Supporting Information for
Mesoscale Molecular Network Formation in Amorphous Organic Materials

Authors: Brett M. Savoie[1*], Kevin L. Kohlstedt[1], Nicholas E. Jackson[1], Lin X. Chen[1,2], Monica O. de la Cruz[1], George C. Schatz[1], Tobin J. Marks[1*], Mark A. Ratner[1*]

Affiliations:

[1]Department of Chemistry, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, United States.

[2]Chemical Sciences and Engineering Division, Argonne National Laboratory, Lemont, Illinois, 60439 United States.

*Correspondence to: brettsavoie@u.northwestern.edu, ratner@northwestern.edu, t-marks@northwestern.edu

Additional Methodological Details

Networking Algorithm. $N(\mathbf{X};\mathbf{R})$ operates on a matrix **X** by searching the row of a seed vertex, **S**, for entries that satisfy the list of conditions in **R**. This generates a list of columns, **C**. Each column within **C** is then searched for elements that satisfy the connectivity conditions, generating a new set(s) of seed vertices, **S'**. At each step, the list of vertices already incorporated into a network, $\mathbf{V_N}$, is updated to avoid looping conditions. The algorithm continues until a null set is returned, at which point a new network is seeded using a vertex from the unincorperated set of vertices, $\mathbf{V}\text{-}\mathbf{V_N}$. Once all vertices have been assigned to a network (i.e. $\mathbf{V}\text{-}\mathbf{V_N}$ is null) the algorithm terminates. Note that subnetworks returned by $N(\mathbf{X};\mathbf{R})$ are uniquely defined by the conditions **R** and independent of the choice of **S**.

Molecular Dynamics. All trajectories were initialized as diffuse orthogonal arrays of 64 molecules, with randomized orientations to avoid unintentional ordering effects. Five independent trajectories were simulated for each material. All simulations were performed on isothermal-isobaric ensembles (constant nPT), using a Nose-Hoover barostat and thermostat and periodic boundary conditions. One fs timestep integrations were used along with a particle-particle-particle-mesh (pppm) solver and a 2 nm cutoff for pair-wise interactions. In all cases the simulation program consisted of a 2 ns annealing period at 550 K, followed by cooling to 298 K over 100 ps, finishing with 6 ns of room temperature simulation. Throughout the simulation the pressure was maintained at 1 atm. 10 snapshots spaced at 100 ps were extracted from the end of each trajectory, for a total of 50 snapshots for each material. All simulations were performed within the LAMMPS software suite (1). MD trajectories may overstate the disorder of real films, since long timescale ordering effects like annealing and molecular diffusion are

not captured within the accessible time window. Nevertheless, the MD simulations provide geometries that can be compared against one another, while allowing for the potential influence of experimental processing conditions not considered here. Details about the individual runs and force field parameters are included below.

Force Field Parameters. Standard OPLS stretching, bending, pair-wise, and torsional potentials were used in all simulations. The total force field energy, E_{FF}, is expressed as

$$E_{FF} = \sum_{bonds} k_{bond}\left(l - l_0\right)^2 + \sum_{angles} k_{angle}\left(\theta - \theta_0\right)^2 + \sum_{dihedrals} \frac{1}{2}\left\{V_1\left(1+\cos\phi\right) + V_2\left(1-\cos 2\phi\right) + V_3\left(1+\cos 3\phi\right) + V_4\left(1-\cos 4\phi\right)\right\} + \sum_{i>j}\left\{\frac{q_i q_j}{r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]\right\} \quad \textbf{[S1]}$$

where k and V terms are harmonic and OPLS dihedral force constants, respectively. r_{ij} are the interatomic separations, q_i are the atomic charges, and ε_{ij} and σ_{ij} are the Lennard-Jones variables for each pair-wise interaction. A 2 nm cutoff was used for all Coulomb and pair-wise interactions. LAMMPS pair style "lj/charmm/coul/long" was used to ensure smooth pair-wise potentials for separations between 1.9-2.0 nm. For PDI derivatives an additional harmonic improper potential was used based on previous work, to stiffen the imide/side-chain junction (2). Charges for fullerene derivatives were taken from earlier work by Cheung et al. (3) Charges on the PDI core were calculated using B3LYP/6-31G* optimized geometries and electron densities, post-processed using CHELPG (4)—all calculated within Q-Chem (5). OPLS parameters are primarily taken from (6) and supporting publications from Jorgensen's group, as tabulated and distributed with the BOSS 4.8 software (7). All parameters have been tabulated in a spreadsheet available for download on the journal website.

Cluster Generation. While the MD simulations were performed using periodic boundary conditions, in applications like solar cells, finite clusters are the relevant functional unit. For example, typical devices are less than 100 nm thick, meaning only ~100 molecules lie in a straight path between electrodes (8). Moreover, in bulk-heterojunction donor-acceptor blends, filamentous domains of each component are formed that typically only extend a few nanometers in thickness (9). We extract finite clusters from the periodic simulation cells by first unwrapping the

molecular coordinates of each snapshot. The use of periodic boundary conditions results in some molecules along the boundary getting divided across the simulation cell. The coordinates were unwrapped by translating the smaller fraction of all split molecules by the approriate simulation cell length(s). This ensures only whole molecules are obtained in each snapshot. The geometries obtained by this method can be considered as clusters extracted from the amorphous bulk of the respective materials.

Electronic Structure. The electronic structure of all MD snapshots was evaluated using the tight-binding hamiltonian developed by Hoffmann (10). The tight-binding framework utilizes an atomic orbital (AO) basis to construct the system hamiltonian, $\mathbf{H}_{AO}$, based on the overlap between AOs. The AO basis of the total system is expressed as $\sum_{m}^{N}\sum_{i}\left|\varphi_i^m\right\rangle$ where the index *m* runs over the *N* molecules in the system, and *i* runs over each molecule's AOs. To perform the network analysis it is necessary to transform $\mathbf{H}_{AO}$ into the MO basis. The MOs in the system were obtained by solving the generalized eigenproblem for every molecule in the system. For an individual molecule, *m*, the eigenproblem takes the form

$$\mathbf{H}_m\mathbf{V}_m = \mathbf{E}_m\mathbf{S}_m\mathbf{V}_m \qquad \textbf{[S2]}$$

where $\mathbf{H}_m$ is the molecular tight-binding hamiltonian, $\mathbf{S}_m$ is the non-diagonal molecular overlap matrix that arises from the non-orthogonality of the AO basis, $\mathbf{V}_m$ is the eigenvector matrix holding coefficients, c_{ij}^m, of the MOs for *m* in its columns, and $\mathbf{E}_m$ is the eigenvalue matrix with the corresponding MO eigenenergies along the diagonal. The *j*th MO of *m* is expressed in the *total system* AO basis trivially as

$$\left|\phi_{j,AO}^m\right\rangle = \sum_{n}^{N}\sum_{i}\delta_{mn}c_{ij}^n\left|\varphi_i^n\right\rangle \qquad \textbf{[S3]}$$

where the index *n* runs over all *N* molecules, the index *i* runs over all AOs on molecule *n*, φ_i^n is the *i*th AO belonging to molecule *n*, and the coefficients c_{ij}^n correspond to the elements of $\mathbf{V}_n$. The summation over *n* is necessary to ensure that each MO is properly indexed to $\mathbf{H}_{AO}$. The matrix elements of **H** were then constructed as

$$\mathbf{H} = \sum_{i}\sum_{j}\left\langle\phi_{i,AO}\middle|\mathbf{H}_{AO}\middle|\phi_{j,AO}\right\rangle\left|\phi_i\right\rangle\left\langle\phi_j\right| \qquad \textbf{[S4]}$$

where the indices *i* and *j* run over all MOs in the system, and the molecule indices have been dropped for clarity. The diagonal elements of Eq. **S4** correspond to the MO energies, ε_i, and off-diagonal elements correspond to the

MO couplings, V_{ij}, from which Eq. **1** is obtained. All eigenvalue evaluations and network analyses were performed using a code written by BMS. The YAEHMOP package (11) was used to construct the initial $\mathbf{H}_{AO}$ and $\mathbf{S}_{AO}$ matrices for the total system with default parameters for all atoms.

Additional Commentary on Data

Convergence with Respect to Number of Snapshots

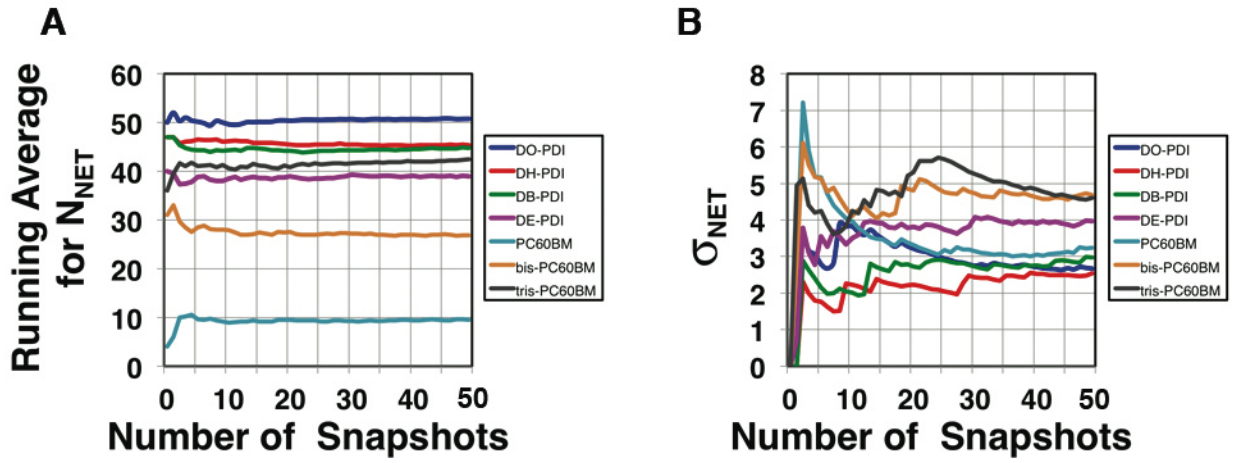


Fig. S1. (A) Running average for the number of networks in each material's simulation cell, N_{NET}, using a coupling threshold of 20 meV. (B) The running average for the standard deviation in the number of networks across all snapshots, σ_{NET} .

Snapshots were drawn sequentially from each trajectory and analyzed by network analysis until the running averages for the number of networks, N_{NET}, and standard deviation in the number of networks, σ_{NET}, converged for each material. Using 50 snapshots fulfilled these criteria for all materials. Convergence of both the mean and distribution suggests that enough snapshots were incorporated such that most of the important geometries *within the simulation window* (with respect to the variable N_{NET}) have been considered. The possibility of longer time ordering effects are neglected and cannot be ruled out as being important for a given material.

Convergence with Respect to Number of Trajectories

Below the glass transition temperature, the molecular reorganization is rather slow, allowing relatively little conformational phase space to be explored in a single trajectory. For each material system, five independent trajectories were simulated to partially ameliorate this problem. In each case a prior annealing simulation at 550 K was performed to avoid the formation of physically unrepresentative, kinetically arrested structures.

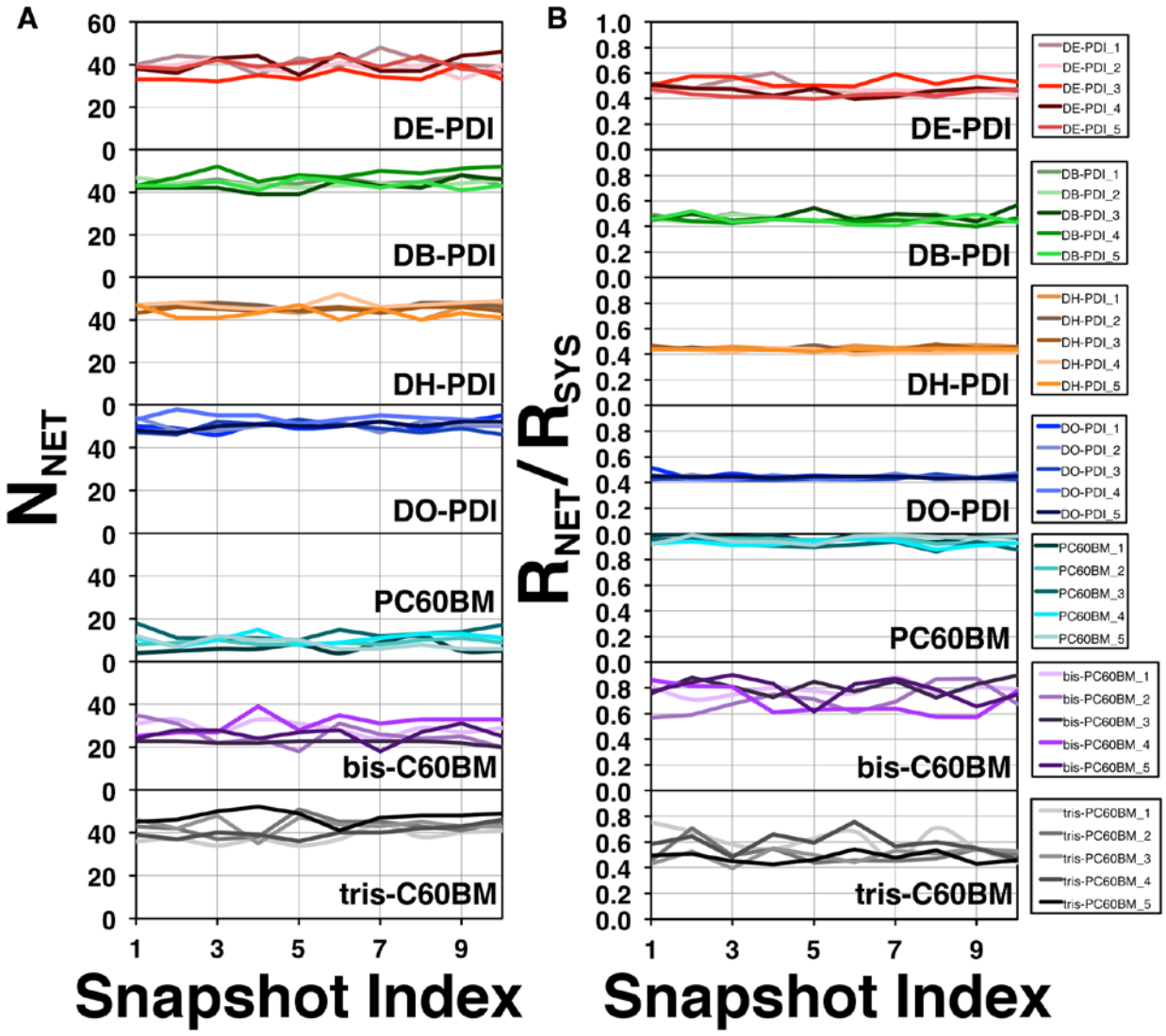


Fig. S2. (A) Comparison of N_{NET} between snapshots sampled from separate trajectories. (B) Comparison of the percolation ratio between snapshots sampled from separate trajectories. Both N_{NET} and the percolation ratio were calculated using a coupling threshold of 20 meV. Snapshots were sampled at 100 ps intervals, thus the x-axis can also be interpreted as a time.

However, it is not generally possible to guarantee convergence with respect to the number of trajectories. The simulations do exhibit a small but significant difference between the network properties for the various trajectories (Fig. S2). There is also dynamical evolution of the transport networks. Although it is beyond the scope of the present work, it should be highlighted that the timescale of the dynamical switching of connected neighbors is intrinsically related to the timescale of charge transport in non-percolating systems.

Dynamic percolation and additional network variables. In the present analysis, we have only incorporated the site energies and couplings into the network analysis; this is the minimal construction capable of describing mesoscopic charge transport between molecules. The framework can, and should, be generalized for the investigation of specific phenomenea. For instance, in analyzing networks for energy tranfer, long-range dipolar couplings should supplant the one-electron couplings.

Alternatively, for a realistic description of carrier mobility in amorphous systems, site reorganization energies, and dynamic percolation effects (i.e., the time-dependent evolution of the molecular networks) can be incorporated into the network construction. For example, the instantaneous network structure only carries meaning if a charge can traverse it on a comparable timescale to the geometric reorganizations that dynamically break and form connections. The figure of merit for determining which threshold matters is thus the timescale of charge transfer compared to the timescale of network reorganization. The percolation threshold must correspond to a charge transfer rate that is competitive with the network reorganization timescale, otherwise the reorganization timescale will determine the transport timescale. Percolative networks established below this threshold are meaningless instantaneously, since networks will dynamically evolve much faster than charge can traverse them. A full account of dynamic percolation is beyond the scope of the present work. This is the topic of a paper currently in preparation.

Surface Effects and the Size Dependence of Percolation. Deeper insight into the nature of fragmentation at higher thresholds can be obtained by examining sample snapshots from the network analysis shown in Fig. 3D. It can be seen that fragmentation begins predominantly along the surface of the cluster. This is a persistent effect that we have observed across many of the snapshots studied. This “surface effect” can be simply understood by considering the reduced number of neighbors for molecules situated at the cluster surface. Fewer neighbors represents fewer opportunities to establish electrical connections. Such surface effects can be parasitic since they have the potential to insulate the cluster surface. This holds importance for applications where charge is injected into the acceptor material, either at an electrode, or at an interface, as in transistors and solar cells.

In general, the percolation threshold is a function of the size of the cluster simulation. Due to the critical role of the surface defects, shape will also play a role, with higher surface area structures being more difficult to percolate. This tends to also suggest that systems that percolate moderate sized clusters will also percolate the bulk, since the surface makes the former more difficult to percolate.

The Role of State Degeneracy. The C_{60} derivatives are distinguished from the PDI derivatives by their three low-lying MOs arising from a three-fold degenerate fullerene LUMO. Functionalization breaks this degeneracy. Monofunctionalization strongly perturbs one of the three states, leaving two nearly isoenergetic LUMOs and a higher energy LUMO+2 ~ 0.3 eV above the others. Higher functionalization levels lead to further splitting in the low-lying unoccupied states, and the occurrence of regioisomers (12) (See below for details about the handling of regioisomers in the MD simulations). In our network analysis, we include all states within 0.3 eV of the vacuum LUMO level of each molecule. For the $PC_{60}BM$ clusters this leaves ~190 states percolating the cluster volume; for bis-$PC_{60}BM$ clusters ~140; for tris-$PC_{60}BM$ clusters ~160; whereas the PDI derivatives only have a single low-lying electronic state that can participate in the transport, resulting in 64 states within each cluster. With more states percolating the volume, the connectivity between molecules increases, as is reflected in the percolative behavior of the fullerene materials.

Number of Connections Per Molecule. The number of networks, N_{NET}, and the percolation ratio consider only the membership of molecules in the respective networks; the number of connected neighbors per molecule is not directly reflected in the largest network size or total number of networks. Consider that a minimally connected linear chain in which each molecule only has two connected neighbors can percolate an entire volume. By the metrics of R_{NET} and N_{NET}, a minimally connected system is indistinguishable from a better connected branched structure, because the network membership is identical, and the size of the networks would be the same. However, the two systems differ in at least two measurable ways. Transport in the well-connected network should be affected by the existence of multiple paths between sites, and the well-connected system should also be more robust to structural disorder. With multiple connected neighbors, there are built-in redundancies for transport. The well-connected system can tolerate disruptions between several sites without total collapse of percolation, whereas a single disruption in the minimally connected system destroys percolation. This second property relates directly to the processing demands and failure behavior of a given material. It is also interesting to speculate that with multiple paths between sites, transport interference effects on the many-molecule scale might play a role in some organic semiconductors.

In the main text, note that fragmentation occurs when the average number of connections falls below 2. Two is important, because the number of connections per molecule for a minimally connected linear chain grows as $2(N-1)N^{-1}$ (two connections per molecule on the chain interior, one connection for each terminal molecule; when

N is large this converges to two). Thus, when the average number of connected neighbors falls beneath ~2, network fragmentation is guaranteed; only structures with > 2 connections per molecule are capable of non-fragmented percolation of a volume. This point will be further explored in future work, but here it is sufficient to note that this metric offers a rigorous mathematical delineation between percolative and non-percolative molecular structures.

Data Visualization. The MO's and molecular structures shown in Fig. 1 were rendered in Blender. All snapshots and networks were visualized in VMD (13, 14).

Chemical Abbreviations.

PC60BM: [6,6]-phenyl-C_{61}-butyric acid methyl ester

Bis-PC60BM: bis-[6,6]-phenyl-C_{61}-butyric acid methyl ester

Tris-PC60BM: tris-[6,6]-phenyl-C_{61}-butyric acid methyl ester

DE-PDI: *N,N′*-Bis(1-ethylpropyl)perylene-3,4,9,10-tetracarboxylic diimide

DB-PDI: *N,N′*-Bis(1-butylpentyl)perylene-3,4,9,10-tetracarboxylic diimide

DH-PDI: *N,N′*-Bis(1-hexylheptyl)perylene-3,4,9,10-tetracarboxylic diimide

DO-PDI: *N,N′*-Bis(1-octylnonyl)perylene-3,4,9,10-tetracarboxylic diimide

Treatment of Regioisomers. For bis-addition to the fullerene, there are 8 distinguishable 6'-6 bonds that the second adduct can react with. We use the labeling scheme of Hirsch (15), referring to 6'-6 bonds on the same hemisphere as the first adduct as "cis", those on the opposite hemisphere as "trans", and those on the equator as "e", as shown in Figure S3. Numbering of these bonds refers their distance from the equator, with cis-1 and trans-1 being the furthest available 6'-6 bonds from the equator.

Considering the degeneracy of each position (1-fold for trans-1, 2-fold for trans-2,3,4 and cis-2,3,4 and 4-fold for e) there are 8 distinguishable regioisomers in total. In addition, each regioisomer has two distinguishable diasteriomers due to the asymmetry of the phenyl butyric acid methyl ester adduct. When generating the initial geometries for the bis-adducts we place the regioisomers according to the relative abundances published by Djojo et al. for bis-(4-phenyloxazoline) with fullerene, and assume there is no preference between diasteriomers (15, 16). Regioisomer (abundance): cis-1 (0.00), cis-2 (0.02), cis-3 (0.06), e (0.38), trans-1 (0.02), trans-2 (0.13), trans-3 (0.30), trans-4 (0.09)

Since no abundance information is available for the tris adducts, we employ assumptions in line with the known experimental data. In our simulations we discount the cis isomers as being sterically unfavorable, as has

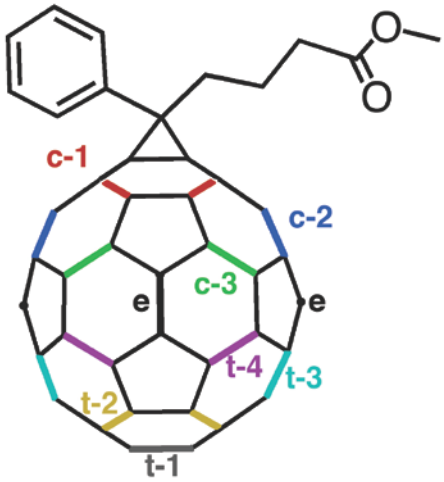


Fig. S3. Labels show the several possible 6'-6 bonding relationships, named according the scheme of Hirsch. Here "c" refers to cis, "t" to trans, and "e" to equitorial. The same pairwise substitutional relationships apply for tris adducts, although chiral center formation must be considered for some properties.

been confirmed for isolated yields of bis-adducts. Similarly, the trans-1 substitution is unfavorable, since the addition of the first adduct deactivates the trans 6'-6 bond. For these reasons, we follow the reasoning of Djojo et al. and limit the simulated clusters to include only tris-adducts possessing trans and e positional relationships (17). This leads to ten possible distinguishable regioisomers: (e,e,e), (e,e,t1), (e,e,t1)', (e,t2,t3), (e,t2,t4), (e,t3,t4), (t2,t4,t4), (t3,t3,t3), (t3,t3,t4), (t4,t4,t4). As above, the occurrence of diasteriomers is considered randomly distributed.

Pairwise Correlation Analysis. The principal concern of this paper is to establish the relationship between structural order and electrical connectivity. Part of this investigation concerns the degree to which even local orientational ordering, such as stacking, correlates with electrical connectivity. Such local orientational ordering might be present even in the absence of translational ordering, and remain undetectable by X-ray diffraction. To this end, we have calculated the orientational correlation factors, $\langle \mathbf{n} \rangle_{\bar{y}}$, between molecules within each network and across all materials. The correlation between two vectors takes the form of a simple normalized dot product:

$$\langle \mathbf{n} \rangle_{\bar{y}} = \frac{\mathbf{n}_i \cdot \mathbf{n}_j}{|n_i||n_j|} \qquad \textbf{[S5]}$$

where $\mathbf{n}_i$ and $\mathbf{n}_j$ are defined alignment vectors for each molecule being compared. When the vectors are perfectly correlated (aligned), $\langle \mathbf{n} \rangle_{ij} = 1$, and anti-correlated (anti-aligned), $\langle \mathbf{n} \rangle_{ij} = -1$. Depending the symmetry of the molecule, anti-correlation may not be physically meaningful (e.g. the 2-D PDI π-system), in which case the absolute value of the dot product in (8) is used and $\langle \mathbf{n} \rangle_{ij} \in [0,1]$. The correlation factor can also be interpreted as a dihedral angle, θ_{ij}, between the alignment vectors $\mathbf{n}_i$ and $\mathbf{n}_j$ via $\theta_{ij} = \cos^{-1} \langle \mathbf{n} \rangle_{ij}$.

It is critical to meaningfully choose the alignment vector for each material. For fullerenes, we used the vector defined by one of the functionalized fullerene carbons, and its pair on the opposite end of the fullerene cage (Fig. S4). This vector tracks the location of the functionalization on the surface of the fullerene moiety. The correlation factors calculated using this alignment vector reflect the relative orientation of sidechains between compared molecules. Unfortunately, for the higher adducts it is impossible to preserve this vector relative to the secondary and tertiary functionalizations, since several regioisomers are present in the simulation cell, and this partially convolutes the interpretation of the correlation factor for these materials.

For the PDI-derivatives, the vector normal of the π-system was used as the alignment vector. This choice was made to assay the alignment between the π-systems of the compared molecules. The normal vector was determined for each molecule by the cross product of the long and short axes of the PDI moiety, where the long axis is defined by the vector connecting the imide nitrogen atoms, and the short axis by the vector connecting two bay carbons (Fig. S5). All correlation calculations were performed using networks obtained using a coupling threshold of 20 meV.

A correlation matrix for each material is shown in Figs. S4 and S5 (middle-left column) of the largest network from a representative snapshot. In this plot, each square corresponds to a correlation factor, colored according to its magnitude, and the *x* and *y* indices correspond to the molecules in the investigated network. More pixels denote larger networks, and monochromaticity denotes consistency in relative orientations. The diagonal represents self-correlations and is therefore one. This limited dataset already foreshadows many of the conclusions developed in the main text. The network size decreases with increasing functionalization and all fullerene derivatives form more expansive networks than the PDI–based materials. Moreover, the relative orientations of the

butyric acid functionalizations are randomized indicating the very relaxed orientational criteria for establishing electrical connections between fullerenes.

The spatial dependence of the correlation factors is plotted using normalized 3-D histograms in Figs. S4 and S5 (middle-right column). If the *x*-axis of this plot were collapsed, the resulting data plot would correspond to the pair-distribution function between molecules within the same network, tabulated over all networks in all snapshots. The contraction to short values of R_{ij} observed for tris-$PC_{60}BM$ and all PDI derivatives, reflects the fewer long-range interactions due to network contraction in these systems.

The spatially resolved histograms also clearly quantify the topological influence on network formation and shape. The narrow bunching of counts in the bottom left of each PDI histogram (Fig S5, middle-right), reflects the strong correlation between π-system alignment and electrical connectivity. The alignment implies that the PDI molecules in the same network strongly favor "stacking" interactions. *This should not be interpreted as reflecting superior ordering in PDI-derivatives*; rather, it is only evidence that the PDI materials have more stringent orientational constraints for establishing electrical connections, whereas the fullerene materials can establish connections at many different orientations.

Finally, the local order can be assessed by calculating the correlation histograms restricted to pairs of molecules within the first nearest neighbor shell. Based on the spatially resolved histograms, centroid-centroid separation criteria of $R_{ij} < 1.5$ nm and $R_{ij} < 0.7$ nm were used for the C_{60}-derivatives and PDI-derivatives, respectively. This histogram reflects how neighbors within each network are statistically oriented relative to one another. The wide distribution of orientations in C_{60}-derivatives rigorously quantifies their weak orientational preference, even locally. For PDI-derivatives, electrical connections between neighbors are almost exclusively limited to stacked molecules.

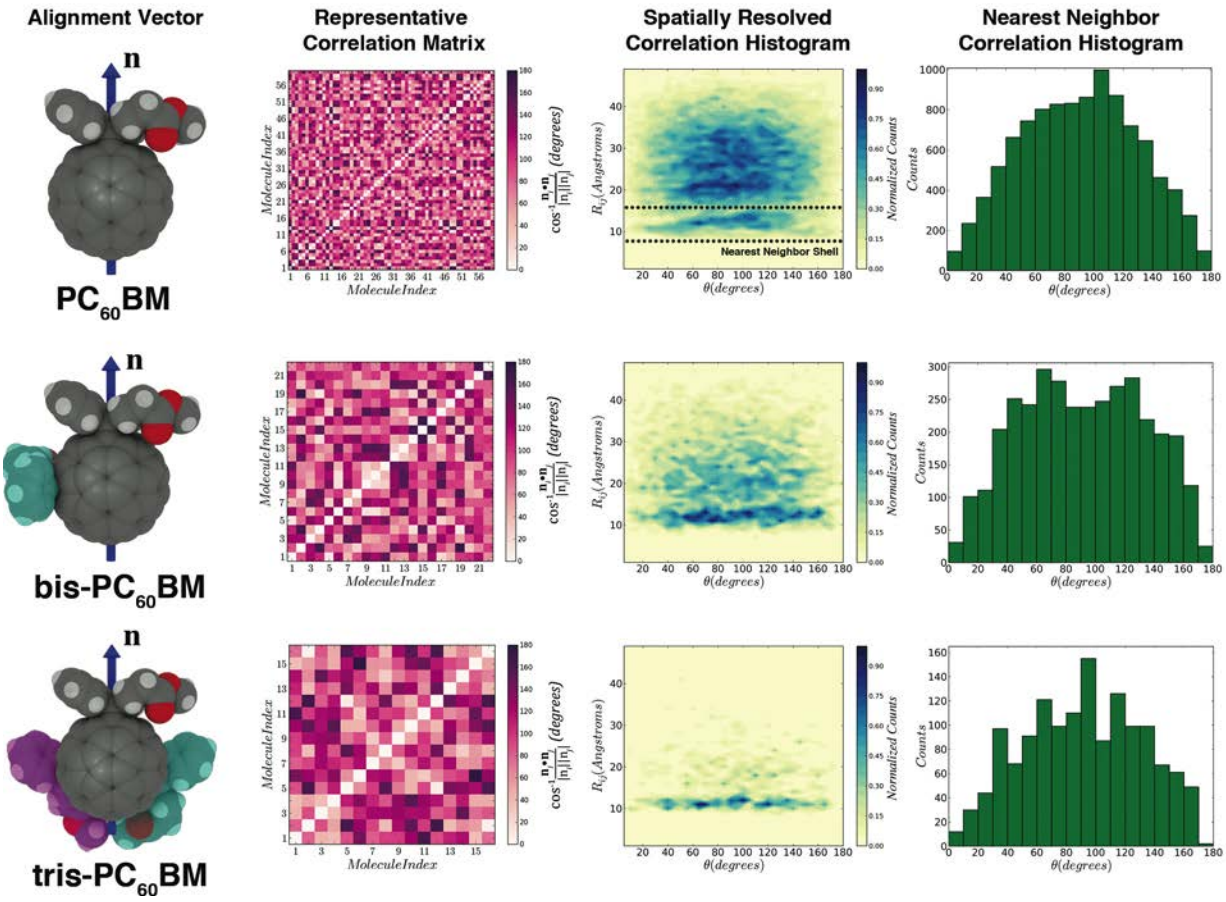


Fig. S4. (Leftmost) Steepest decent minimized geometries of each molecule, with the orientational vector indicated by **n**. (middle, left) A correlation matrix calculated between molecules from the largest network of a representative snapshot. (middle, right) The correlation factors across all snapshots were calculated then formed into histograms based on the radial separation of the molecules being compared. A histogram bin size of 1 angstrom in R_{ij} and 5° in θ were used to generate these plots. This calculation, in principle, quantifies the lengthscale of correlations; although, for the C_{60}-derivatives there are no correlations to observe. (Rightmost) The correlation histogram calculated across all snapshots, but only between molecules with a centroid-centroid separation <1.5 nm and using a bin size of 10° in θ. This provides a measure of the nearest-neighbor correlation factors. The wide range of correlation factors reflects the absence of a strongly preferred orientation between two connected C_{60}-derivatives. A weak preference for avoiding the extrema is partly due to the small conformation space reflected by the extrema. In the 0° case, it probably also reflects a preference to avoid steric interaction between the functionalizations.

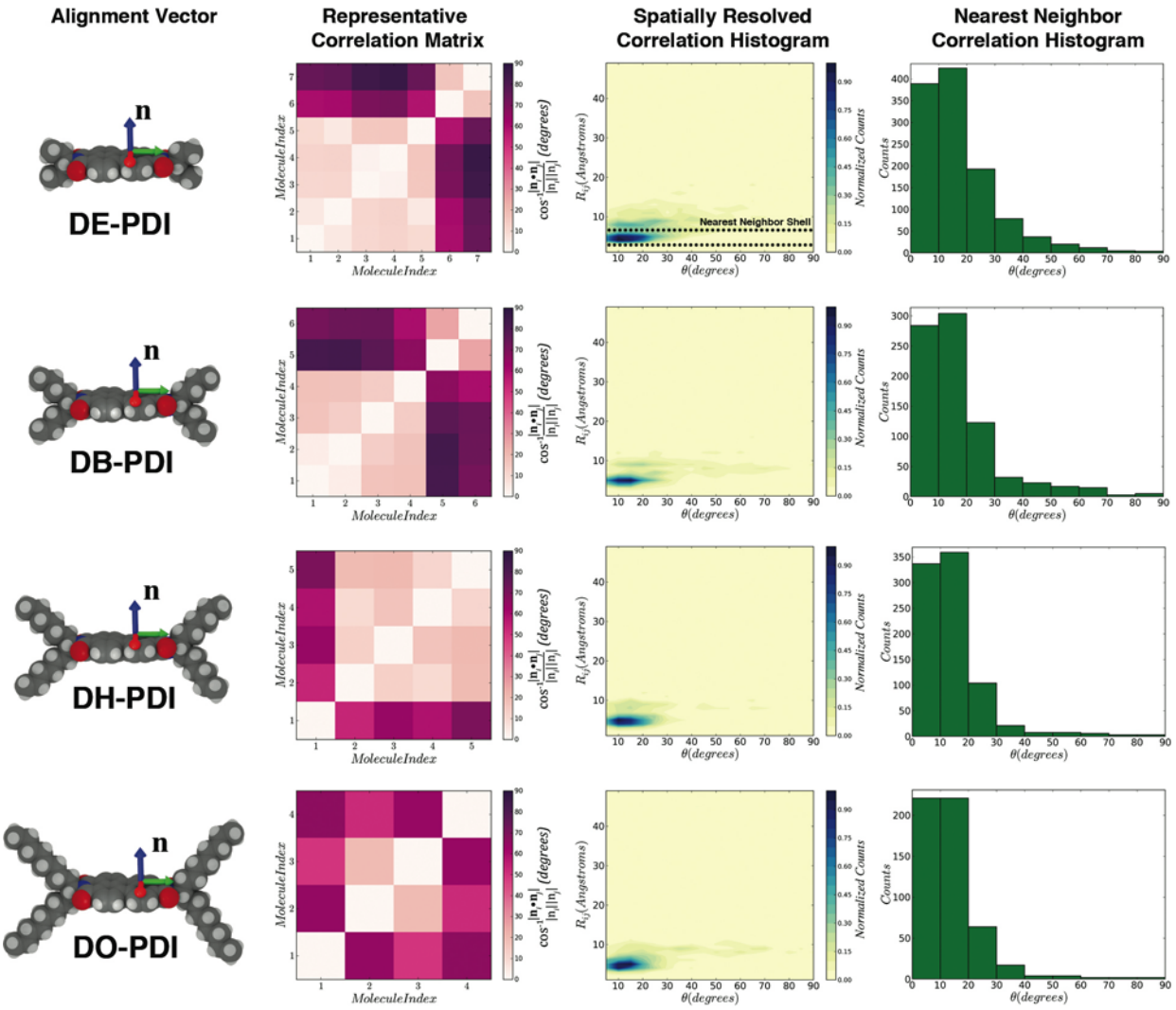


Fig. S5. (Leftmost) Steepest decent minimized geometries of each molecule, with the orientational vector indicated by **n**. (middle, left) A correlation matrix calculated between molecules from the largest network of a representative snapshot. (middle, right) The correlation factors across all snapshots were calculated then formed into histograms based on the radial separation of the molecules being compared. The PDI's show a strong preference for alignment; however, long-range correlations are rare because of the limited extension of the networks. (Rightmost) The correlation histogram was calculated across all snapshots, but only between molecules with a centroid-centroid separation <0.70 nm. This provides a measure of the nearest-neighbor correlation factors and indicates, both, the strong preference of PDI-derivatives to form stacks and that stacking frustration by the sidechains limits network formation.

Spatial Distribution of Electronic Couplings and MOs.

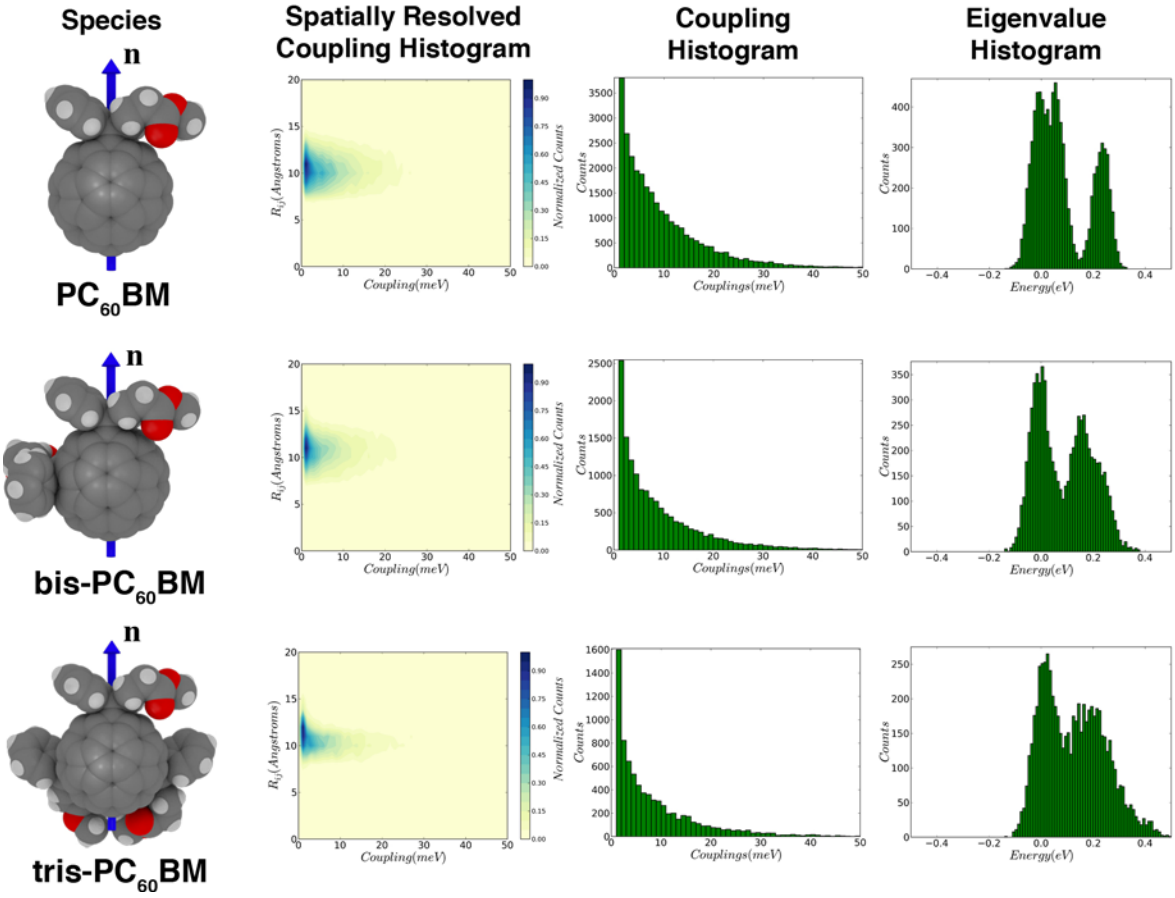


Fig. S6. (Leftmost) Steepest decent minimized geometries of each molecule. For bis-$PC_{60}BM$ and tris-$PC_{60}BM$, only a single regioisomer is shown. (Middle, left) The absolute value of the couplings, summed over all snapshots, displayed in a 3-D histogram based on the radial separation of the coupled molecules' centroids. (Middle, right) A histogram of the pair-wise electronic coupling magnitudes for all states, in all snapshots. Since any individual state is only coupled to a small number of proximate states, most couplings are negligible. To visualize the meaningful couplings, only couplings >1 meV have been included. (Rightmost) The eigenenergy histogram calculated across all snapshots and centered about the vacuum LUMO level of each molecule ($PC_{60}BM$: -9.903 eV, bis-$PC_{60}BM$: -9.861 eV, tris-$PC_{60}BM$: -9.839 eV calculated using the EH hamiltonian). For the bis-$PC_{60}BM$ and tris-$PC_{60}BM$ eigenenergy histograms, the "e" and "t2-t4-t4" regioisomer (see Fig. S3) LUMO energies were used to center the distribution. For the bis-$PC_{60}BM$ series, the "e" isomer is the most abundant. For the tris-$PC_{60}BM$ series, whose regioisomeric abundances are not known, the "t2-t4-t4" isomer has a LUMO energy near the mean of the LUMOs of all of the regioisomers. All histogram energies and couplings correspond to diagonal and off-diagonal components of localized Hamiltonians assembled for each snapshot. 1 meV, 1 Å, and 10 meV bins were used to generate histograms for the coupling magnitude, centroid-centroid separation, and eigenenergies, respectively.

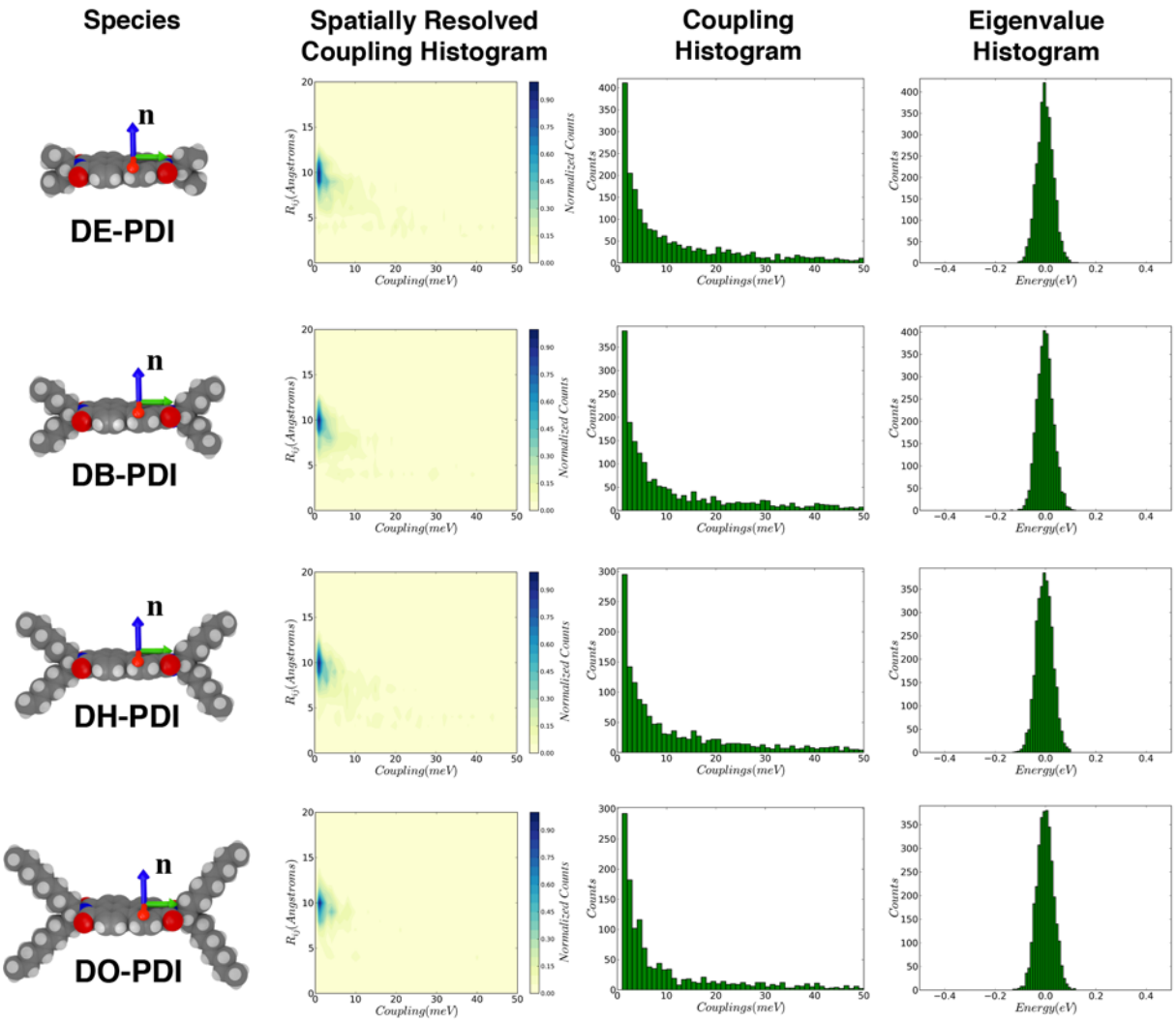


Fig. S7. (Leftmost) Steepest decent minimized geometries of each molecule. (middle, left) The absolute value of the couplings, summed over all snapshots, displayed in a 3-D histogram based on the radial separation of the coupled molecules' centroids. (Middle, right) A histogram of the pair-wise electronic coupling magnitudes for all states, in all snapshots. Since only any individual state is only coupled to a small number of proximate states, most couplings are negligible. To visualize the meaningful couplings, only couplings >1 meV have been included. (Rightmost) The eigenenergy histogram calculated across all snapshots and centered about the vacuum LUMO level of each molecule (-10.65 eV for all PDI derivatives using the EH hamiltonian). All histogram energies and couplings correspond to diagonal and off-diagonal components of localized Hamiltonian assembled for each snapshot. 1 meV, 1 Å, and 10 meV bins were used to generate histograms for the coupling magnitude, centroid-centroid separation, and eigenenergies, respectively.

Surface Area and Volume Analysis.

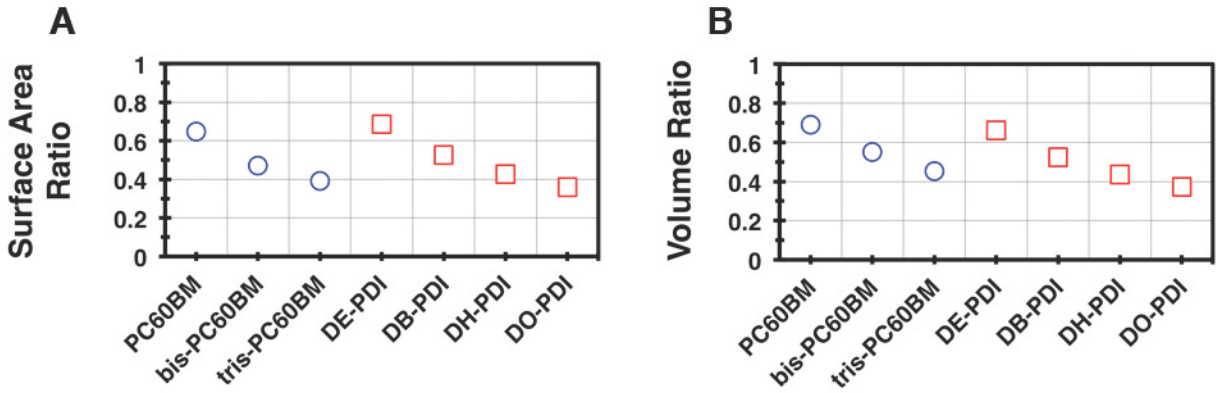
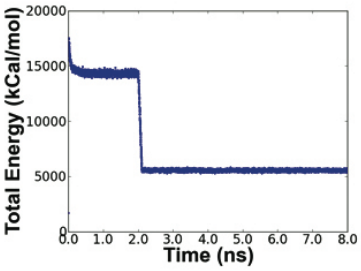
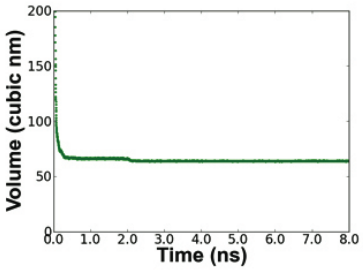


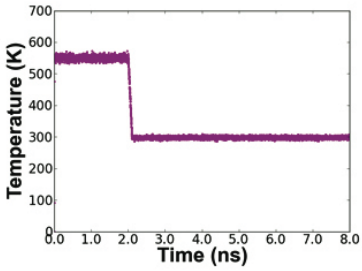
Fig. S8. (A) Ratio of the surface area of each reference molecule to the surface area of each functionalized molecule. (B) Ratio of the volume of each reference molecule to the volume of each functionalized molecule. The reference molecule for the $PC_{60}BM$ family is C_{60}. The reference molecule for the PDI family is H-PDI (side-chain replaced by a hydrogen). Areas and volumes were calculated for the solvent excluded surfaces around each molecule, using the Chimera software (18, 19). The surface area and volume ratios of each derivative with its parent are the simplest measures for how functionalization occludes and dilutes the conjugated moiety, respectively. Neither measure accurately reproduces the trends observed in the connections per molecule and contraction results. For example, DE-PDI has a more accessible conjugated moiety per molecule compared with $PC_{60}BM$, yet DE-PDI has more contracted and fragile network properties. Moreover, both the surface area and volume ratio data suggest that bis-$PC_{60}BM$ behavior should be closer to tris-$PC_{60}BM$ than $PC_{60}BM$; whereas, the network properties of tris-$PC_{60}BM$ is qualitatively different from both.

$PC_{60}BM$ Run Details

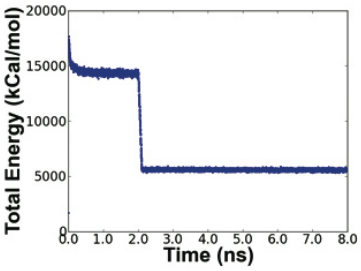
Trajectory 1

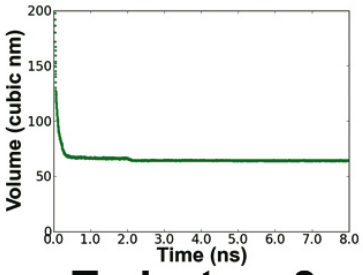


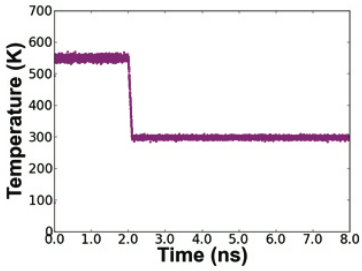




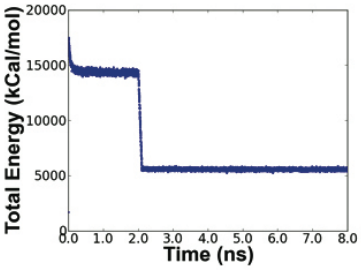
Trajectory 2

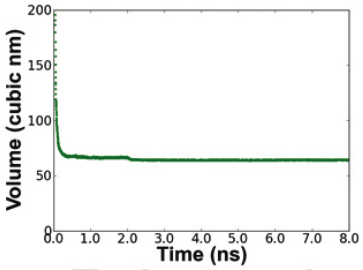


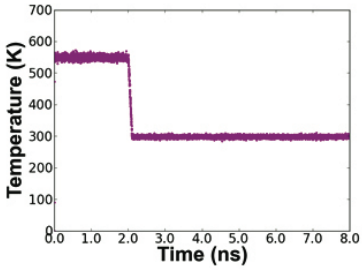




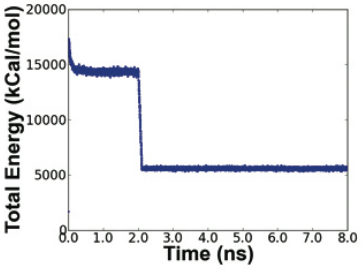
Trajectory 3

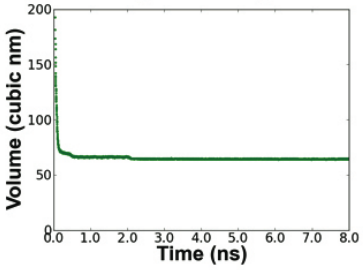


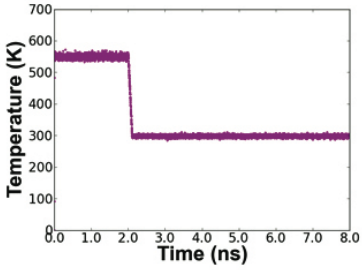




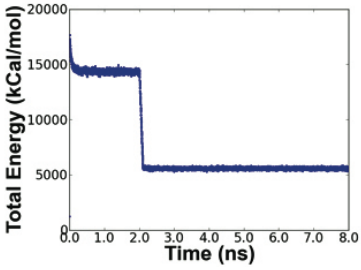
Trajectory 4

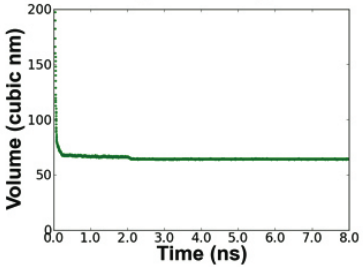


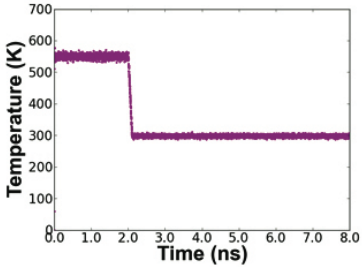




Trajectory 5

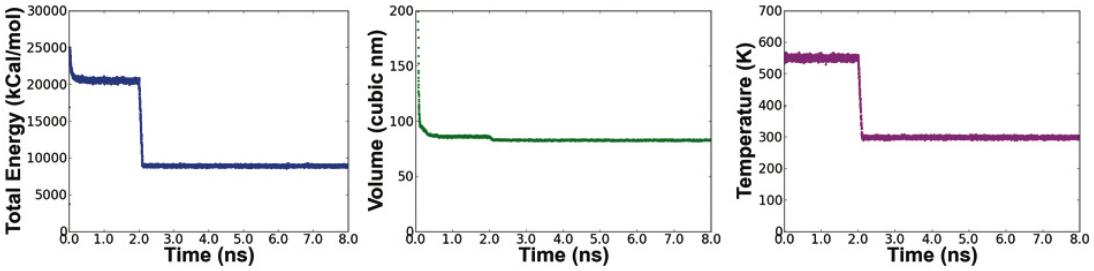




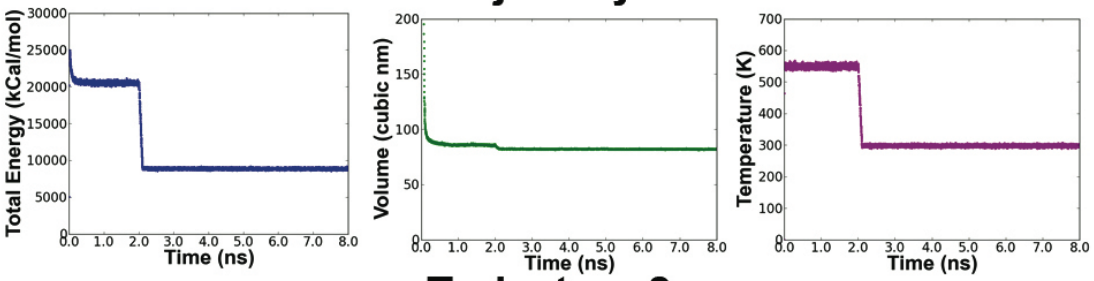


bis-$PC_{60}BM$ Run Details

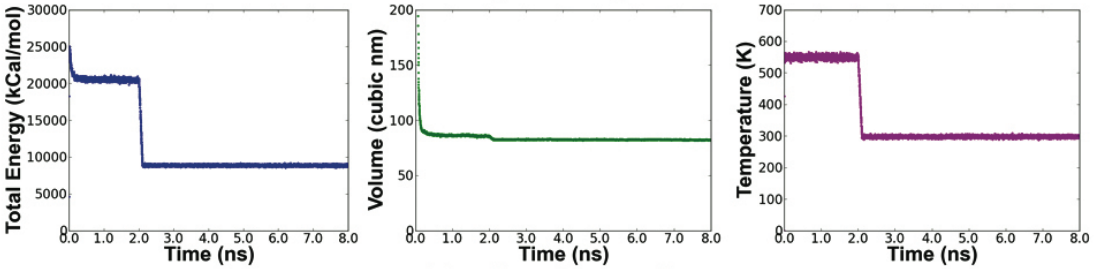
Trajectory 1



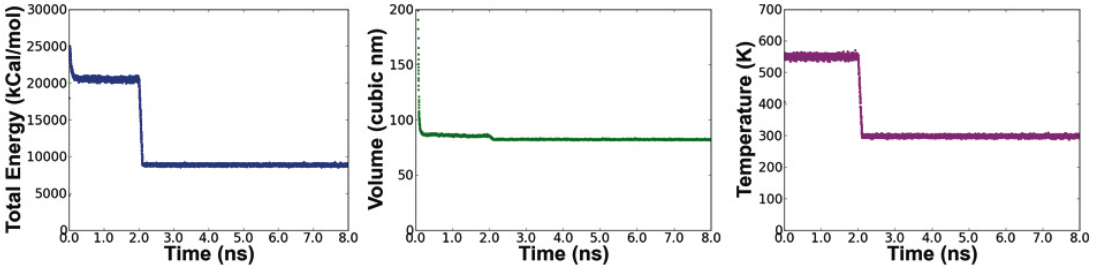
Trajectory 2



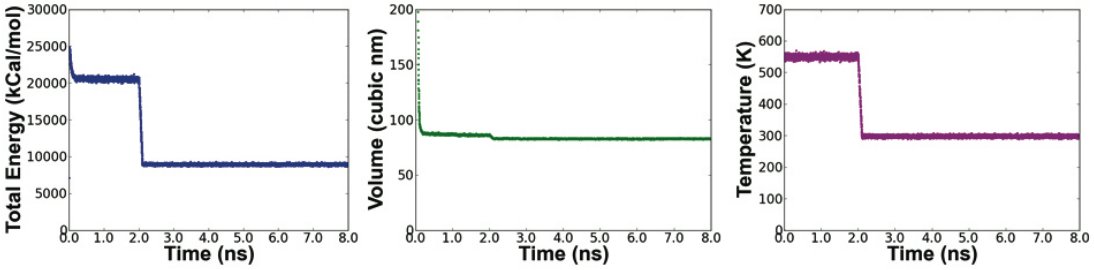
Trajectory 3



Trajectory 4

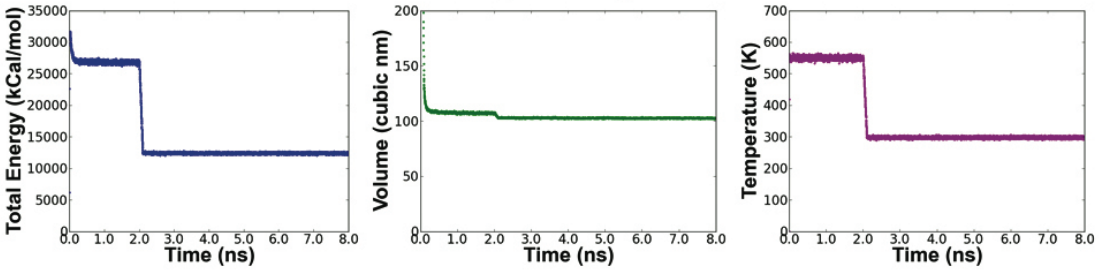


Trajectory 5

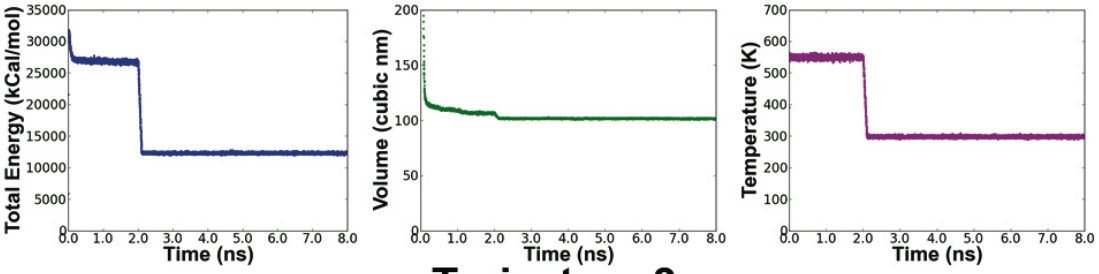


tris-$PC_{60}BM$ Run Details

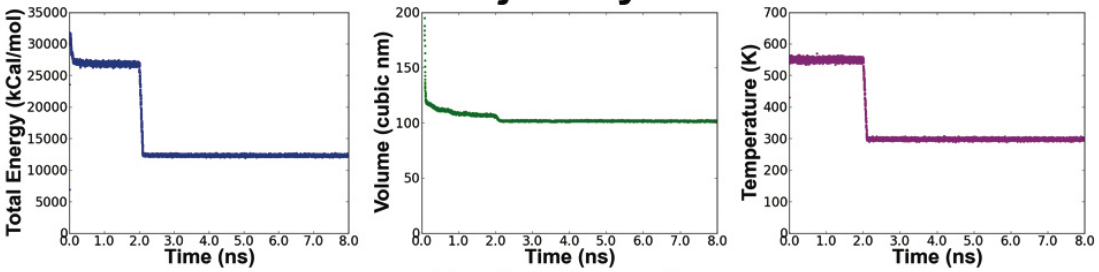
Trajectory 1



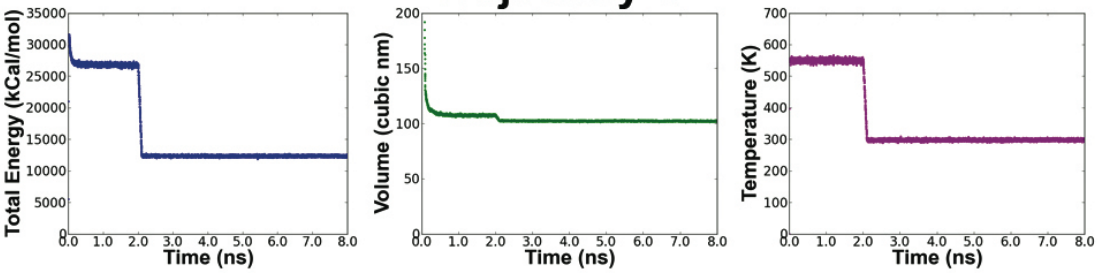
Trajectory 2



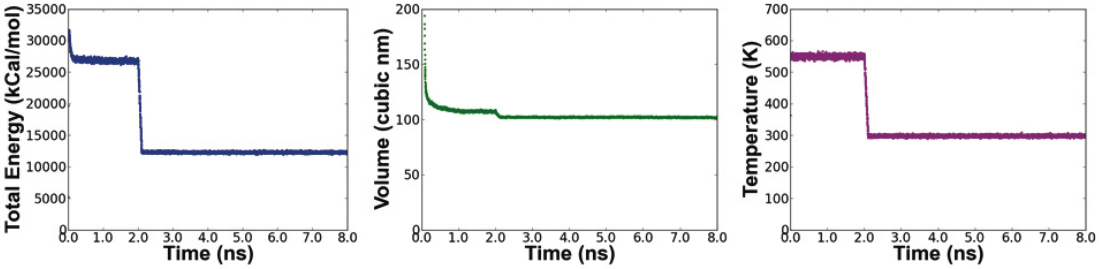
Trajectory 3



Trajectory 4

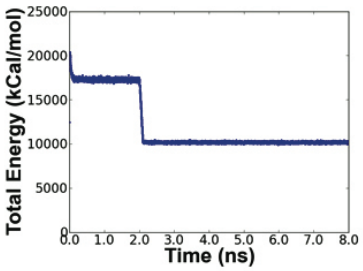


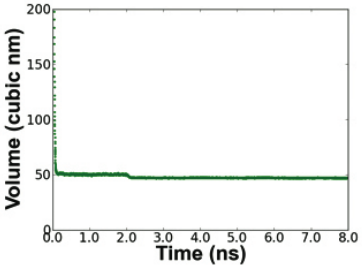
Trajectory 5

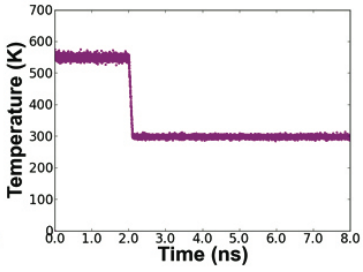


DE-PDI Run Details

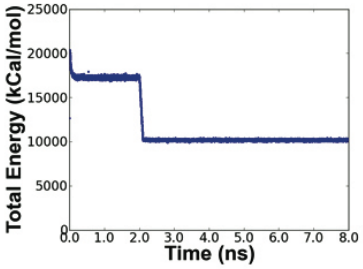
Trajectory 1

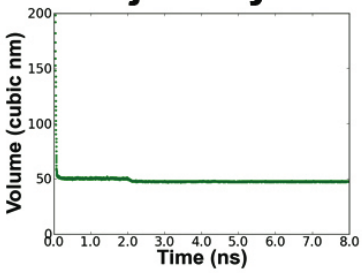


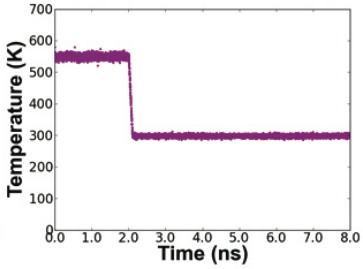




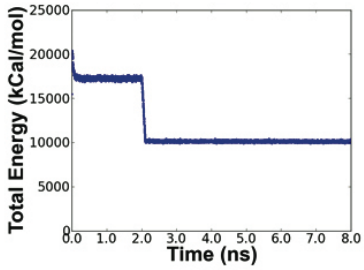
Trajectory 2

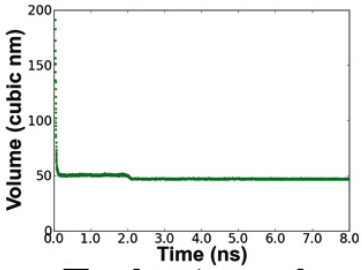


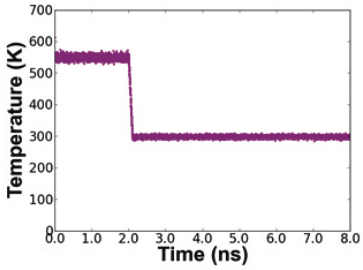




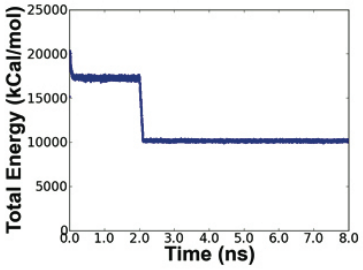
Trajectory 3

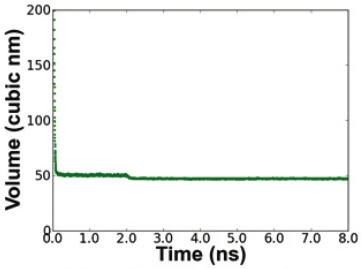


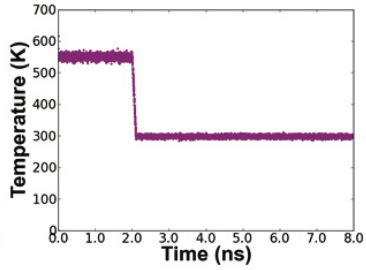




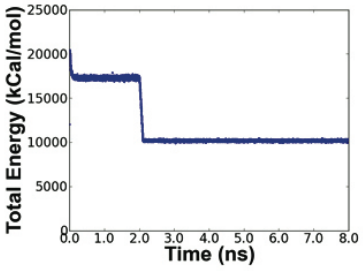
Trajectory 4

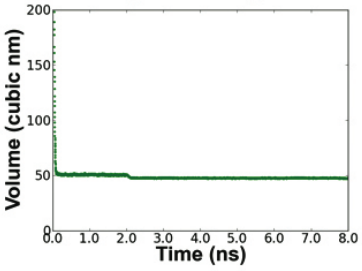


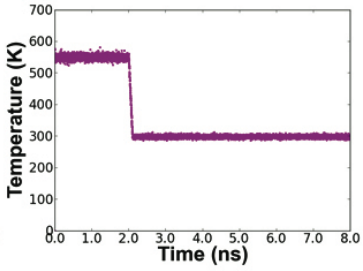




Trajectory 5

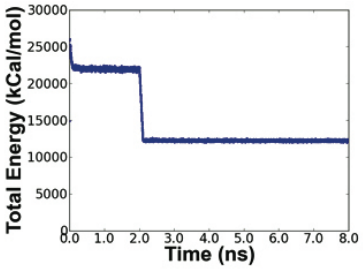


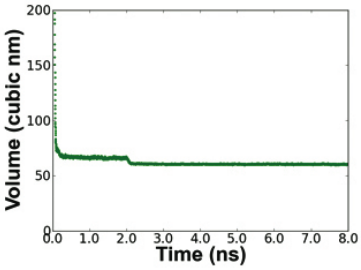


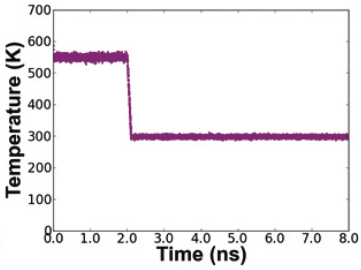


DB-PDI Run Details

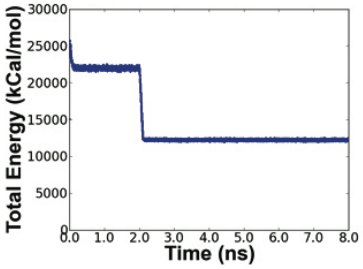
Trajectory 1

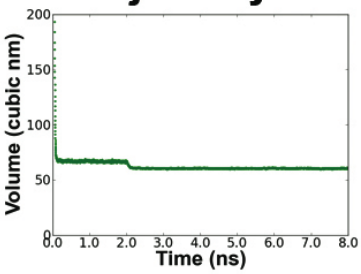


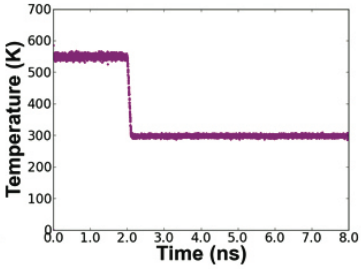




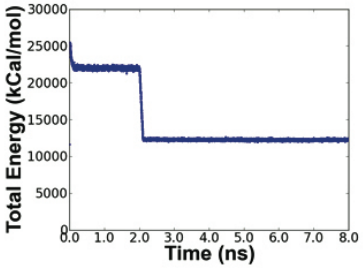
Trajectory 2

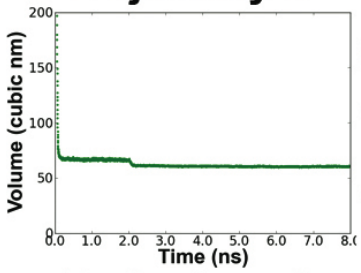


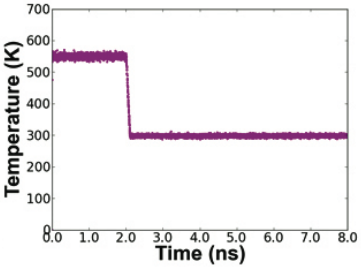




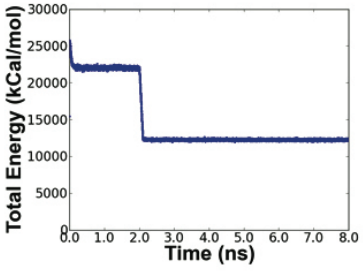
Trajectory 3

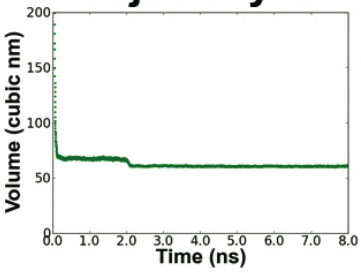


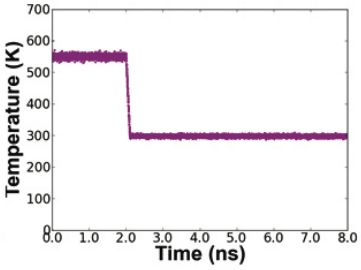




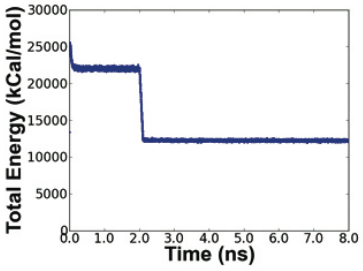
Trajectory 4

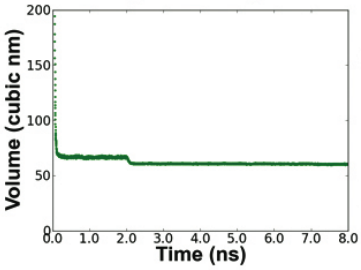


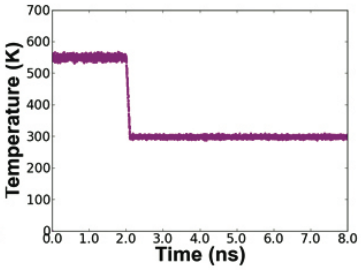




Trajectory 5

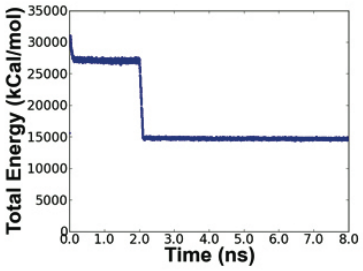


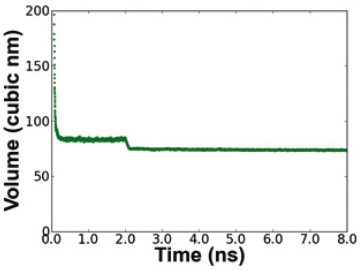


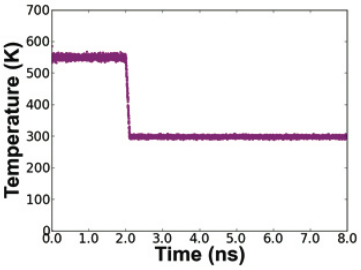


DH-PDI Run Details

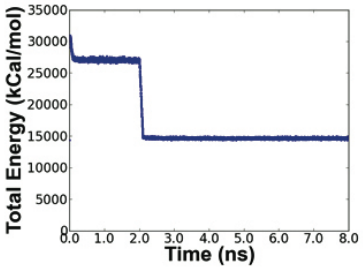
Trajectory 1

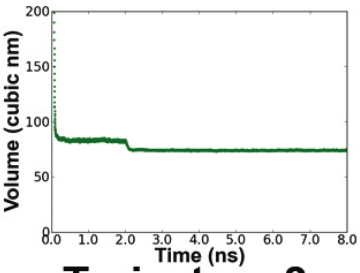


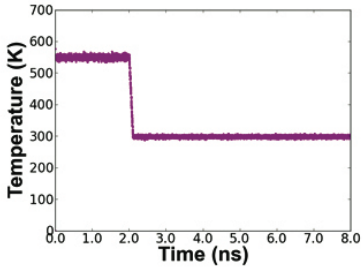




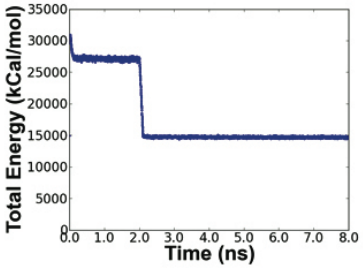
Trajectory 2

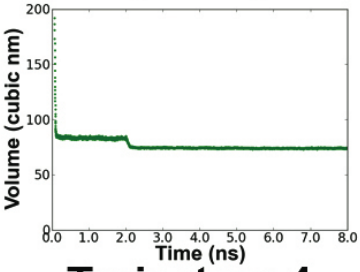


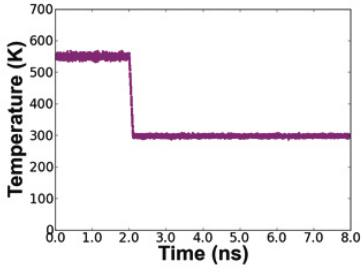




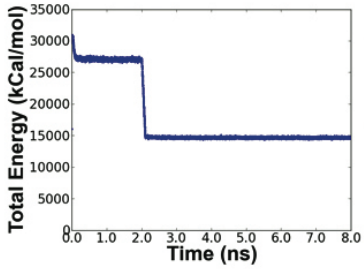
Trajectory 3

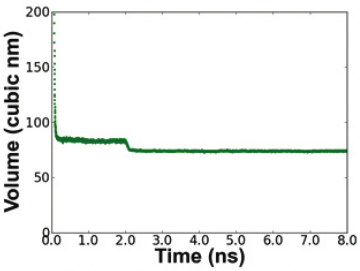


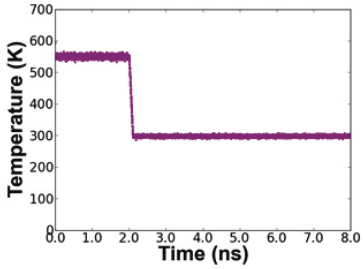




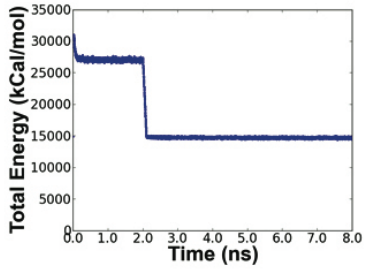
Trajectory 4

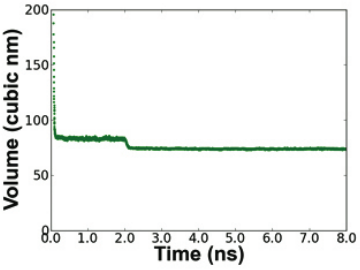


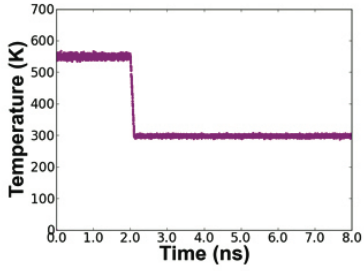




Trajectory 5

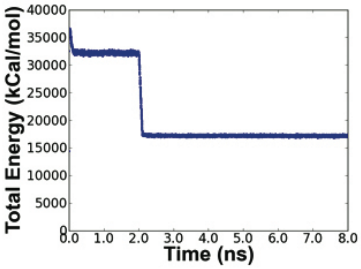


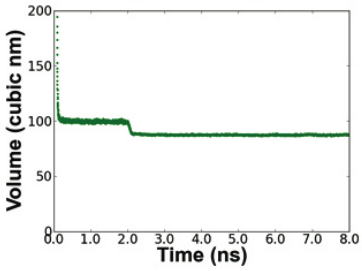


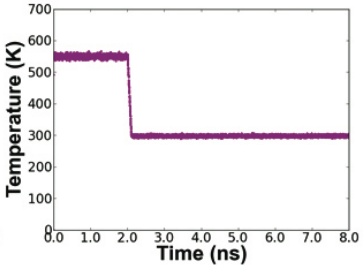


DO-PDI Run Details

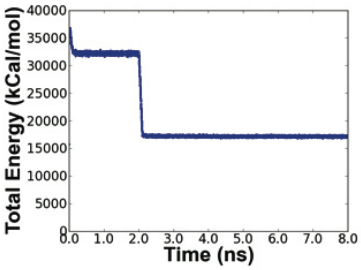
Trajectory 1

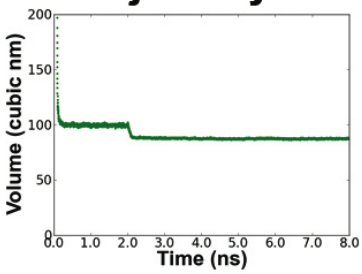


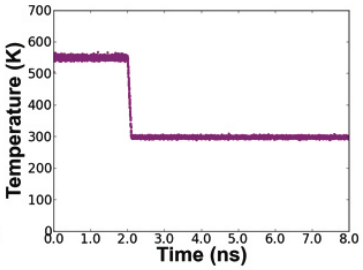




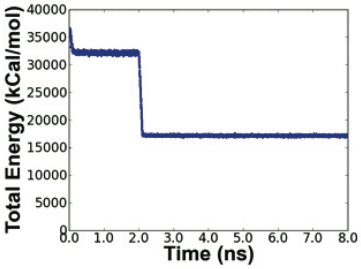
Trajectory 2

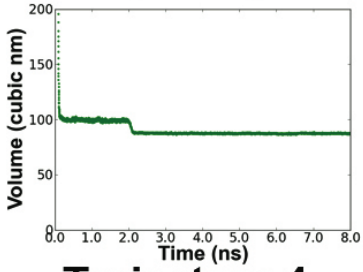


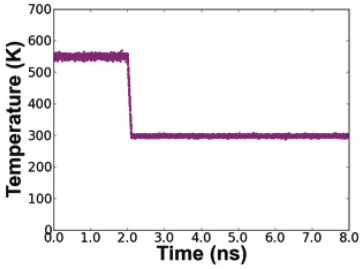




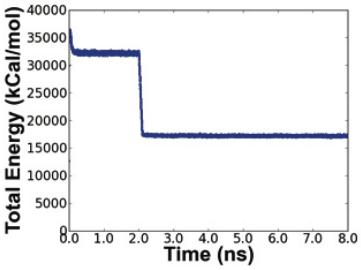
Trajectory 3

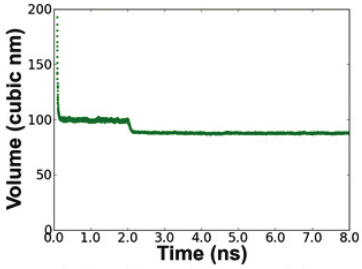


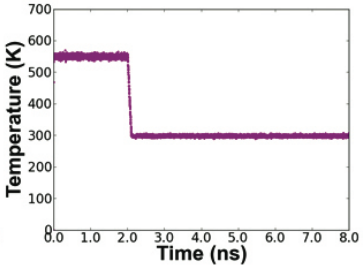




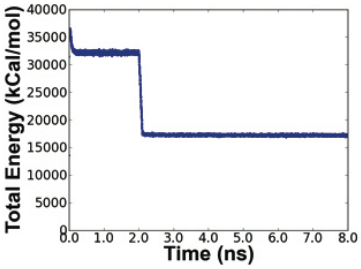
Trajectory 4

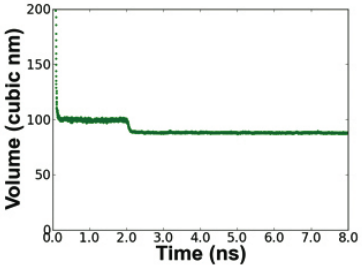


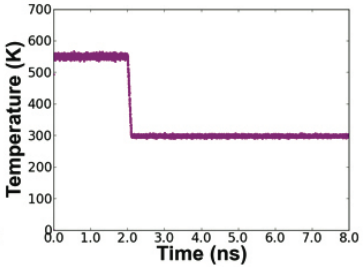




Trajectory 5







Supplementary References:

1. Plimpton S (1995) Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J Comp Phys* 117:1-19.
2. Marcon V, *et al.* (2009) Understanding Structure−Mobility Relations for Perylene Tetracarboxydiimide Derivatives. *J Am Chem Soc* 131(32):11426-11432.
3. Cheung DL & Troisi A (2010) Theoretical Study of the Organic Photovoltaic Electron Acceptor PCBM: Morphology, Electronic Structure, and Charge Localization. *J Phys Chem C* 114(48):20479-20488.
4. Breneman CM & Wiberg KB (1990) Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis. *J Comput Chem* 11:361-373.
5. Shao Y, *et al.* (2006) Advances in methods and algorithms in a modern quantum chemistry program package. *Phys Chem Chem Phys* 8(27):3172-3191.
6. Jorgensen W, Maxwell D, & Tirado-Rives J (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* 118:11225-11236.
7. Jorgensen WL & Tirado-Rives J (2005) Molecular modeling of organic and biomolecular systems using BOSS and MCPRO. *J Comput Chem* 26(16):1689-1700.
8. Li G, Zhu R, & Yang Y (2012) Polymer solar cells. *Nature Photon* 6(3):153-161.
9. Bartelt J, *et al.* (2013) The Importance of Fullerene Percolation in the Mixed Regions of Polymer–Fullerene Bulk Heterojunction Solar Cells. *Adv Energy Mater* 3:364-374.
10. Hoffmann R (1963) An extended Hückel theory. I. hydrocarbons. *J Chem Phys* 39:1397-1412.
11. Landrum GA & Glassy WV (2001) The YAEHMOP Project).
12. Frost J, Faist M, & Nelson J (2010) Energetic disorder in higher fullerene adducts: A quantum chemical and voltammetric study. *Adv Mater* 22:4881-4884.
13. Stone J (1998) An Efficient Library for Parallel Ray Tracing and Animation. Masters Masters Thesis (University of Missouri-Rolla).
14. Humphrey W, Dalke A, & Schulten K (1996) VMD - Visual Molecular Dynamics. *J Mol Graphics* 14:33-38.
15. Hirsch A (1999) *Principles of Fullerene Reactivity* pp 1-65.
16. Djojo F, Herzog A, Lamparth I, Hampel F, & Hirsch A (1996) Regiochemistry of Twofold Additions to [6, 6] Bonds in C60: Influence of the Addend-Independent Cage Distortion in 1, 2-Monoadducts. *Chem Eur J* 2:1537-1547.
17. Djojo F, Hirsch A, & Grimme S (1999) The Addition Patterns of C60 Trisadducts Involving the Positional Relationships e and trans-n (n= 2–4): Isolation, Properties, and Determination of the Absolute Configuration of Tris(malonates) and Tris[bis(oxazolines)]. *Eur J Org Chem* 1999(11):3027-3039.
18. Pettersen E, *et al.* (2004) UCSF Chimera--a Visualization System for Exploratory Research and Analysis. *J Comput Chem* 25:1605-1612.
19. Sanner M, Olson A, & Spehner J (1996) Reduced Surface: an Efficient Way to Compute Molecular Surfaces. *Biopolymers* 38:305-320.