

Here some preliminary data analysis for IET in bis-peri.

I have analyzed the following parameters:

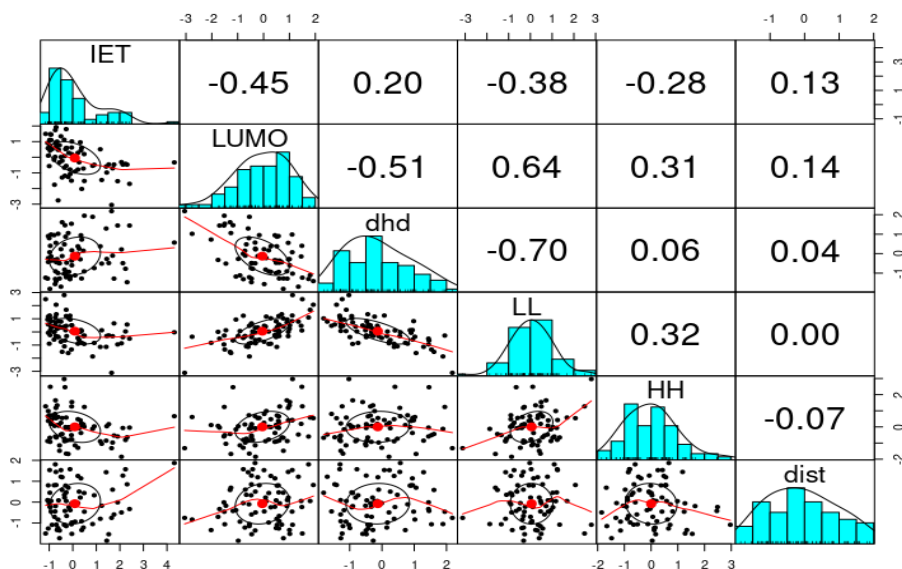
- IET – time interval for  $P_{\text{dye}}=0.3$  (70% of electronic charge is already in the TiO<sub>2</sub> cluster), femtosecond units.
- LUMO – energy of the LUMO orbital (range: -9.678 to -9.178), eV units
- dhd – dihedral angle between linkers (0 means linkers are in the plane of Pe whereas 90° means the length direction of the linkers are perpendicular)
- LL (rotation) – rotation angle of loose-linker (LL) relative to the plane of the Pe (0 means C<sub>2</sub>H<sub>2</sub> is in the Pe plane, 90° means C<sub>2</sub>H<sub>2</sub> is perpendicular to Pe plane)
- HH – distance of H atoms in neighboring linkers, angstrom units.
- dist – distance of LL COOH group to the Ti atom in the surface, angstrom units

Below, the summary of the raw data

IET	LUMO	dhd	LL	HH	dist
Min. : 5.50	Min. : -9.678	Min. : 2.956	Min. : -0.3102	Min. : 1.949	Min. : 2.100
1st Qu.: 22.25	1st Qu.: -9.448	1st Qu.: 11.820	1st Qu.: 25.4207	1st Qu.: 2.140	1st Qu.: 2.164
Median : 32.50	Median : -9.387	Median : 18.013	Median : 31.6325	Median : 2.236	Median : 2.223
Mean : 40.66	Mean : -9.392	Mean : 19.140	Mean : 31.5828	Mean : 2.244	Mean : 2.227
3rd Qu.: 50.00	3rd Qu.: -9.322	3rd Qu.: 26.415	3rd Qu.: 37.4014	3rd Qu.: 2.323	3rd Qu.: 2.274
Max. : 166.50	Max. : -9.178	Max. : 38.657	Max. : 60.2697	Max. : 2.650	Max. : 2.418

Prior to analysis the data has been centered (mean is subtracted) and rescaled (divided by variance).

Then I performed a correlation analysis, as shown in the graph.



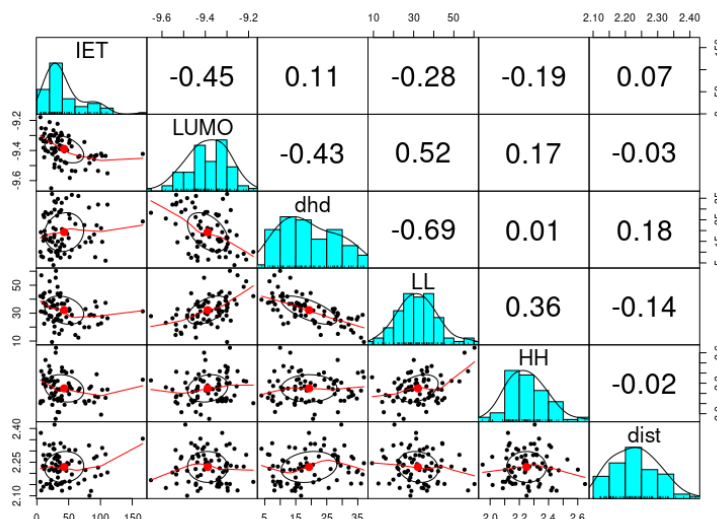
Analysing the scaled data (the raw data is shown in the nex graph, for comparison):

- IET vs LUMO  
moderate anticorrelation (-0.45), IET (↑) LUMO(↓)  
low energy LUMO (close to CB edge) => slow IET  
high energy LUMO (above CB edge) => fast IET
- IET vs dhd  
weak correlation (0.20), IET (↑) dhd(↑)  
linkers open up (separate) => slow IET  
linkers close together => fast IET
- IET vs LL(rotation)  
moderate anticorrelation (-0.38), IET (↑) LL(↓)  
C2H2 in the plane of Perylene => slow IET  
C2H2 rotate perpendicular to plane of Perylene => fast IET
- IET vs HH (distance)  
weak anticorrelation (-0.28), IET (↑) HH-dist(↓)  
HH distance between linkers *decrease* => slow IET  
HH distance between linkers *increase* => fast IET
- distance of LL COOH group to the Ti atom in the surface is weakly correlated with any other variable, therefore, we will not analyze it.

The strongest correlations, caused by steric effects:

- LUMO vs dhd  
moderate anticorrelation (-0.51), LUMO (↑) dhd(↓)  
low energy LUMO (close to CB edge, very negative) => linkers open up (separate)  
high energy LUMO (above CB edge, less negative) => linkers close together (small dihedral)

- LUMO vs LL(rotation)  
moderate-strong correlation (0.64), LUMO (↑) LL (↑)  
low energy LUMO (close to CB edge, very negative) => C2H2 in the plane of Perylene  
high energy LUMO (above CB edge, less negative) => C2H2 perpendicular to plane of Perylene
- dhd vs LL(rotation)  
strong anticorrelation (-0.7), dhd (↑) LL (↓)  
linkers close together (small dihedral) => C2H2 perpendicular to plane of Perylene  
linkers open up (separate) => C2H2 in the plane of Perylene



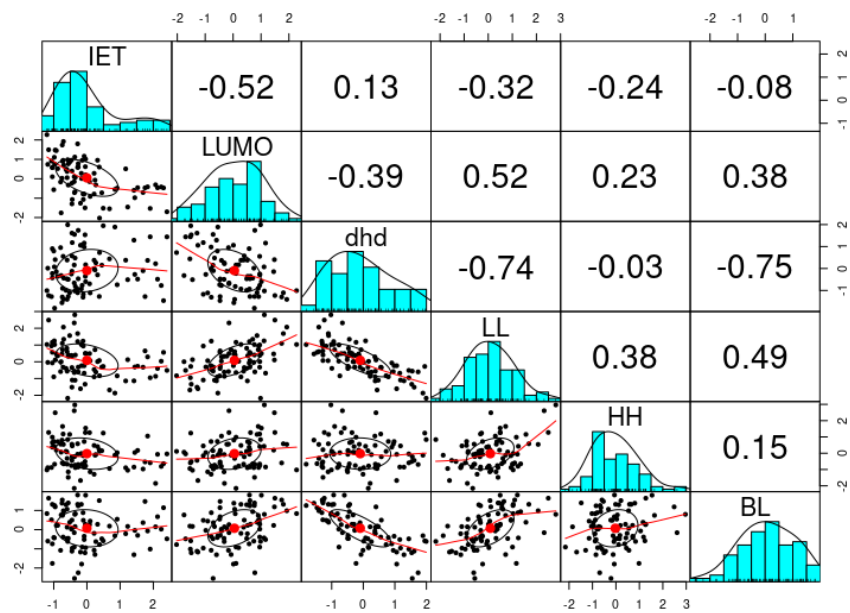
In a second set of data I replaced the *dist* parameter (distance between LL-COOH and TiO<sub>2</sub> surface) by the BL(rotation) parameter: – rotation angle of bound-linker (BL) relative to the plane of the Pe (0 means C2H2 is in the Pe plane, 90° means C2H2 is perpendicular to Pe plane).

The results are shown next. In summary the correlations remain approximately the same. There is no significant correlation between IET and BL(rotation).

The correlations due to steric interactions between linkers and perylene remain strong for BL as well.

The strongest correlations so far involving IET times are associated with the LUMO position, namely, higher LUMO faster IET and lower LUMO (close to CB edge) slower IET.

The correlations are plotted in the next panel for scaled data.



## Principal Components Analysis

For the training data generated from the raw data we have the following table.

Standard deviations (1, ..., p=7):

1.6570975	1.1304489	1.0095744	0.9696017	0.6939062	0.6594892	0.3167231
-----------	-----------	-----------	-----------	-----------	-----------	-----------

PC loadings:

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
IET	0.23483832	<b>0.5248208</b>	<b>-0.482125272</b>	0.35938010	-0.32046576	0.45150569	-0.03549042
LUMO	<b>-0.44912105</b>	-0.2017889	0.003563817	-0.36487722	0.04041131	0.78393574	-0.09074747
dhd-dihedral	<b>0.51889423</b>	-0.3537977	0.050251876	0.04365521	0.12108309	0.30154595	0.70395586
LL-rotation	<b>-0.51185265</b>	-0.1127649	-0.015590612	0.14607897	-0.65698795	-0.16298231	0.49549209
HH-distance	-0.14552245	<b>-0.5365751</b>	-0.077976812	<b>0.76191479</b>	0.14944735	0.09450353	-0.27027840
BL-rotation	<b>-0.43572979</b>	<b>0.4020773</b>	-0.101102504	0.21560892	0.64435233	-0.03001949	0.41913412
dist (COOH-Ti)	-0.02931981	-0.3104459	<b>-0.865142823</b>	-0.29574374	0.10885728	-0.23294223	0.02674407

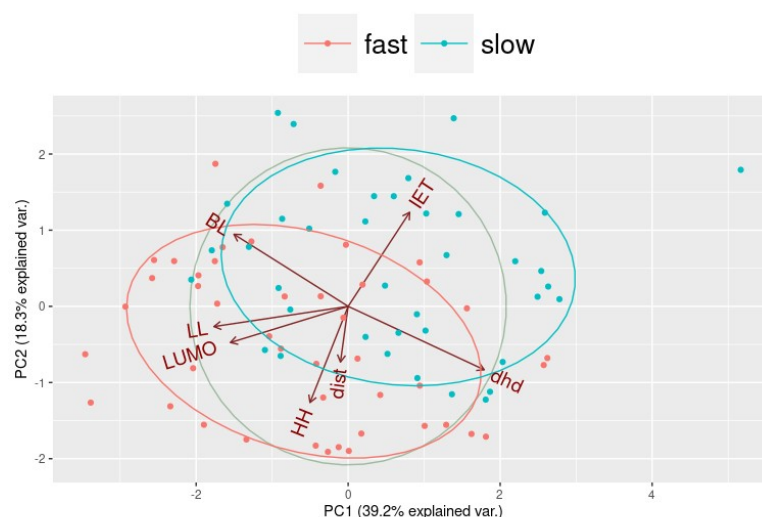
Importance of components:

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Standard deviation	1.6571	1.1304	1.0096	0.9696	0.69391	0.65949	0.31672
Proportion of Variance	0.3923	0.1826	0.1456	0.1343	0.06879	0.06213	0.01433
Cumulative Proportion	0.3923	0.5748	0.7205	0.8548	0.92354	0.98567	1.00000

In summary:

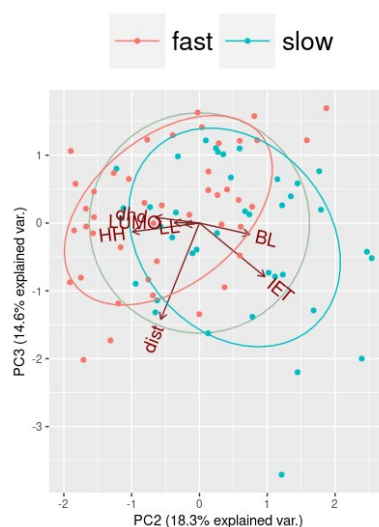
- the features with highest correlation are highlighted, up to PC4 that comprises 85% of the variability of the data.
- PC1 contains mainly the steric effects of the structure and the LUMO energy; see to bi-plots.
- PC2 shows some prominence of IET time and its relation with HH-distance (H atoms belonging to different linkers) and BL (rotation angle of bound linker). Slow IET is associated with HH proximity and rotation of BL perpendicularly to Perylene plane; and vice-versa.
- PC3 also shows relevance of IET time and strong correlation with distance between LL-COOH and the surface (that is, LL-anchor and surface). That means that IET becomes faster as LL-anchor approaches the surface. This is an indication that some IET may be occurring through LL. All the other features are irrelevant here.
- PC4 shows loadings that are somewhat contradictory to PC2 and PC3, but in this case IET has less relevance. So I believe we should give precedence to PC2 and PC3.

In the following we show PC biplots that illustrate the data above. We divided the data into two categories, using the median of the IET times distribution, 32.5 fs, as a criterium: the entries with IET < 32.5 fs are labeled “fast” and those with IET > 32.5 fs are labeled slow.



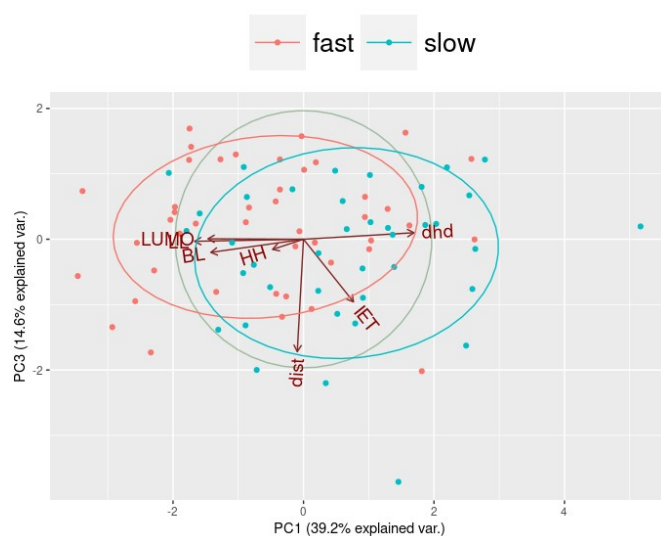
Biplot of PC1 and PC2, ellipses have 68% of data. Fast corresponds to IET < 32.5 fs.

PC1 and PC2 account for most of the variability. The above plot shows that IET and the dihedral angle dhd are quite uncorrelated to each other. On the other hand we have the steric effects (HH-dist, COOH-Ti dist, and LL-rotation) strongly correlated with the LUMO energy, that is, LUMO is higher (above the CB edge) as the aforementioned variables increase, thereby decreasing the conjugation of the molecular system. On the other hand, IET becomes slower (bigger injection times) when that happens. The variables BL-rotation and dhd-dihedral are totally uncorrelated to the process.



Biplot of PC2 and PC3, ellipses have 68% of data. Fast corresponds to IET < 32.5 fs.

Bi-plot of PC2 and PC3 loadings corroborates the PC1-PC2 biplot. Furthermore, it shows a correlation between IET times and BL-rotation, which means that IET becomes slow if BL rotates out of the plane, as expected. In particular, the slow IET outliers show pronounced BL-rotations. Recall that IET is prominent in both PC2 and PC3 PC loadings, however, this plot is not very enlightening. Slow outliers have high BL character.



Biplot of PC1 and PC3, ellipses have 68% of data. Fast corresponds to IET < 32.5 fs.

Bi-plot of PC1 and PC3 loadings also corroborates the PC1-PC2 biplot. The main feature here is the moderate correlation between IET, COOH-Ti distance and dhd-dihedral. As these geometrical parameters increase the IET becomes slower, indicating that the loose linker may contribute to the IET. Slow outliers have high (dist,dhd) character.