Are Coarse-Grained Structures as Good as Atomistic Ones for Calculating the Electronic Properties of Organic Semiconductors?

Otello Maria Roscioni, Matteo Ricci, Claudio Zannoni, Gabriele d'Avino

Are Coarse-Grained Structures as Good as Atomistic Ones for Calculating the Electronic Properties of Organic Semiconductors?

Otello Maria Roscioni,[*,†] Matteo Ricci,[†] Claudio Zannoni,[‡] and Gabriele D'Avino[*,¶]

†*MaterialX LTD, BS50HE Bristol, United Kingdom*

‡*Dipartimento di Chimica Industriale "Toso Montanari", Università di Bologna, Viale Risorgimento 4, IT-40136 Bologna, Italy*

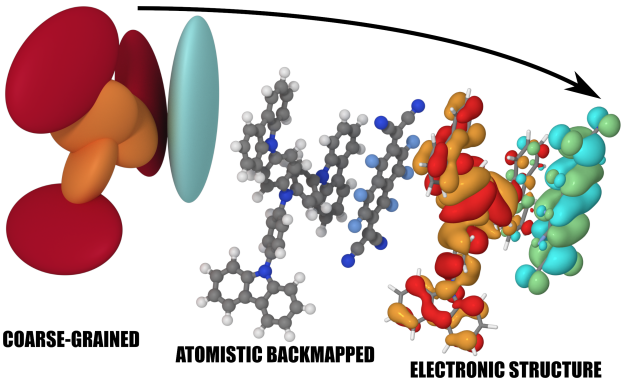
¶*Grenoble Alpes University, CNRS, Grenoble INP, Institut Néel, 25 rue des Martyrs, 38042 Grenoble, France*

E-mail: om.roscioni@materialx.co.uk; gabriele.davino@neel.cnrs.fr

Abstract

The quality of amorphous molecular morphologies obtained with a recently-introduced coarse-grained model, representing molecules in terms of connected anisotropic beads (Phys.Chem.Chem.Phys., 2019, 21, 26195), is benchmarked against reference atomistic data. Typical small-molecule organic semiconductors in their pristine and doped form are chosen as a challenging and technologically relevant case study for our comparison, which includes both structural features and the resulting electronic properties, such as charge carrier energy levels, energetic disorder and intermolecular charge transfer couplings. Our analysis shows that our accurate coarse-grained model leads to molecular glasses that are very similar to native atomistic samples, with the discrepancy being further reduced upon backmapping. The electronic properties computed for backmapped morphologies are almost indistinguishable from the atomistic reference, especially for multi-branched poly(hetero)cyclic hydrocarbons usually employed as organic semiconductors. This study provides a proof of principle for highly-accurate large-scale simulations of complex molecular systems at a reduced computational cost.

TOC Graphic



1 Introduction

Atomistic molecular dynamics (MD) corresponds to a bottom-up modelling and simulation approach aimed at calculating properties of materials, where molecules are described using all the constituting atoms, connected according to the structural formula and represented as classical spherical particles subject to a certain set of interactions (Force Field)[1–4] tuned at least in part with Quantum Mechanics (QM) calculations. Atomistic MD simulations have proved to yield reliable morphologies and their changes according to thermodynamic state for a variety of organic functional materials or systems of biological interest. Indeed, these simulations can provide realistic physical properties for a variety of soft materials, e.g. order parameters and phase transition temperatures for low molar mass liquid crystals or lipid membranes with structural results that compare well with experimental ones (X-ray, NMR, Raman, etc.)[5,6] In addition, the detailed atomistic information of a condensed-phase molecular system can readily be used as an input for subsequent QM calculations. This strategy has been largely applied in the field of organic semiconductors we are interested in this work, where charge transfer rates based on Marcus theory, fed with electronic structure calculation inputs (e.g. transfer integrals, energy levels, etc.), provided fair estimates of charge transport properties.[7–12]

In many ways atomistic MD with properly tuned force fields together with appropriate QM calculations thus provide a "gold standard" for predicting the electronic and transport properties needed for designing or optimizing at molecular level the materials for organic electronics.[8,13] Unfortunately, molecules of interest for organic electronics, even the so called "low molar mass" ones, are rather complex in terms of the number of atoms (of the order of one hundred atoms or more as we see for example in Fig.1), and of the sample size required (say nanometers). Thus, even if efficient atomistic MD packages exist, like LAMMPS,[3,14] GROMACS,[15] NAMD[16] the cost in terms of computational resources and actual turnaround time makes the all-atoms approach hardly feasible for the predictive workflow expected by industry.[17]

In view of this, coarse-grained (CG) molecular models, in which the full atomistic detail is replaced by a representation of molecules in terms of a relatively small number of effective particles (beads), represent a great opportunity to extend the reach of classical molecular dynamics simulations to large and complex molecular systems and to length and time scales that are prohibitive at the atomistic level, e.g. to model static and dynamic properties of biological systems like lipids,[18] proteins,[19,20] DNA,[21] functional materials[22–24] or polymers.[25–28] While CG models are often adequate to provide a basic understanding of the relation between molecular structure and phase organization, they inevitably lead to losing the information about atomic positions, which would be also needed for the calculation of any electronic property of interest at the QM level. To remedy this, it is important to develop reversible CG models, where atomic positions can be back-mapped from the position and orientations of the CG beads.[24,29,30]

In the context of speeding up the statistical sampling of quantum properties though coarse graining, it is worth mentioning a series of novel machine learning approaches to obtain electronic properties directly from a reduced subset of structural degrees of freedom,[31,32] in some case obtained from CG simulations,[33] bypassing the atomistic back-mapping. While these techniques are definitely promising in the perspective of drastically reducing the computational cost associated with electronic structure calculations, they still hinge on the availability of atomistic structures to build training and test sets, as well as accurate CG force fields.

In this work, we use the MOLC model,[24] a CG force field that is particularly suited for poly(hetero)cyclic aromatic compounds, as it is based on replacing relatively rigid conjugated moieties with overlapping biaxial Gay-Berne ellipsoids[6,34,35] connected with directional bonds and decorated with off-center charges.[36] As a result, CG molecules retain some of the internal degrees of freedom, such as rotations around selected bonds, and include anisotropic electrostatic interactions. Using anisotropic, instead of the spherical beads as, for instance, in the MARTINI package often used for modelling biomolecules,[37] or in the VOTCA package for organic semiconductors[8,38] adds some technical complications, but has advantages for

back-mapping and in representing molecular shapes. Indeed a main feature of MOLC is the ability to reproduce the molecular excluded volume and thus molecular packing, which in turn yields accurate mass density and other condensed-phase properties for both amorphous solids[24] and liquids.[39] Somewhat similar approaches have also been developed recently by Jackson, de Pablo and coworkers,[27,40] and by Khot and Savoie.[41]

A legitimate question for any CG approach concerns the price one has to pay for downgrading the level of resolution, i.e. the possible sacrifice of the fine accuracy of the resulting structure.[42] This is particularly important whenever one is interested in the electronic properties of a given system and especially in using the simulated structure as inputs for electronic structure calculations. Electronic properties are notably highly sensitive to the molecular geometry and the supramolecular packing in condensed phases,[7,43] making accurate morphologies a necessary prerequisite for carrying out reliable electronic structure calculations. In the present paper, we hence propose a comparative study of CG morphologies against reference atomistic ones, by taking the electronic structure as a severe benchmarking criterion.

Our study is conducted on non-polymeric organic semiconductors, representing a technologically relevant example of disordered and possibly heterogeneous molecular systems. We specifically focus on two well known and experimentally studied hole-transport materials, 2,2',7,7'-tetrakis(N,N-diphenylamino)-9,9-spirobifluorene (spiro-TAD)[44,45] and tris(4-carbazoyl-9-ylphenyl)amine (TCTA),[45,46] in their pristine form and mixed with the p-type dopant 2,2'-(perfluoronaphthalene-2,6-diylidene)dimalononitrile (F6TCNNQ, see Figure 1).[47] We emphasize that both spiro-TAD and TCTA are fairly-large flexible molecules presenting several soft degrees of freedom that in our model are associated with inter-bead torsional and bending angles. These features make the application of the MOLC model particularly relevant and challenging at the same time.

In order to investigate the electronic properties important for charge transport in organic semiconductors, a comprehensive multiscale procedure is deployed. This starts with the creation and equilibration of pristine and doped samples of organic semiconductors with

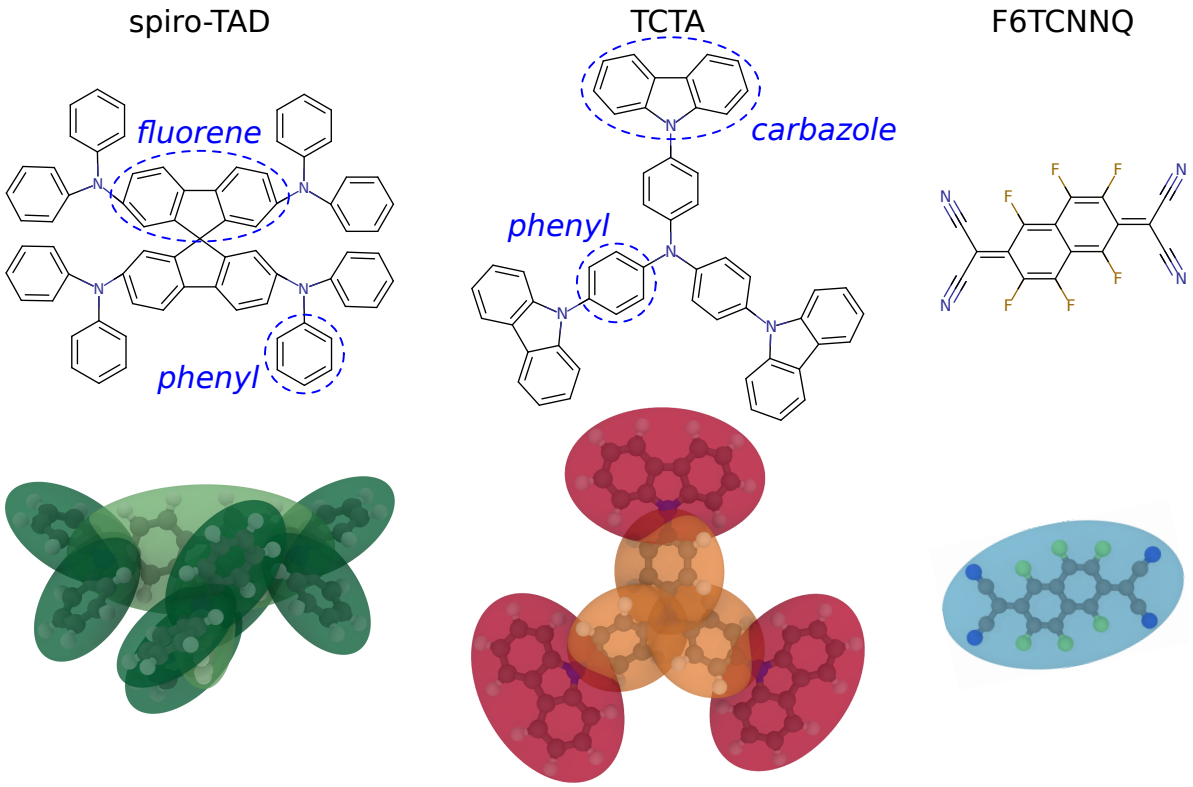


Figure 1: Atomic structure of the organic semiconductors spiro-TAD and TCTA, and the dopant F6TCNNQ. The bottom panel compares the CG structure (semi-transparent ellipsoidal beads) and the classic ball-and-stick representation.

the MOLC CG force field, followed by back-mapping to atomistic coordinates and short annealing at room temperature with an all-atom (AA) force field to redistribute the thermal energy on the new degrees of freedom. The resulting morphologies are then used to perform electronic-structure calculations to compute charge transport energy levels and intermolecular charge transfer integrals.

The results obtained with this protocol are compared with reference calculations, in which organic semiconductor samples were obtained from independent MD calculations with atomic detail. Our analysis demonstrates an excellent agreement between CG and AA results, both in terms of the structural features and the resulting electronic properties, providing a solid argument in support of accurate and cost-effective simulations with the MOLC model.

2 Methods

2.1 MD simulations

The organic semiconductors studied in this work are parameterized in a bottom-up fashion using the Advanced Topology Builder (ATB) repository, version 3.0,[48–50] to assign AA force field parameters based on GROMOS 54A7,[51,52] and atomic charges computed with the semiempirical molecular-orbital program MOPAC.[53,54] The MOLC CG force field is then parametrized using GROMOS-ATB as a reference for the inter- and intra-molecular interactions, following the protocol described in reference.[24]

The spiro-TAD and TCTA semiconductors have been modelled with 10 and 6 CG beads (biaxial Gay-Berne ellipsoids), respectively, while for the F6TCNNQ dopant we opted for a single bead representation. A pictorial view of the relationship between molecular fragments and CG beads is shown in Figure 1. Point charges suitably placed inside the ellipsoids were used to account for long-range electrostatic interactions: typically, the set of charges employed is smaller than the number of atoms the bead represents, reducing the computational cost associated with the full electrostatic treatment. The optimal number of CG charges is chosen as to simultaneously reproduce the molecular electrostatic potential, dipole and quadrupole moments of the associated molecular fragment through a genetic algorithm optimisation procedure.[24,36] The model for spiro-TAD includes 30 charges (3 per fluorene bead, 3 per phenyl bead), while TCTA has been described with 24 charges (4 per carbazole bead, 4 per phenyl bead). For F6TCNNQ, we tested two models with respectively 6 and 9 charges, finding that the latter, more detailed, model is necessary to properly describe the molecular electrostatic potential at contact distance and the molecular packing in the amorphous phase (see Supporting Information, SI, Figures S1-S2, Table S1). AA and CG simulations were carried out with the program LAMMPS,[14,55] version 30 October 2019, *in house* modified to include the MOLC user package.[56]

Each CG sample was prepared by placing 1000 randomly-oriented molecules in the sim-

ulation box, distributed on the nodes of a cubic grid, with 3D periodic boundary conditions applied. The sample was slowly compressed at 1 K to suppress thermal motions, yielding an amorphous solid. The compression step was carried out with a timestep of 10 fs, with the temperature controlled via a Langevin thermostat with a damping parameter of 10 ps. Each sample was then relaxed with isobaric-isothermal conditions (NPT ensemble) at temperature $T = 300$ K, pressure $P = 1$ atm, using a timestep of 20 fs, until thermal equilibrium was reached. A production trajectory of 5 ns was finally collected using a Nosé-Hoover barostat with a stress-damping parameter of 200 ps. The CG samples were back-mapped to their atomistic representation with the open-source program Backmap.[57] The back-mapped samples were then relaxed with the AA GROMOS-ATB force field at $T = 300$ K and $P = 1$ atm for 5 ns using a timestep of 2 fs, a Nosé-Hoover barostat with a damping parameter of 2 ps and a Langevin thermostat with a damping parameter of 0.2 ps. The last configuration was finally minimised at 0 K to remove thermal fluctuations and used to carry out electronic-structure calculations.

In order to have a purely-atomistic reference against which benchmarking the back-mapped morphologies, we carried out AA simulations using the same protocol described for CG samples, i.e., compressing low-density samples to the desired density and then equilibrating in the NPT ensemble. Along these lines, we obtained native atomistic amorphous morphologies with a completely independent history from the back-mapped CG samples.

Besides bulk morphologies, we have also created thin-film samples that have been used for the calculations of the energy levels to be compared with photoemission experiments. CG samples of pristine spiro-TAD, TCTA, and F6TCNNQ in a thin-film configuration were created by removing the periodic boundary along the z-direction and adding empty space, turning the bulk systems into a free-standing slab. Each system was equilibrated in the canonical ensemble (NVT) at 300 K for 500 ns in order to achieve a complete surface relaxation. The equilibration was carried out with a timestep of 20 fs and a correction term was added to compute the long-range electrostatic interactions for systems with slab geometry.[58]

After back-mapping to the atomic coordinates, the samples were further equilibrated for 10 ns with a timestep of 2 fs before carrying out the final energy minimisation.

2.2 Electronic properties

Following the common assumption of charge carriers localized on individual molecules and a perturbative treatment of inter-molecular interactions,[43] the nth energy level of molecule i can be written as

$$E_n^i = \varepsilon_n^i + \Delta_n^i \tag{1}$$

where ε_n^i is the molecular gas-phase energy and Δ_n^i is the environmental contribution. In the following, we will focus on the levels relevant to hole (n = HOMO) and electron (n = LUMO) transport. Please note that these levels are intended as electron addition (ionization potential, IP=$-E_{\mathrm{HOMO}}$) and removal (electron affinity, EA=$-E_{\mathrm{LUMO}}$) energies, and should not be confused with their analogue in mean-field theories (e.g. DFT). The relevant densities of states (DOS) are built as distributions of site energies obtained upon sampling over the molecules of our amorphous morphologies. The width of these distributions, measured by the standard deviation σ, quantifies the energetic disorder in our samples.

Gas-phase energy levels depend on the molecular geometries extracted from MD snapshots. This intra-molecular contribution is computed as

$$\begin{aligned} \varepsilon_n^i &= \varepsilon_n^{(GW)}(\tilde{\boldsymbol{r}}_0) + \Delta\varepsilon_n^{(SE)}(\tilde{\boldsymbol{r}}_i) \\ &= \varepsilon_n^{(GW)}(\tilde{\boldsymbol{r}}_0) + \left[\varepsilon_n^{(SE)}(\tilde{\boldsymbol{r}}_i) - \varepsilon_n^{(SE)}(\tilde{\boldsymbol{r}}_0)\right] \end{aligned} \tag{2}$$

where $\tilde{\boldsymbol{r}}_0$ indicates the molecular equilibrium geometry, $\tilde{\boldsymbol{r}}_i$ is the geometry of molecule i in the amorphous sample, and the superscripts indicate the level of theory employed in the calculations. This approach hence combines the many-body GW formalism,[59,60] with the semiempirical (SE) ZINDO method.[61] The GW calculations have been performed once at the equilibrium geometry (obtained at PBE0/cc-pVDZ level) for each compound, adopting

a partial self-consistency on the eigenvalues (evGW) and extrapolating the results in the complete basis set limit.[62] This permits to obtain absolute values for the energy levels with a typical accuracy of 0.1 eV,[63,64] as also demonstrated by a recent comparison with reference CCSD(T) data on medium-size molecules.[65,66] The gas-phase evGW energy levels for the three molecules under examination are reported in the SI, Table S2. ZINDO calculations are instead employed as a cost-effective mean to assess the variation of energy levels with molecular geometry in large samples, assuming that the intrinsic inaccuracy of ZINDO cancels out in the difference $\left[\varepsilon_n^{(SE)}(\tilde{\boldsymbol{r}}_i) - \varepsilon_n^{(SE)}(\tilde{\boldsymbol{r}}_0)\right]$ (see Eq. 2). Energy levels calculated according to Eq. 1 and Eq. 2 correspond to vertical IE and EA, without accounting for structural relaxation. The spread of gas-phase molecular levels, i.e the standard deviation σ_{conf}, quantifies the energetic disorder of conformational origin. Density functional theory (DFT) and semiempirical electronic-structure calculations have been performed with the ORCA software (version 4.2.1),[67,68] GW calculation results have been obtained with the Fiesta package.[63,64]

The intermolecular contribution is sourced from electrostatic interactions in the polarizable molecular environment, and it is here evaluated with classical microlectrostatic (ME) calculations, as implemented in the MESCal code.[69] The intermolecular interactions consist of an electrostatic and an induction contribution[43]

$$\Delta_n^i = \Delta_n^{E,i} + \Delta_n^{I,i} \tag{3}$$

The electrostatic term $\Delta_n^{E,i}$ corresponds to the energy necessary to charge a molecule in the field of permanent and induced multipoles of all other molecules in the neutral sample. Due to the inhomogeneity of the electrostatic landscape in disordered samples, this term represents the leading source of environmental energetic disorder, σ_{env}.[70] $\Delta_n^{E,i}$ has been evaluated with self-consistent ME calculations of atomistic resolutions, employing atomic charges from electrostatic potential fitting (ESP scheme)[71] and molecular polarizability tensor computed

at the PBE0/cc-PVTZ theory level. ME calculations on the neutral systems have been performed accounting for periodic boundary conditions in 3D, except those on slab samples reported in Section 3 that employed 2D periodicity to simulate photoemission spectra at the sample surface. The induction term $\Delta_n^{I,i}$ quantifies the reaction of the polarizable environment to an added charge. This term represents an important contribution to the absolute value of charge transport levels, but it is weakly sensitive to the sample morphology, hence it provides an almost negligible contribution to the energetic disorder. It has been evaluated at the molecular equilibrium geometries once and for all for each chemical species. In this case we performed ME calculations in an effective isotropic polarizable medium (cubic lattice of polarizable points) with dielectric constant arbitrarily set at $\epsilon_r = 3$, a typical value for organic molecules.

Intermolecular charge transfer couplings have been computed at the ZINDO level with the standard dimer projection method.[72]

3 Results and discussion

3.1 Structural properties

The multiscale protocol begins with the generation of amorphous samples of pristine spiro-TAD, TCTA, and F6TCNNQ, plus samples of spiro-TAD and TCTA doped with F6TCNNQ in 2% and 10% molar concentrations, similar to the ones used in experiments.[44] For the sake of illustration, Figure 2 shows two samples of F6TCNNQ-doped spiro-TAD and TCTA, displaying the morphology obtained from CG MD simulations and the atomistic one obtained after back-mapping and thermalization. We characterize the CG samples in terms of their density and radial distribution functions (RDFs), and we compare these observables with the corresponding AA samples.

The computed mass density values for the different samples are reported in Table 1. The densities obtained with the MOLC model are within 7% of reference AA data, similar to what

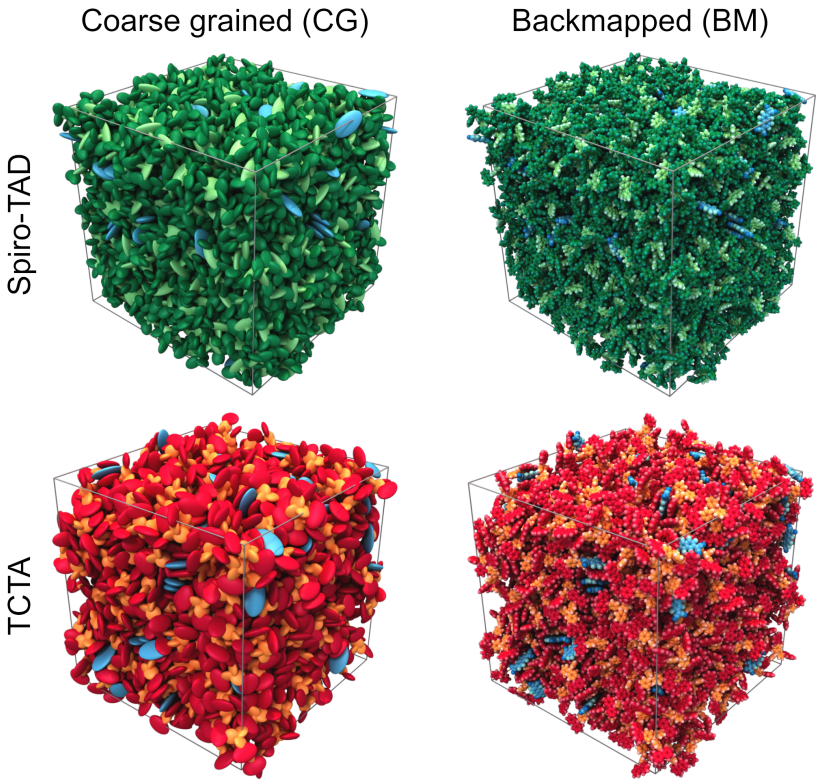


Figure 2: Rendering of spiro-TAD and TCTA samples doped with F6TCNNQ (10% molar concentration) as obtained from MD simulations with the CG model and after atomistic backmapping and relaxation. The different beads and fragments are visualized according to the colour scheme defined in Figure 1.

Table 1: Density (g/cm^3) of amorphous samples of organic semiconductors from coarse-grained (CG), back-mapped (BM), and reference atomic (AA) simulations.

Sample	**CG**	**BM**	**AA**
F6TCNNQ	1.216(2)	1.221(2)	1.231(2)
Spiro-TAD	1.131(1)	1.082(1)	1.091(1)
Spiro-TAD/F6TCNNQ: 2%	1.129(1)	1.120(1)	1.097(1)
Spiro-TAD/F6TCNNQ: 10%	1.124(1)	1.090(1)	1.094(1)
TCTA	1.083(1)	1.147(1)	1.162(1)
TCTA/F6TCNNQ: 2%	1.087(1)	1.148(1)	1.163(1)
TCTA/F6TCNNQ: 10%	1.083(1)	1.150(1)	1.164(1)

was reported for the CG model for α-NPD.[24] A much smaller, $\sim 1\%$, relative difference has been obtained between the GC and AA sample densities for pure F6TCNNQ. The significant, although not systematic, density deviations between the two models suggest that a reduction in the number of degrees of freedom comes at the expense of a quantitative description of the

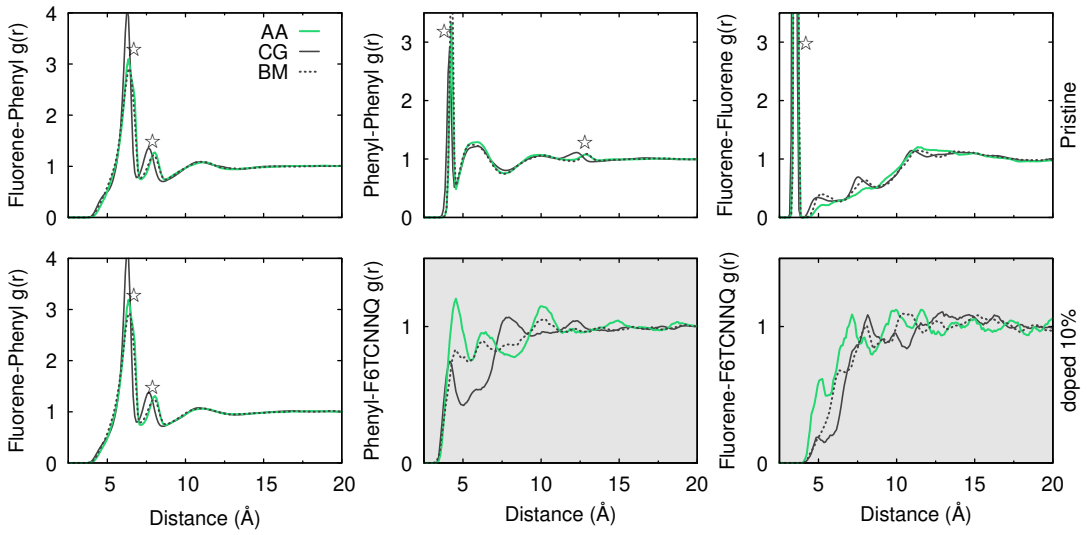


Figure 3: Radial distribution functions $g(r)$ computed between molecular fragments in the pristine and F6TCNNQ-doped (10% molar concentration) spiro-TAD. Molecular fragments are defined in Fig. 1. The gray-shaded panels refer to doped samples, and specifically to the coordination of the F6TCNNQ dopant with the host molecule. The different data sets correspond to all-atom (green solid line), coarse-grained (black solid line), and back-mapped atomic (black dotted line) models. The stars mark intra-molecular peaks of the host molecule.

stress in CG samples. Possible improvement strategies could include density and cohesive energy as additional targets in the model parameterization. The density of the BM samples is typically within $\sim$ 2% to that of the reference AA simulation, hinting at the fact that the molecular packing of their parent CG sample is an excellent initial guess and that short annealing at room temperature is enough to retrieve the correct density.

To further characterize the relationship between the CG and AA samples, we analyse the RDFs computed between the CG bead centroids, as well as the AA and CG RDFs, obtained with the center of mass of the molecular fragments as defined in Figure 1. The RDFs of pristine spiro-TAD (Figure 3) and TCTA (Figure 4), and F6TCNNQ (SI, Figure S2) show the telltale signature of an amorphous morphology, characterised by broad peaks followed by a flat profile, indicating a homogeneous distribution of molecules at long distance. The intra-molecular peaks are marked with a star, while the other peaks show the average distance between molecules in the first coordination shell. Overall, a good agreement is found between

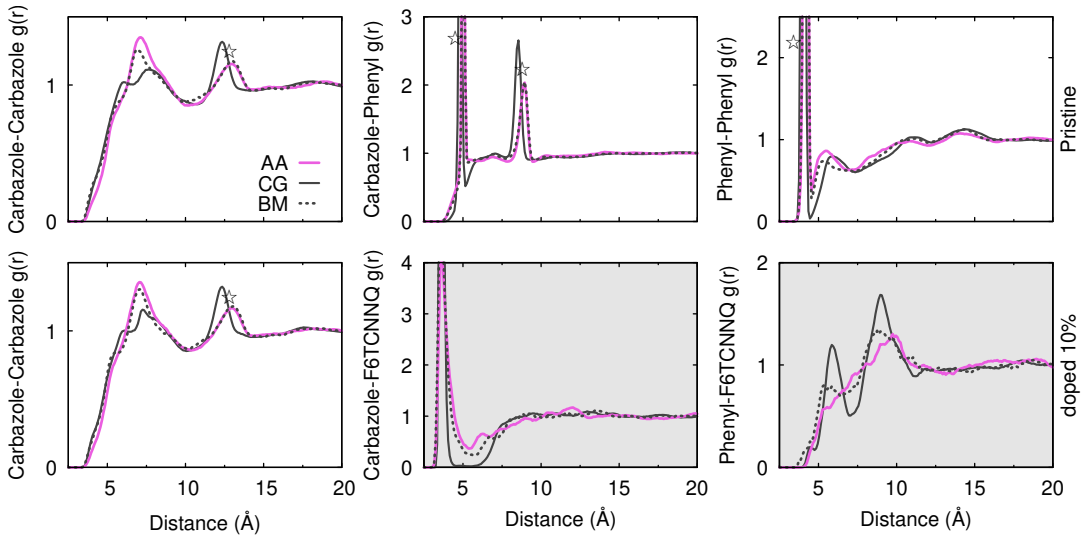


Figure 4: Radial distribution functions $g(r)$ computed between molecular fragments in the pristine and F6TCNNQ-doped (10% molar concentration) TCTA. Molecular fragments are defined in Fig. 1. The gray-shaded panels refer to doped samples, and specifically to the coordination of the F6TCNNQ dopant with the host molecule. The different data sets correspond to all-atom (magenta solid line), coarse-grained (black solid line), and back-mapped atomic (black dotted line) models. The stars mark intra-molecular peaks of the host molecule.

the RDF from CG and AA simulations. The peaks in the CG samples are generally sharper than their AA counterparts as the internal degrees of freedom are suppressed in the CG beads. This difference disappears in the RDF of BM samples, which are almost indistinguishable from the AA samples.

In doped samples, the RDF between the F6TCNNQ dopant and the host molecular fragments (gray-shaded panels) are in general noisy due to the low concentration of dopant molecules. For doped spiro-TAD samples (see Figure 3), the RDF between phenyl–F6TCNNQ and fluorene–F6TCNNQ show a shallow profile in all samples suggesting a lack of preferential orientation of dopant molecules and absence of a coordination shell. For doped TCTA samples, a sharp peak at 3.7 Å is observed for the carbazole–F6TCNNQ pair (bottom-central panel in Figure 4), a distance typical of π–π stacking. This neat feature is present in all samples, even though the peak appears sharper in the CG one. The RDF for the phenyl–F6TCNNQ fragments (bottom-right panel in Figure 4) is relatively shallow and close to

1, similar to what seen for the spiro-TAD samples. However, in all samples a small peak is present at 9-10 Å distance that is consistent with the distance of a dopant molecule coordinated with the carbazole fragment. The amorphous matrix of both spiro-TAD and TCTA is not appreciably influenced by the dopant presence, as the RDF fluorene–phenyl and carbazole–carbazole are almost identical to those of the pristine samples. The same results were found for the samples doped with 2% of F6TCNNQ, shown in SI Figures S3-S4.

The analysis of the position and intensity of the main peaks of the RDFs reveals substantial similarities between the morphology of the CG and AA samples obtained independently from one another. This result provides a strong indication that, despite significant differences in density, the CG and AA force fields yield consistent morphologies, albeit not yet attaining the accuracy level of CG model obtained, for instance, by iterative Boltzmann inversion[73] or force matching[74] techniques. The overall quality of MOLC CG morphologies is further attested by the fact that, upon atomistic back-mapping and a short re-equilibration run, the discrepancy in the RDF against the atomistic reference becomes practically negligible.

In the next section we make one step forward in our benchmark of CG morphologies, by comparing the electronic properties computed for the BM and AA samples.

3.2 Electronic structure properties

We start the analysis of the electronic properties from the energy levels of localized charge carriers in pristine and doped materials. The distribution of these quantities, mapping the density of states (DOS) for hole and electron carriers, are shown in Figure 5. The breakdown of the energy levels into intra- and inter-molecular contributions (see Eq. 2), as well as the associated energetic disorder, is reported in Table 2 for the HOMO (LUMO) levels of host (dopant) molecules – see SI Table S3 for complementary data. Our results for TCTA and spiro-TAD are in good agreement with those reported by Kotadiya *e*t al., obtained with a similar approach.[13]

We first observe a very good agreement between the set of DOS obtained with AA and BM

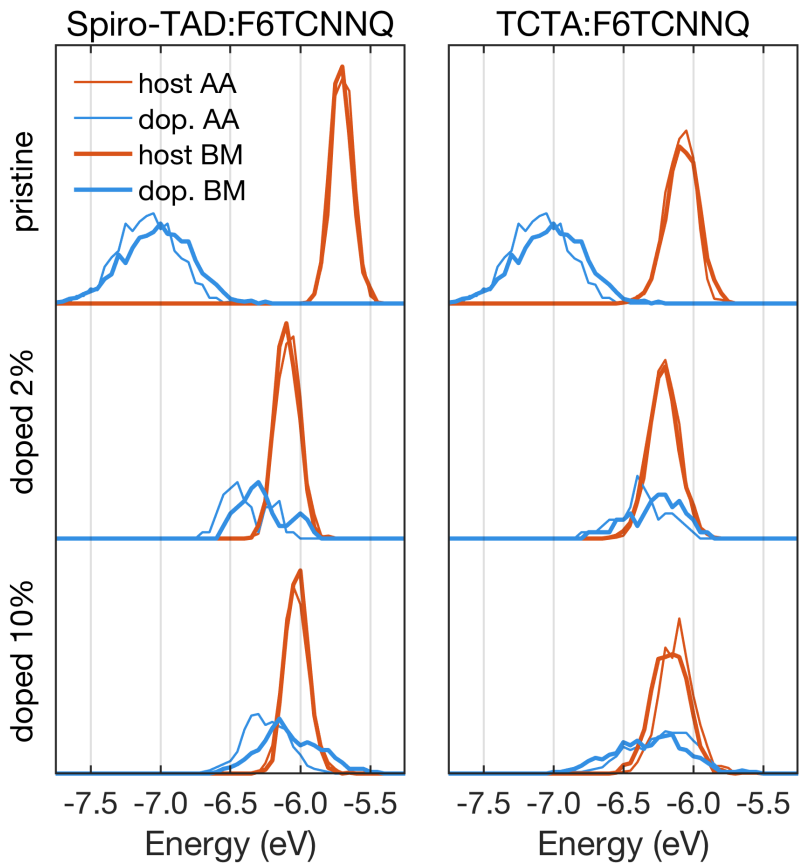


Figure 5: Densities of states for holes at host, spiro-TAD and TCTA, molecules (orange lines) and electrons at the F6TCNNQ dopant (light blue lines) for pristine and doped bulk systems. Results from QM and ME calculations (see Eq. 2) performed on the native atomistic (thin lines) and back-mapped (thick lines) morphologies. We note the very good agreement between the DOSs computed with AA and BM samples. The DOSs of F6TCNNQ LUMOs in doped samples are noisy because of the poor statistics, especially in the 2% concentration sample.

morphologies, both in terms of mean values and distribution spread. The main significant discrepancy between the two data sets regards the standard deviation of the LUMO levels of the pristine F6TCNNQ, which is 20% higher in the BM sample, with respect to the AA one. The data in Table 2 allow to safely ascribe this difference to the environmental contribution that is indeed sensitive to the details of molecular packing inherited from the CG simulation history.

Table 2: Charge transport energy levels (mean values, eV units) and energetic disorder (in meV) of pristine systems. Results from QM and ME calculations (see Eq. 2) calculations employing samples obtained with native atomistic simulations (AA) and the back-mapped MOLC model (BM). H and L label HOMO and LUMO levels, respectively. Energy levels account for a geometry-independent induction term Δ^I, equal to -1.00 eV for electrons in F6TCNNQ and 0.54 and 0.61 eV for holes in spiro-TAD and TCTA, respectively.

material	sample	level	$\langle E \rangle$	$\langle \varepsilon \rangle$	$\langle \Delta^E \rangle$	σ	σ_{conf}	σ_{env}
F6TCNNQ	BM	L	-7.03	-4.55	-1.47	249	14	251
F6TCNNQ	AA	L	-7.09	-4.55	-1.54	206	14	207
Spiro-TAD	BM	H	-5.71	-6.48	0.23	77	33	69
Spiro-TAD	AA	H	-5.70	-6.48	0.24	76	36	67
TCTA	BM	H	-6.08	-6.73	0.04	117	44	105
TCTA	AA	H	-6.09	-6.73	0.03	104	45	91

The DOSs in Figure 5 display a marked dependence of the energy levels on the chemical composition of the system. The most striking variation regards the 0.7-0.8 eV decrease of the F6TCNNQ LUMO level between its pristine phase and when it acts as a dopant impurity in the hole transporting materials spiro-TAD and TCTA. Such a pronounced host-dependence of the acceptor level of dopant impurities has been disclosed previously on the basis of accurate embedded *GW* calculations.[62] We also note a smaller variation of the host HOMOs upon doping, with the mean values of their distributions shifting to higher binding energies by about 0.4 and 0.2 eV in spiro-TAD and TCTA, respectively. This effect can be attributed to the short-range interactions between hole carriers and the quadrupole moment of the dopant molecules, as discussed in a recent study.[75]

We complement the characterization of molecular site energy levels by discussing the corresponding spatial correlation function

$$C(r) = \frac{1}{\sigma^2} \langle \, (E(r_i) - \langle E(r_i) \rangle) \, (E(r_i + r) - \langle E(r_i) \rangle) \, \rangle \tag{4}$$

where $E(r_i)$ is the energy level of molecule i, σ^2 is the variance of site energy fluctuations, and angle brackets indicate the average over the molecules in the sample. $C(r)$ decays from 1 (full correlation) at $r = 0$ to 0 (no correlation) a large distance for our pure samples shown in

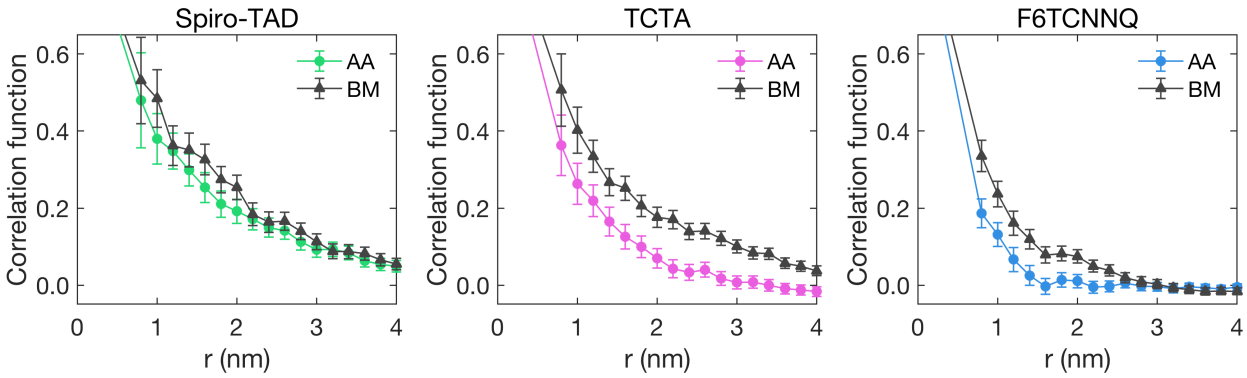


Figure 6: Spatial correlation function of electrostatic site energies in bulk pristine samples.

Figure 6, similar to what reported for other amorphous organic semiconductors,[8,76] and approximately following the $1/r$ behavior expected for randomly oriented dipoles on a lattice.[77] The comparison between AA and BM morphologies reveals qualitative similarities between the two approaches, albeit the MOLC model seems to produce longer-range correlations that are inherited in the BM samples. Quantitatively, significant differences exist between TCTA and F6TCNNQ, with the latter showing a larger relative difference between AA and BM.

While charge carries energy levels mostly reflect the impact of electrostatic interactions developing at long range, intermolecular charge transfer couplings J, being related to the quantum overlap between molecular wavefunctions, are instead ideal probes for nearest-neighbor molecular contacts. The distributions of intermolecular couplings, shown in Figure 7, reveal an excellent agreement between AA and BM morphologies for the three pristine samples. The distributions rapidly drops upon increasing J, with a very small fraction of dimers exceeding 10 meV. These are typical feature of amorphous molecular solids.[8,76] The region below 0.1 meV is loosely relevant as it is highly sensitive to the criteria employed in the selection of molecular pairs to be computed, hence not shown.

The information contained in charge transfer integrals is then exploited to quantify the effective electronic connectivity of the different samples. Following Ref.,[76] we define the electronic connectivity as the average number of neighbors to which every molecule is coupled by a transfer integral exceeding a given threshold. This is shown for pure TCTA, spiro-TAD and

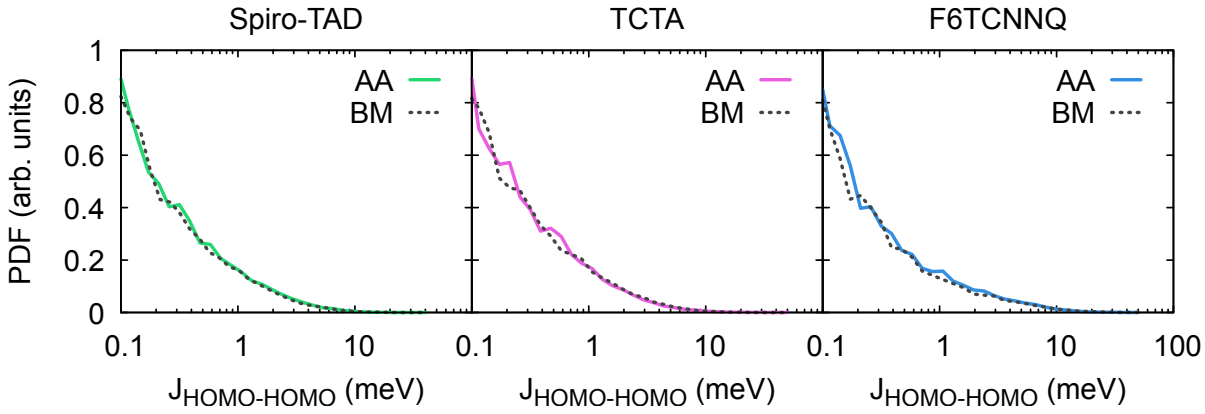


Figure 7: Probability distribution function (PDF) of intermolecular charge transfer couplings (absolute values) between frontier molecular orbitals (see axis label) for pristine spiro-TAD, TCTA, and F6TCNNQ samples.

F6TCNNQ samples in Figure 8, as a function of the threshold coupling J_{thr}. The comparison between the connectivity curves of the different materials reveals that TCTA presents a higher connectivity than spiro-TAD, an observation that can be related to the denser molecular packing of the former, less rigid, molecule (see Table 1). F6TCNNQ presents a higher connectivity than the two hosts for $J_{\mathrm{thr}} > 2$ meV, while the situation is reversed for smaller J_{thr}. This can be ascribed to the smaller size of the dopant molecules and to the consequently more localized frontier orbitals, which can give rise to a small number of strong intermolecular coupling. On the other hand, branched molecules as TCTA and spiro-TAD can instead establish a larger number of weaker contacts with other molecules. As for it concerns the comparison between AA and BM morphologies for a given material, we remark also in this case an excellent agreement, with very similar electronic connectivities obtained with the two molecular models, with the only appreciable deviation obtained for F6TCNNQ at low couplings ($J_{\mathrm{thr}} < 2$ meV) .

The results reported in this section overall prove that the electronic properties calculated on atomistic structures, obtained upon back-mapping CG morphologies simulated with the MOLC model (BM), are almost indistinguishable from those obtained from native atomistic samples (AA). The comparison among different materials show that the loss of detail intro-

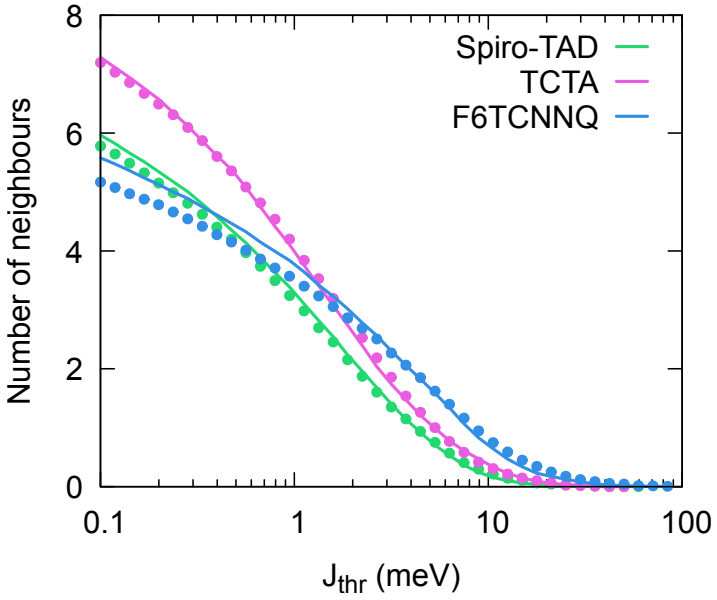


Figure 8: Intermolecular electronic connectivity for the three pristine samples, defined as the average number of neighbors having a charge transfer coupling larger in magnitude than a given threshold. Solid lines and circles correspond to AA and BM morphologies, respectively.

duced upon coarse graining has a larger impact on the electronic properties of F6TCNNQ, described by a single bead. Here, the most significant flaw of BM morphologies is a 20% overestimation of the energetic disorder as compared to AA simulations (see Table 2). The discrepancy between the electronic properties of BM and AA amorphous phases in instead negligible for complex multi-bead molecules such as TCTA and spiro-TAD, i.e. the systems where the MOLC model is mostly advantageous.

3.3 Photoemission energy levels

We finally attempt a comparison between our multiscale simulations for charge transport energy levels with available experimental data from photoemission spectroscopies.[44,46,78–81] To such an aim, we performed a dedicated set of simulations for the organic-vacuum interface, since both direct ultraviolet photoelectron spectroscopy (UPS) and inverse photoemission spectroscopy (IPES) are surface-sensitive techniques, probing only the first 1-2 nanometers of the sample. Such an approach turned out to be essential to obtain a quantitative description

of energy levels in molecular crystals and ordered films.[43,82–85] These slab simulations have been performed with the MOLC model, with subsequent back mapping to atomistics to run electronic the structure calculations. The simulated photemission spectra have been calculated as histograms of molecular site energies weighted by a position-dependent factor $\exp(-z/\lambda)$, where z is the distance of a given molecule from the surface, and $\lambda = 1$ nm is the electron mean free path.

Figure 9 shows the calculated photoemission spectra for occupied (holes) and unoccupied (electrons) states, along with experimental values (vertical bars) for ionization potential (IP) and electron affinity (EA), both taken at the respective band onset. We remark a good agreement between simulations and experiments as for it concerns occupied states, with calculations capturing the trends between different materials. It is worth remarking the significant difference between the IP of F6TCNNQ that would be obtained from bulk simulations (6.5 eV, see Figure 5) and the surface value of 7.7 eV. This substantial difference, which is key for reproducing the experiment, is determined by the electrostatic environmental contribution, $\Delta_E = -0.26$ eV at the surface vs. -1.47 eV in the bulk (see Table 2). This large difference in the electrostatic energy landscape results from a large surface reconstruction at the interface to vacuum, which can be appreciated from the visual inspection of the surface as obtained upon cutting the bulk sample, and after thermal annealing, as shown in Figure S5. The difference in Δ_E between bulk and surface are instead quite modest (< 0.2 eV) for TCTA and spiro-TAD, consistent with a much less important surface reconstruction.

Our calculations do not achieve comparable accuracy for unoccupied states, namely overestimating EA by ~ 1 eV for F6TCNNQ and underestimating EA of TCTA and spiro-TAD by ~ 1.2 eV, as compared to IPES data by Zhang and Kahn.[44] A much better agreement is obtained with the EA of TCTA reported by Yoshida and Yoshizaki[46] and obtained with low-energy IPES (LEIPS), a technique that is considered being more accurate than conventional IPES.[86] In general, the comparison between experiments and calculation for photoemission measurements at amorphous films surfaces reveals a satisfactory agreement on occupied en-

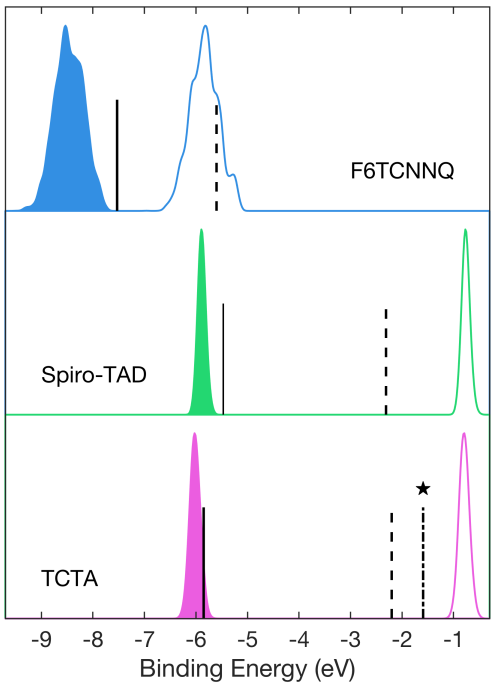


Figure 9: Simulated photoemission spectra of pristine F6TCNNQ, spiro-TAD and TCTA samples. Filled and empty areas refer to occupied and unoccupied states, respectively. Vertical black lines mark to the IP (valence band top, full line) and the EA (conduction band minimum, dashed line) measured in Ref.[44] The bar labeled with the star marks the EA of TCTA reported in Ref.[46] Other works have reported consistent data for the IP of TCTA[78,79] and spiro-TAD.[80,81]

ergy levels, calling at the same time for an improved in the description of unoccupied states.

4 Conclusions

We have presented a comparative analysis of morphologies obtained from atomistic and coarse-grained molecular dynamics, the latter based on the MOLC model.[24] This study covered both the structure of pristine and doped organic semiconductors and the electronic properties governing charge transport in these systems, implementing a multiscale simulation workflow that includes molecular dynamics, atomistic back-mapping for coarse-grained morphologies, classical microelectrostatics, DFT and many-body *GW* electronic structure calculations.

Our study reveals a substantial agreement between the morphologies simulated with the MOLC CG model and reference atomistic ones. Most interestingly, the reintroduction of the atomistic detail in CG samples (back-mapping) leads to structures that are almost indistinguishable, in a statistical sense, from those obtained from native atomistic simulations. The same conclusion holds for electronic properties that are known to be highly sensitive to molecular geometries, meaning that atomistic samples obtained upon back-mapping can be used with a high level of confidence for more refined analyses, especially for branched poly-heterocyclic compounds the MOLC model is designed for. Besides the methodological benchmark, we have attempted a comparison of calculated charge transport energy levels with literature data from photoemission spectroscopy, obtaining a very good accord on occupied levels. The agreement with experiments remains qualitative for unoccupied states, calling for additional research effort.

In summary, the MOLC coarse-grained molecular model emerges as very accurate approach for the simulation of complex molecular systems. This enables an important saving of computational resources (two orders of magnitude reduction in the time per simulated nanosecond) and, most interestingly, it paves the way towards the simulation of systems and phenomena whose length and time scales are out of the reach of atomistic techniques, such as those relative to the processing of organic semiconductors over large area, among others.

Acknowledgement

This work has received early-stage support from the European Union's Horizon 2020 Framework Programme under grant agreement No. 646176 (EXTMOS). Work in Grenoble has been supported by the French "Agence Nationale de la Recherche", project RAPTORS (ANR-21-CE24-0004-01). High-performance computing resources from GENCI-TGCC (Grant no. 2020-A0090910016) are acknowledged.

Supporting Information Available

The following files are available free of charge.

- `supp_info.pdf`: additional structural and electronic characterisation.
- `osc_multiscale.zip`: template files for multiscale simulations of Spiro-TAD, TCTA, and F6TCNNQ.

References

(1) Riniker, S. Fixed-charge atomistic force fields for molecular dynamics simulations in the condensed phase: An overview. *Journal of Chemical Information and Modeling* **2018**, *58*, 565–578.

(2) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Amer. Chem. Soc.* **1996**, *118*, 11225–11236.

(3) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(4) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*, 2nd ed.; Oxford University Press: Oxford, 2017.

(5) Tiberio, G.; Muccioli, L.; Berardi, R.; Zannoni, C. Towards in silico liquid crystals. Realistic transition temperatures and physical properties for *n*–cyanobiphenyls via molecular dynamics simulations. *ChemPhysChem* **2009**, *10*, 125–136.

(6) Zannoni, C. *Liquid Crystals and their Computer Simulations*; Cambridge University Press: Cambridge, 2022.

(7) Coropceanu, V.; Cornil, J.; da Silva Filho, D. A.; Olivier, Y.; Silbey, R.; Brédas, J. L. Charge transport in organic semiconductors. *Chem. Rev.* **2007**, *107*, 926–952.

(8) Ruhle, V.; Lukyanov, A.; May, F.; Schrader, M.; Vehoff, T.; Kirkpatrick, J.; Baumeier, B.; Andrienko, D. Microscopic simulations of charge transport in disordered organic semiconductors. *J. Chem. Theory Comput.* **2011**, *7*, 3335–3345.

(9) Brédas, J.-L., Marder, S. R., Eds. *The WSPC Reference on Organic Electronics: Organic Semiconductors*; World Scientific: Singapore, 2016.

(10) Friederich, P.; Fediai, A.; Kaiser, S.; Konrad, M.; Jung, N.; Wenzel, W. Toward design of novel materials for organic electronics. *Advanced Materials* **2019**, *31*, 1808256.

(11) Baggioli, A.; Casalegno, M.; Raos, G.; Muccioli, L.; Orlandi, S.; Zannoni, C. Atomistic simulation of phase transitions and charge mobility for the organic semiconductor Ph-BTBT-C10. *Chemistry of Materials* **2019**, *31*, 7092–7103.

(12) Li, H. Y.; Brédas, J. L. Developing molecular-level models for organic field-effect transistors. *Natl. Sci. Rev.* **2021**, *8*.

(13) Kotadiya, N. B.; Mondal, A.; Xiong, S.; Blom, P. W. M.; Andrienko, D.; Wetzelaer, G.-J. A. H. Rigorous characterization and predictive modeling of hole transport in amorphous organic semiconductors. *Adv. Electron. Mater.* **2018**, *4*, 1800366.

(14) LAMMPS Molecular Dynamics Simulator. https://www.lammps.org.

(15) Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B.; the GROMACS development team, GROMACS user manual version 2019. http://www.gromacs.org, 2019.

(16) Phillips, J. C.; Braun, R.; Wang, W.; J.Gumbart,; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(17) Mityashin, A.; Cornil, J.; Zannoni, C.; Heremans, P. New predictive modeling approach to boost the development of organic electronics. *EE-Times Europe* **2013**,

(18) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.

(19) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI coarse-grained force field: Extension to proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.

(20) Shen, H.; Li, Y.; Ren, P.; Zhang, D.; Li, G. An anisotropic coarse-grained model for proteins based on Gay-Berne and electric multipole potentials. *J. Chem. Theory Comput.* **2014**, *10*, 731–750.

(21) Brandner, A.; Schüller, A.; Melo, F.; Pantano, S. Exploring DNA dynamics within oligonucleosomes with coarse-grained simulations: SIRAH force field extension for protein-DNA complexes. *Biochem. Biophys. Res. Commun.* **2018**, *498*, 319–326.

(22) Alessandri, R.; Uusitalo, J. J.; de Vries, A. H.; Havenith, R. W. A.; Marrink, S. J. Bulk heterojunction morphologies with atomistic resolution from coarse-grain solvent evaporation simulations. *J. Am. Chem. Soc.* **2017**, *139*, 3697–3705.

(23) Huang, D. M.; Faller, R.; Do, K.; Moulé, A. J. Coarse-grained computer simulations of polymer/Ffllerene bulk heterojunctions for organic photovoltaic applications. *Journal of Chemical Theory and Computation* **2010**, *6*, 526–537.

(24) Ricci, M.; Roscioni, O. M.; Querciagrossa, L.; Zannoni, C. MOLC. A reversible coarse grained approach using anisotropic beads for the modelling of organic functional materials. *Phys. Chem. Chem. Phys.* **2019**, *21*, 26195–26211.

(25) Sun, Q.; Faller, R. Systematic coarse-graining of atomistic models for simulation of polymeric systems. *Comput. Chem. Eng.* **2005**, *29*, 2380–2385.

(26) Gartner, T. E. I.; Jayaraman, A. Modeling and simulations of polymers: A roadmap. *Macromolecules* **2019**, *52*, 755–786.

(27) Cohen, A. E.; Jackson, N. E.; de Pablo, J. J. Anisotropic coarse-grained model for conjugated polymers: investigations into solution morphologies. *Macromolecules* **2021**,

(28) Kempfer, K.; Devémy, J.; Dequidt, A.; Couty, M.; Malfreyt, P. Development of coarse-grained models for polymers by trajectory matching. *ACS Omega* **2019**, *4*, 5955–5967.

(29) Peter, C.; Delle Site, L.; Kremer, K. Classical simulations from the atomistic to the mesoscale and back: coarse graining an azobenzene liquid crystal. *Soft Matter* **2008**, *4*, 859–869.

(30) Peter, C.; Kremer, K. Multiscale simulation of soft matter systems - from the atomistic to the coarse-grained level and back. *Soft Matter* **2009**, *5*, 4357–4366.

(31) Jackson, N. E.; Bowen, A. S.; Antony, L. W.; Webb, M. A.; Vishwanath, V.; de Pablo, J. J. Electronic structure at coarse-grained resolutions from supervised machine learning. *Sci. Adv.* **2019**, *5*.

(32) Jackson, N. E.; Bowen, A. S.; de Pablo, J. J. Efficient Multiscale Optoelectronic Prediction for Conjugated Polymers. *Macromolecules* **2020**, *53*, 482–490.

(33) Simine, L.; Allen, T. C.; Rossky, P. J. Predicting optical spectra for optoelectronic polymers using coarse-grained models and recurrent neural networks. *Proceedings of the National Academy of Sciences* **2020**, *117*, 13945–13948.

(34) Berardi, R.; Fava, C.; Zannoni, C. A generalized Gay–Berne intermolecular potential for biaxial particles. *Chem. Phys. Lett.* **1995**, *236*, 462–468.

(35) Berardi, R.; Fava, C.; Zannoni, C. A Gay–Berne potential for dissimilar biaxial particles. *Chem. Phys. Lett.* **1998**, *297*, 8–14.

(36) Berardi, R.; Muccioli, L.; Orlandi, S.; Ricci, M.; Zannoni, C. Mimicking electrostatic interactions with a set of effective charges: a genetic algorithm. *Chem. Phys. Lett.* **2004**, *389*, 373–378.

(37) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.

(38) Ruhle, V.; Junghans, C.; Lukyanov, A.; Kremer, K.; Andrienko, D. Versatile object-oriented toolkit for coarse-graining applications. *J. Chem. Theory Comput.* **2009**, *5*, 3211–3223.

(39) Bellussi, F. M.; Roscioni, O. M.; Ricci, M.; Fasano, M. Anisotropic electrostatic interactions in coarse-grained water models to enhance the accuracy and speed-up factor of mesoscopic simulations. *J. Phys. Chem. B* **2021**, *125*, 12020–12027.

(40) Jackson, N. E. Coarse-Graining organic semiconductors: the path to multiscale design. *J. Phys. Chem. B* **2021**, *125*, 485–496.

(41) Khot, A.; Savoie, B. M. Top–Down Coarse-Grained Framework for Characterizing Mixed Conducting Polymers. *Macromolecules* **2021**, *54*, 4889–4901.

(42) Stieffenhofer, M.; Scherer, C.; May, F.; Bereau, T.; Andrienko, D. Benchmarking coarse-grained models of organic semiconductors via deep backmapping. *Frontiers in Chemistry* **2022**, *10*.

(43) D'Avino, G.; Muccioli, L.; Castet, F.; Poelking, C.; Andrienko, D.; Soos, Z. G.; Cornil, J.; Beljonne, D. Electrostatic phenomena in organic semiconductors: fundamentals and implications for photovoltaics. *J. Phys.: Condens. Matter* **2016**, *28*, 433002.

(44) Zhang, F.; Kahn, A. Investigation of the high electron affinity molecular dopant F6-TCNNQ for hole-transport materials. *Adv. Funct. Mater.* **2018**, *28*, 1703780.

(45) Schwarze, M.; Gaul, C.; Scholz, R.; Bussolotti, F.; Hofacker, A.; Schellhammer, K. S.; Nell, B.; Naab, B. D.; Bao, Z. N.; Spoltore, D.; Vandewal, K.; Widmer, J.; Kera, S.; Ueno, N.; Ortmann, F.; Leo, K. Molecular parameters responsible for thermally activated transport in doped organic semiconductors. *Nat. Mater.* **2019**, *18*, 242–248.

(46) Yoshida, H.; Yoshizaki, K. Electron affinities of organic materials used for organic light-emitting diodes: A low-energy inverse photoemission study. *Org. Electron.* **2015**, *20*, 24–30.

(47) Koech, P. K.; Padmaperuma, A. B.; Wang, L.; Swensen, J. S.; Polikarpov, E.; Darsell, J. T.; Rainbolt, J. E.; Gaspar, D. J. Synthesis and Application of 1,3,4,5,7,8-Hexafluorotetracyanonaphthoquinodimethane (F6-TNAP): A Conductivity Dopant for Organic Light-Emitting Devices. *Chemistry of Materials* **2010**, *22*, 3926–3932.

(48) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An automated force field topology builder (ATB) and repository: Version 1.0. *J. Chem. Theory Comput.* **2011**, *7*, 4026–4037.

(49) Canzar, S.; El-Kebir, M.; Pool, R.; Elbassioni, K.; Malde, A. K.; Mark, A. E.; Geerke, D. P.; Stougie, L.; Klau, G. W. Charge group partitioning in biomolecular simulation. *J. Comput. Biol.* **2013**, *20*, 188–198.

(50) Koziara, K. B.; Stroet, M.; Malde, A. K.; Mark, A. E. Testing and validation of the Automated Topology Builder (ATB) version 2.0: prediction of hydration free enthalpies. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 221–233.

(51) Poger, D.; Van Gunsteren, W. F.; Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **2010**, *31*, 1117–1125.

(52) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843–856.

(53) Chirgwin, B. H.; Coulson, C. A.; Randall, J. T. The electronic structure of conjugated systems. VI. *Proc. R. Soc. London, Ser. A* **1950**, *201*, 196–209.

(54) Stewart, J. J. P. MOPAC: A semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–103.

(55) Plimpton, S. Fast parallel algorithms for short-range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(56) USER-MOLC package is available on GitHub. https://github.com/matteoeghirotta/lammps-30Oct19.

(57) Backmap: an open-source program to reverse map coarse-grained trajectories to atomic coordinates. https://github.com/matteoeghirotta/backmap_legacy.

(58) Yeh, I.-C.; Berkowitz, M. L. Ewald summation for systems with slab geometry. *J. Chem. Phys.* **1999**, *111*, 3155–3162.

(59) Hedin, L. New Method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **1965**, *139*, A796.

(60) Onida, G.; Reining, L.; Rubio, A. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.* **2002**, *74*, 601–659.

(61) Ridley, J.; Zerner, M. An intermediate neglect of differential overlap technique for spectroscopy: Pyrrole and the azines. *Theor. Chim. Acta* **1973**, *32*, 111–134.

(62) Li, J.; Duchemin, I.; Roscioni, O. M.; Friederich, P.; Anderson, M.; Da Como, E.; Kociok-Köhn, G.; Wenzel, W.; Zannoni, C.; Beljonne, D.; Blase, X.; D'Avino, G. Host dependence of the electron affinity of molecular dopants. *Mater. Horiz.* **2019**, *6*, 107–114.

(63) Blase, X.; Attaccalite, C.; Olevano, V. First-principles *GW* calculations for fullerenes, porphyrins, phtalocyanine, and other molecules of interest for organic photovoltaic applications. *Phys. Rev. B* **2011**, *83*, 115103.

(64) Faber, C.; Boulanger, P.; Attaccalite, C.; Duchemin, I.; Blase, X. Excited states properties of organic molecules: from density functional theory to the GW and Bethe-Salpeter Green's function formalisms. *Phil. Trans. R. Soc. A* **2014**, *372*, 20130271.

(65) Knight, J. W.; Wang, X.; Gallandi, L.; Dolgounitcheva, O.; Ren, X.; Ortiz, J. V.; Rinke, P.; Körzdörfer, T.; Marom, N. Accurate Ionization Potentials and Electron Affinities of Acceptor Molecules III: A Benchmark of GW Methods. *Journal of Chemical Theory and Computation* **2016**, *12*, 615–626.

(66) Duchemin, I.; Blase, X. Robust Analytic-Continuation Approach to Many-Body GW Calculations. *Journal of Chemical Theory and Computation* **2020**, *16*, 1742–1756.

(67) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.

(68) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

(69) D'Avino, G.; Muccioli, L.; Zannoni, C.; Beljonne, D.; Soos, Z. G. Electronic polarization in organic crystals: a comparative study of induced dipoles and intramolecular charge redistribution schemes. *J. Chem. Theory Comput.* **2014**, *10*, 4959–4971.

(70) Londi, G.; Khan, S.-U.-Z.; Muccioli, L.; D'Avino, G.; Rand, B. P.; Beljonne, D. Fate of low-lying charge-transfer excited states in a donor:acceptor blend with a large energy offset. *J. Phys. Chem. Lett.* **2020**, *11*, 10219–10226.

(71) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic charges derived from semiempirical methods. *J. Comput. Chem.* **1990**, *11*, 431–439.

(72) Valeev, E. F.; Coropceanu, V.; da Silva Filho, D. A.; Salman, S.; Brédas, J.-L. Effect of electronic polarization on charge-transport parameters in molecular organic semiconductors. *J. Am. Chem. Soc.* **2006**, *128*, 9882–9886.

(73) Reith, D.; Pütz, M.; Müller-Plathe, F. Deriving effective mesoscale potentials from atomistic simulations. *Journal of Computational Chemistry* **2003**, *24*, 1624–1636.

(74) Izvekov, S.; Voth, G. A. A Multiscale Coarse-Graining Method for Biomolecular Systems. *The Journal of Physical Chemistry B* **2005**, *109*, 2469–2473.

(75) Privitera, A.; Londi, G.; Riede, M.; D'Avino, G.; Beljonne, D. Molecular quadrupole moments promote ground-state charge generation in doped organic semiconductors. *Advanced Functional Materials* **2020**, *30*, 2004600.

(76) D'Avino, G.; Olivier, Y.; Muccioli, L.; Beljonne, D. Do charges delocalize over multiple molecules in fullerene derivatives? *J. Mater. Chem. C* **2016**, *4*, 3747–3756.

(77) Novikov, S. V.; Vannikov, A. V. Cluster structure in the distribution of the electrostatic potential in a lattice of randomly oriented dipoles. *J. Phys. Chem.* **1995**, *99*, 14573–14576.

(78) D'Andrade, B. W.; Datta, S.; Forrest, S. R.; Djurovich, P.; Polikarpov, E.; Thompson, M. E. Relationship between the ionization and oxidation potentials of molecular organic semiconductors. *Org. Electron.* **2005**, *6*, 11–20.

(79) Yun, D.-J.; Chung, J.; Kim, Y.; Park, S.-H.; Kim, S.-H.; Heo, S. Direct comparative study on the energy level alignments in unoccupied/occupied states of organic semiconductor/electrode interface by constructing in-situ photoemission spectroscopy and Ar gas cluster ion beam sputtering integrated analysis system. *J. Appl. Phys.* **2014**, *116*, 153702.

(80) Mahns, B.; Roth, F.; Grobosch, M.; Lindner, S.; Knupfer, M.; Saragi, T. P. I.; Reichert, T.; Salbeck, J.; Hahn, T. Electronic properties of spiro compounds for organic electronics. *J. Chem. Phys.* **2012**, *136*, 124702.

(81) Belisle, R. A.; Jain, P.; Prasanna, R.; Leijtens, T.; McGehee, M. D. Minimal effect of the hole-transport material ionization potential on the open-circuit voltage of perovskite solar cells. *ACS Energy Lett.* **2016**, *1*, 556–560.

(82) Topham, B. J.; Soos, Z. G. Ionization in organic thin films: Electrostatic potential, electronic polarization, and dopants in pentacene films. *Phys. Rev. B* **2011**, *84*, 165405.

(83) Li, J.; D'Avino, G.; Duchemin, I.; Beljonne, D.; Blase, X. Accurate description of charged excitations in molecular solids from embedded many-body perturbation theory. *Phys. Rev. B* **2018**, *97*, 035108.

(84) Poelking, C.; Tietze, M.; Elschner, C.; Olthof, S.; Hertel, D.; Baumeier, B.; Würthner, F.; Meerholz, K.; Leo, K.; Andrienko, D. Impact of mesoscale order on open-circuit voltage in organic solar cells. *Nat. Mater.* **2015**, *14*, 434–439.

(85) D'Avino, G.; Duhm, S.; Della Valle, R. G.; Heimel, G.; Oehzelt, M.; Kera, S.; Ueno, N.; Beljonne, D.; Salzmann, I. Electrostatic interactions shape molecular organization and electronic structure of organic semiconductor blends. *Chem. Mater.* **2020**, *32*, 1261–1271.

(86) Yoshida, H. Near-ultraviolet inverse photoemission spectroscopy using ultra-low energy electrons. *Chem. Phys. Lett.* **2012**, *539-540*, 180–185.