett.* **88**, 233502 (2006).
- [24] A. Colsmann et al., *Adv. Energy Mater.* **1**, 599 (2011).
- [25] M. Pagliaro, R. Ciriminna, and G. Palmisano, *ChemSusChem* **1**, 880 (2008).
- [26] A. Heeger, Low-cost plastic solar cells: a dream becoming a reality, in *Global Sustainability - A Nobel Cause*, edited by H. J. Schellnhuber, M. Molina, N. Stern, V. Huber, and S. Kadner, Cambridge University Press, Cambridge, UK, 2010.
- [27] S. R. Forrest, *Nature* **428**, 911 (2004).
- [28] G. Marsh, *Renew. Energy Focus* **9**, 62 (2008).
- [29] A. Henemann, *Renew. Energy Focus* **9**, 14, 16 (2008).
- [30] R. Tipnis et al., *Sol. Energy Mater. Sol. Cells* **93**, 442 (2009).
- [31] M. B. Schubert and J. H. Werner, *Mater. Today* **9**, 42 (2006).
- [32] C. J. Brabec, *Sol. Energy Mater. Sol. Cells* **83**, 273 (2004).
- [33] A. Bedeloglu, A. Demir, Y. Bozkurt, and N. S. Sariciftci, *Synth. Met.* **159**, 2043 (2009).
- [34] [http://wgisi.org/files/EquinoxCommunique\\_June9\\_2011.pdf](http://wgisi.org/files/EquinoxCommunique_June9_2011.pdf).
- [35] V. Shrotriya, *Nat. Photonics* **3**, 447 (2009).
- [36] C. H. Peters et al., *Adv. Energy Mater.* **in print**, 491 (2011).
- [37] R. F. Service, *Science* **332**, 293 (2011).
- [38] A. Rivaton et al., *Polym. Degrad. Stab.* **95**, 278 (2010).
- [39] M. Manceau, A. Rivaton, J.-L. Gardette, S. Guillerez, and N. Lemaître, *Sol. Energy Mater. Sol. Cells* **95**, 1315 (2011).
- [40] S. Bertho et al., *Sol. Energy Mater. Sol. Cells* **91**, 385 (2007).
- [41] S. Bertho et al., *Sol. Energy Mater. Sol. Cells* **92**, 753 (2008).
- [42] R. Gaudiana and C. Brabec, *Nat. Photonics* **2**, 287 (2008).
- [43] S. E. Shaheen, D. S. Ginley, and G. E. Jabbour, *MRS Bull.* **30**, 10 (2005).
- [44] <http://cleanenergy.harvard.edu> (accessed August 10, 2011).
- [45] J.-L. Brédas, J. E. Norton, J. Cornil, and V. Coropceanu, *Acc. Chem. Res.* **42**, 1691 (2009).
- [46] P. Heremans, D. Cheyns, and B. P. Rand, *Acc. Chem. Res.* **42**, 1740 (2009).
- [47] J. Chen and Y. Cao, *Acc. Chem. Res.* **42**, 1709 (2009).
- [48] C. W. Schlenker and M. E. Thompson, *Chem. Commun.* **47**, 3702 (2011).
- [49] H. Borchert, *Energy Environ. Sci.* **3**, 1682 (2010).
- [50] N. Blouin et al., *J. Am. Chem. Soc.* **130**, 732 (2008).
- [51] J. Nelson, J. J. Kwiatkowski, J. Kirkpatrick, and J. M. Frost, *Acc. Chem. Res.* **42**, 1768 (2009).
- [52] X.-Y. Zhu, Q. Yang, and M. Muntwiler, *Acc. Chem. Res.* **42**, 1779 (2009).
- [53] L. Wang et al., *Chem. Soc. Rev.* **39**, 423 (2010).
- [54] E. H. Kerns, *J. Pharm. Sci.* **90**, 1838 (2001).
- [55] J. Bajorath, *Nat. Rev. Drug Discovery* **1**, 882 (2002).
- [56] C. McInnes, *Curr. Opin. Chem. Biol.* **11**, 494 (2007).
- [57] P. Gribbon and S. Andreas, *Drug Discovery Today* **10**, 17 (2005).
- [58] K. Appell, J. J. Baldwin, and W. J. Egan, Combinatorial chemistry and high-throughput screening in drug discovery and development, in *Handbook of Modern Pharmaceutical Analysis*, edited by S. Ahuja and S. Scypinski, volume 3 of *Separation Science and Technology*, pages 23–56, Academic Press, San Diego, 2001.
- [59] E. M. Gordon, M. A. Gallop, and D. V. Patel, *Acc. Chem. Res.* **29**, 144 (1996).
- [60] J. A. Ellman, *Acc. Chem. Res.* **29**, 132 (1996).
- [61] E. M. Gordon, R. W. Barrett, W. J. Dower, S. P. A. Fodor, and M. A. Gallop, *J. Med. Chem.* **37**, 1385 (1994).
- [62] The authors are aware of and acknowledge related work by Geoffrey Hutchison (University of Pittsburgh), Michel Côté (University of Montreal), Gerbrand Ceder (Massachusetts Institute of Technology), and Josef Michl (University of Colorado). Their respective research efforts each have a certain degree of overlap with CEP.
- [63] J. Bajorath, *Chemoinformatics: Concepts, Methods, and Tools for Drug Discovery*, Humana Press, Totowa, NJ, 2010.
- [64] A. K. Ghose, T. Herbertz, J. M. Salvino, and J. P. Mal-lamo, *Drug Discovery Today* **11**, 1107 (2006).
- [65] M. M. Gaber, *Scientific Data Mining and Knowledge Discovery: Principles and Foundations*, Springer, Berlin, 2009.
- [66] R. L. Grossman, C. Kamath, P. Kegelmeyer, V. Kumar, and R. Namburu, *Data Mining for Scientific and Engineering Applications*, Springer, Berlin, 2001.
- [67] C. Suh and K. Rajan, *Mater. Sci. Technol.* **25**, 466 (2009).
- [68] K. Rajan, *Annu. Rev. Mater. Res.* **38**, 299 (2008).
- [69] C. Suh and K. Rajan, *QSAR Comb. Sci.* **24**, 114 (2005).
- [70] A. Jain et al., *Comput. Mater. Sci.* **50**, 2295 (2011).
- [71] S. Colella et al., *Chem. Commun.* **46**, 6273 (2010).
- [72] X. Zhan and D. Zhu, *Polym. Chem.* **1**, 409 (2010).
- [73] Y. Liang et al., *J. Am. Chem. Soc.* **131**, 7792 (2009).
- [74] C. L. Chochos and S. A. Choulis, *Prog. Polym. Sci.* **36**, 1326 (2011).
- [75] J. Liu et al., *Polymer* **52**, 2302 (2011).
- [76] A. Mishra, M. K. R. Fischer, and P. B auerle, *Angew. Chem. Int. Ed.* **48**, 2474 (2009).
- [77] P. M. Beaujuge, C. M. Amb, and J. R. Reynolds, *Acc. Chem. Res.* **43**, 1396 (2010).
- [78] A. N. Sokolov et al., *Nature Commun.* **in print**, (2011).
- [79] A. Facchetti, *Chem. Mater.* **23**, 733 (2011).
- [80] A. Pron and P. Rannou, *Prog. Polym. Sci.* **27**, 135 (2002).
- [81] D. Weininger, *J. Chem. Inf. Comput. Sci.* **28**, 31 (1988).
- [82] Marvin 5.3.2 by ChemAxon, 2010, <http://www.chemaxon.com>.
- [83] J. Sadowski, J. Gasteiger, and G. Klebe, *J. Chem. Inf. Comput. Sci.* **34**, 1000 (1994).
- [84] R. Farchioni and G. Grosso, *Organic Electronic Materials: Conjugated Polymers and Low Molecular Weight Organic Solids*, Springer, Berlin, 2001.
- [85] G. Meller and T. Grasser, *Organic Electronics*, Springer, Berlin, 2009.
- [86] W. Paszkowicz, *Mater. Manuf. Processes* **24**, 174 (2009).
- [87] V. Arora and A. Bakhshi, *Chem. Phys.* **373**, 307 (2010).
- [88] R. Giro, M. Cyrillo, and D. S. Galv o, *Chem. Phys. Lett.* **366**, 170 (2002).
- [89] B. V. C. Martins, G. Brunetto, F. Sato, V. R. Coluci, and D. S. Galv o, *Chem. Phys. Lett.* **453**, 290 (2008).
- [90] E.-W. Lameijer, T. B ack, J. Kok, and A. D. Ijzerman, *Natural Comput.* **4**, 177 (2005).

- [91] R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2002.
- [92] V. J. Sykora and D. E. Leahy, *J. Chem. Inf. Model.* **48**, 1931 (2008).
- [93] S. Mitra, S. Datta, T. Perkins, and G. Michailidis, *Introduction to Machine Learning and Bioinformatics*, CRC Press, Boca Raton, 2008.
- [94] Z. R. Yang, *Machine Learning Approaches to Bioinformatics*, World Scientific, Singapore, 2010.
- [95] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, Berlin, 2nd edition, 2007.
- [96] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, San Diego, 2nd edition, 1990.
- [97] R. Olivares-Amaya et al., *Energy Environ. Sci.* **submitted**, (2011).
- [98] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2008.
- [99] Dragon 6 by Talete, 2011, <http://www.talete.mi.it/index.htm>.
- [100] StarDrop 5.0 by Optibrium, 2011, <http://www.optibrium.com/stardrop/>.
- [101] J.-M. André, J. Delhalle, and J.-L. Brédas, *Quantum Chemistry Aided Design of Organic Polymers: An Introduction to the Quantum Chemistry of Polymers and Its Applications*, World Scientific, Singapore, 1991.
- [102] C. Risko, M. D. McGehee, and J.-L. Brédas, *Chem. Sci.* **2**, 1200 (2011).
- [103] L. Wang, Q. Li, Z. Shuai, L. Chen, and Q. Shi, *Phys. Chem. Chem. Phys.* **12**, 3309 (2010).
- [104] T. Benanti and D. Venkataraman, *Photosynth. Res.* **87**, 73 (2006).
- [105] B. A. Gregg, *Excitonic Solar Cells: The Physics and Chemistry of Organic-Based Photovoltaics*, chapter 19, pages 243–257, ACS, Washington, DC, 2003.
- [106] S. Xiao, S. C. Price, H. Zhou, and W. You, *Recent Progress on Highly Efficient Bulk Heterojunction Polymer Solar Cells*, chapter 7, pages 71–80, ACS, Washington, DC, 2010.
- [107] M. C. Scharber et al., *Adv. Mater.* **18**, 789 (2006).
- [108] R. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
- [109] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, Weinheim, 2nd edition, 2001.
- [110] A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- [111] J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).
- [112] C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- [113] A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- [114] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [115] J. P. Perdew, M. Ernzerhof, and K. Burke, *J. Chem. Phys.* **105**, 9982 (1996).
- [116] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **78**, 1396 (1997).
- [117] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
- [118] J. P. Perdew, J. Tao, V. N. Staroverov, and G. E. Scuseria, *J. Chem. Phys.* **120**, 6898 (2004).
- [119] C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
- [120] A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).
- [121] Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.* **120**, 215 (2008).
- [122] Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.* **41**, 157 (2008).
- [123] W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* **51**, 2657 (1969).
- [124] W. J. Hehre, R. Ditchfield, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* **52**, 2769 (1970).
- [125] F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
- [126] G. Zhang and C. B. Musgrave, *J. Phys. Chem. A* **111**, 1554 (2007).
- [127] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
- [128] S. Grimme, *J. Chem. Phys.* **124**, 034108 (2006).
- [129] R. S. Mulliken, *J. Chem. Phys.* **23**, 1833 (1955).
- [130] K. B. Wiberg and P. R. Rablen, *J. Comp. Chem.* **14**, 1504 (1993).
- [131] A. E. Reed and F. Weinhold, *J. Chem. Phys.* **78**, 4066 (1983).
- [132] A. E. Reed, R. B. Weinstock, and F. Weinhold, *J. Chem. Phys.* **83**, 735 (1985).
- [133] A. E. Reed and F. Weinhold, *J. Chem. Phys.* **83**, 1736 (1985).
- [134] J. P. Foster and F. Weinhold, *J. Am. Chem. Soc.* **102**, 7211 (1980).
- [135] A. E. Reed, L. A. Curtiss, and F. Weinhold, *Chem. Rev.* **88**, 899 (1988).
- [136] J. B. Foresman and A. Frisch, *Exploring Chemistry With Electronic Structure Methods*, Gaussian Inc., Pittsburgh, PA, 2nd edition, 1996.
- [137] R. Stowasser and R. Hoffmann, *J. Am. Chem. Soc.* **121**, 3414 (1999).
- [138] S. Hamel, P. Duffy, M. E. Casida, and D. R. Salahub, *J. Electron. Spectrosc. Relat. Phenom.* **123**, 345 (2002).
- [139] D. P. Chong, O. V. Gritsenko, and E. J. Baerends, *J. Chem. Phys.* **116**, 1760 (2002).
- [140] J. Luo, Z. Q. Xue, W. M. Liu, J. L. Wu, and Z. Q. Yang, *J. Phys. Chem. A* **110**, 12005 (2006).
- [141] C.-G. Zhan, J. A. Nichols, and D. A. Dixon, *J. Phys. Chem. A* **107**, 4184 (2003).
- [142] P. Ayers and J. Melin, *Theor. Chem. Acc.* **117**, 371 (2007).
- [143] Z. Shuai, L. Wang, and Q. Li, *Adv. Mater.* **23**, 1145 (2011).
- [144] G. D. Scholes and G. Rumbles, *Nature Mater.* **5**, 683 (2006).
- [145] M. Jaiswal and R. Menon, *Polym. Int.* **55**, 1371 (2006).
- [146] A. Köhler et al., *Nature* **392**, 903 (1998).
- [147] H. A. Pohl, *J. Polym. Sci. C: Polym. Symp.* **17**, 13 (1967).
- [148] N. T. Maitra, *J. Chem. Phys.* **122**, 234104 (2005).
- [149] M. J. G. Peach, P. Benfield, T. Helgaker, and D. J. Tozer, *J. Chem. Phys.* **128**, 044118 (2008).
- [150] J. Fabian, M. Mann, and M. Petiau, *J. Mol. Model.* **6**, 177 (2000).
- [151] J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, *J. Chem. Phys.* **127**, 134309 (2007).
- [152] J. D. Dunitz and A. Gavezzotti, *Chem. Soc. Rev.* **38**, 2622 (2009).
- [153] J. H. Williams, *Acc. Chem. Res.* **26**, 593 (1993).
- [154] P. M. Borsenberger and J. J. Fitzgerald, *J. Phys. Chem.* **97**, 4815 (1993).
- [155] R. H. Young and J. J. Fitzgerald, *J. Phys. Chem.* **99**,

- 4230 (1995).
- [156] S. S. Zade, N. Zamoshchik, and M. Bendikov, *Acc. Chem. Res.* **44**, 14 (2011).
- [157] J. Gierschner, J. Cornil, and H.-J. Egelhaaf, *Adv. Mater.* **19**, 173 (2007).
- [158] S. S. Zade and M. Bendikov, *Org. Lett.* **8**, 5243 (2006).
- [159] J. J. Ladik, *J. Mol. Struct. THEOCHEM* **230**, 127 (1991).
- [160] T. Ameri, G. Dennler, C. Lungenschmied, and C. J. Brabec, *Energy Environ. Sci.* **2**, 347 (2009).
- [161] G. Heimel, I. Salzmann, S. Duhm, and N. Koch, *Chem. Mater.* **23**, 359 (2011).
- [162] A. Troisi, *Chem. Soc. Rev.* **40**, 2347 (2011).
- [163] J.-L. Brédas, D. Beljonne, V. Coropceanu, and J. Cornil, *Chem. Rev.* **104**, 4971 (2004).
- [164] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One and Two Electron Atoms*, Plenum, New York, 1977.
- [165] R. Singh and M. Mishra, *J. Chem. Sci.* **121**, 867 (2009).
- [166] M. B. Smith and J. Michl, *Chem. Rev.* **110**, 6891 (2010).
- [167] H. Houili, J. D. Picon, L. Zuppiroli, and M. N. Bussac, *J. Appl. Phys.* **100**, 023702 (2006).
- [168] C. She et al., *J. Phys. Chem. A* **114**, 3384 (2010).
- [169] A. T. B. Gilbert, N. A. Besley, and P. M. W. Gill, *J. Phys. Chem. A* **112**, 13164 (2008).
- [170] D. Jacquemin et al., *J. Chem. Phys.* **126**, 144105 (2007).
- [171] T. Le Bahers, T. Pauporte, G. Scalmani, C. Adamo, and I. Ciofini, *Phys. Chem. Chem. Phys.* **11**, 11276 (2009).
- [172] M. Head-Gordon, A. M. Grana, D. Maurice, and C. A. White, *J. Phys. Chem.* **99**, 14261 (1995).
- [173] R. L. Martin, *J. Chem. Phys.* **118**, 4775 (2003).
- [174] G. R. Hutchison, M. A. Ratner, and T. J. Marks, *J. Phys. Chem. A* **106**, 10596 (2002).
- [175] R. S. Sánchez-Carrera, S. Atahan, J. Schrier, and A. Aspuru-Guzik, *J. Phys. Chem. C* **114**, 2334 (2010).
- [176] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **106**, 1063 (1997).
- [177] NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database No 101, Rel. 15a, April 2010, Editor: Russell D. Johnson III, <http://cccbdb.nist.gov/> (accessed August 10, 2011).
- [178] D. L. Cheung and A. Troisi, *Phys. Chem. Chem. Phys.* **10**, 5941 (2008).
- [179] E. San-Fabián, A. Guijarro, J. Vergés, G. Chiappe, and E. Louis, *Eur. Phys. J. B* **81**, 253 (2011).
- [180] J. A. Vergés, E. San-Fabián, G. Chiappe, and E. Louis, *Phys. Rev. B* **81**, 085120 (2010).
- [181] T. Trevethan, A. Shluger, and L. Kantorovich, *J. Phys. Condens. Matter* **22**, 084024 (2010).
- [182] [www.worldcommunitygrid.org](http://www.worldcommunitygrid.org) (accessed August 10, 2011).
- [183] D. Clery, *Science* **308**, 773 (2005).
- [184] Y. Shao et al., *Phys. Chem. Chem. Phys.* **8**, 3172 (2006).
- [185] D. P. Anderson, BOINC: A system for public-resource computing and storage, in *Proceedings of the 5th IEEE/ACM International Workshop on Grid Computing*, GRID '04, pages 4–10, Washington, DC, 2004, IEEE Computer Society.
- [186] [www.nersc.gov](http://www.nersc.gov), (accessed June 15, 2011).
- [187] C. Catlett et al., *Adv. Parallel Comput.* **16**, 225 (2008).
- [188] D. M. Beazley, *Python Essential Reference*, Addison-Wesley, Upper Saddle River, NJ, 4th edition, 2009.
- [189] H. P. Langtangen, *Python Scripting for Computational Science*, Springer, Berlin, 3rd edition, 2010.