Electronic Supplementary Material (ESI) for CrystEngComm.

The Interplay among Molecular Structures, Crystal Symmetries and Lattice Energy Landscapes Revealed by Unsupervised Machine Learning: A Closer Look at Pyrrole Azaphenacenes

Supporting Information

Jack Yang,*,†,‡ Nathan Li,† and Sean Li†

†*Advanced Materials and Manufacturing Futures Institute, School of Material Science and Engineering, University of New South Wales, Sydney, New South Wales 2052, Australia*

‡*Australian Nuclear Science and Technology Organization, New Illawarra Rd, Lucas Heights, New South Wales, 2234, Australia.*

E-mail: jianliang.yang1@unsw.edu.au

Sketch–Maps

Table 1

Molecule	ΔE_{latt} (kJ/mol)	Number of Structures
1	10	774
2	20	202
3	10	475
4	25	887
5	20	627
6	15	1144
7	40	310
8	25	923
9	15	364
10	15	703
11	15	1349
12	15	635
13	15	1174
14	15	844
15	15	1395
16	15	1574
17	15	513
18	15	686
19	15	855
20	15	451
21	15	1266
22	15	782
23	15	306
24	15	1195

Molecule 1

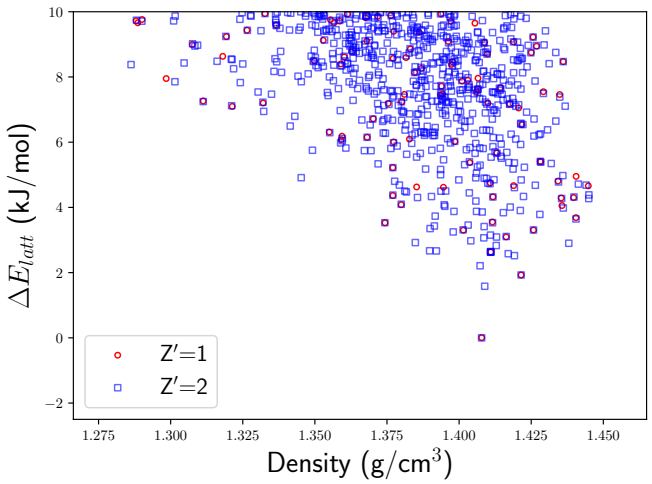

Figure 1: Energy–density plot for the predicted crystal structures for molecule **1** in both $Z' = 1$ and $Z' = 2$ settings.

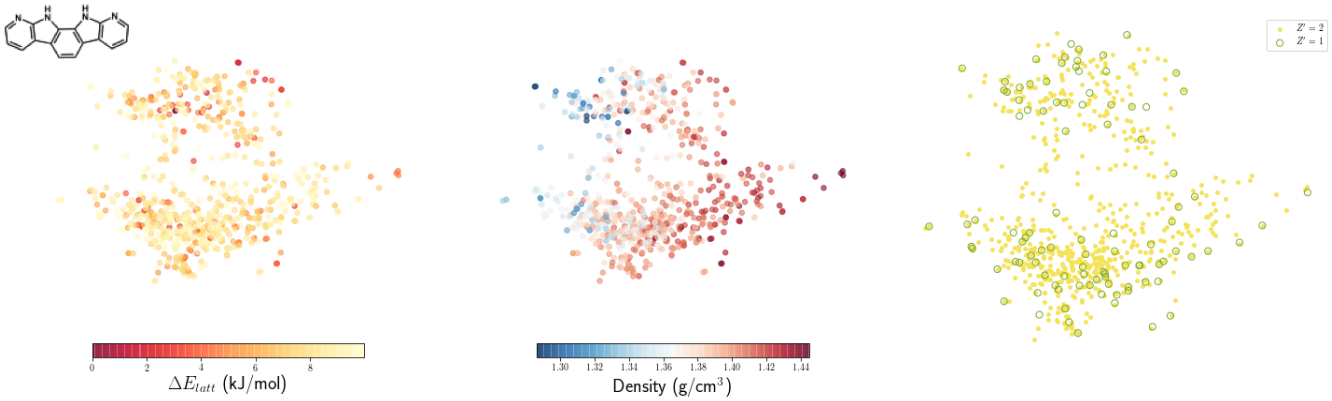

Figure 2: Sketch–map representation of the lattice for molecule **1**, with each point colour–coded according to (left) lattice energy difference to the global minimum, (middle) density and (right) Z' number. SOAP–REMatch kernel were calculated with $r_c = 3$ Å.

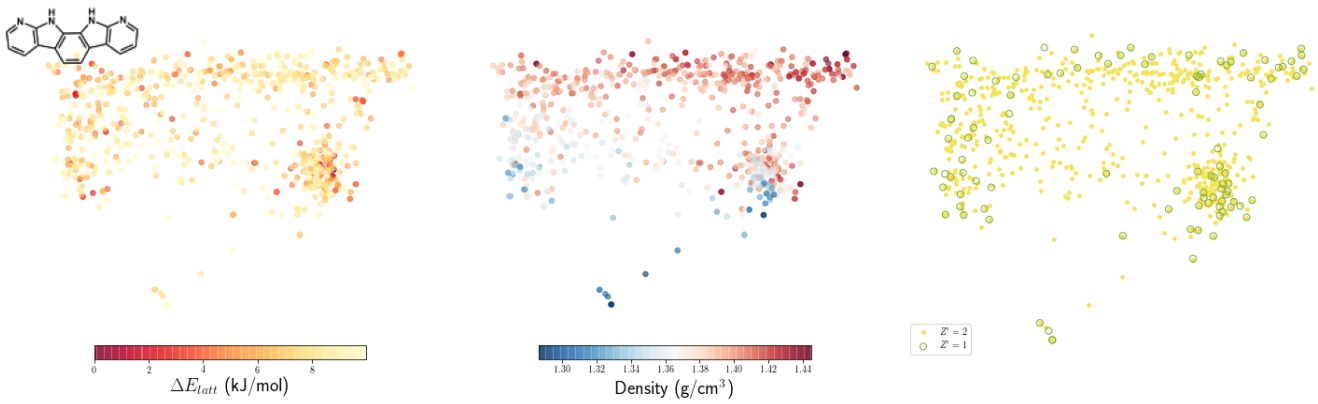

Figure 3: Sketch–map representation of the lattice for molecule **1**. SOAP–REMatch kernel were calculated with $r_c = 5$ Å.

Molecule 2

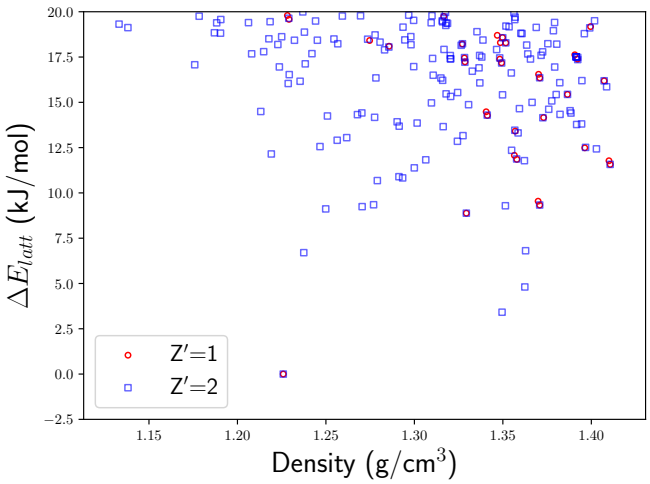

Figure 4: Energy–density plot for the predicted crystal structures for molecule **2** in both $Z' = 1$ and $Z' = 2$ settings.

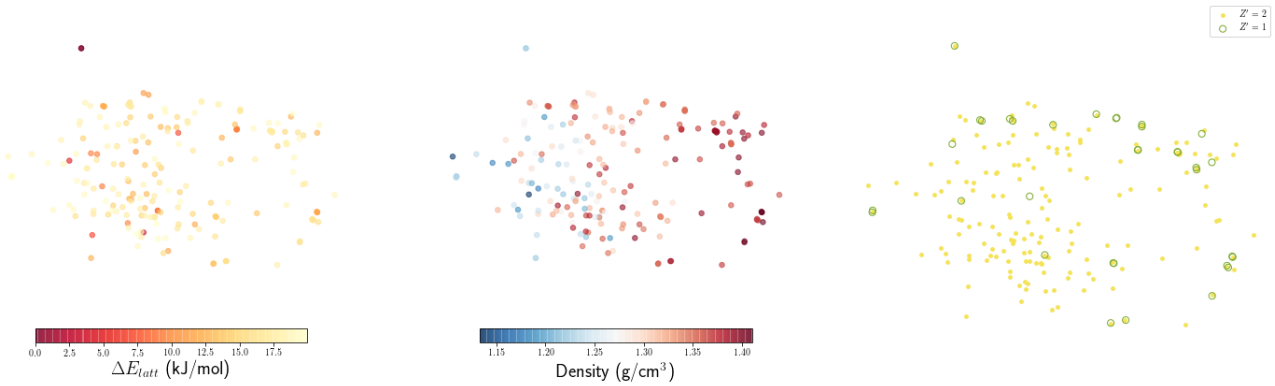

Figure 5: Sketch–map representation of the crystal packing landscapes for molecule **2**, with $r_c = 3$ Å.

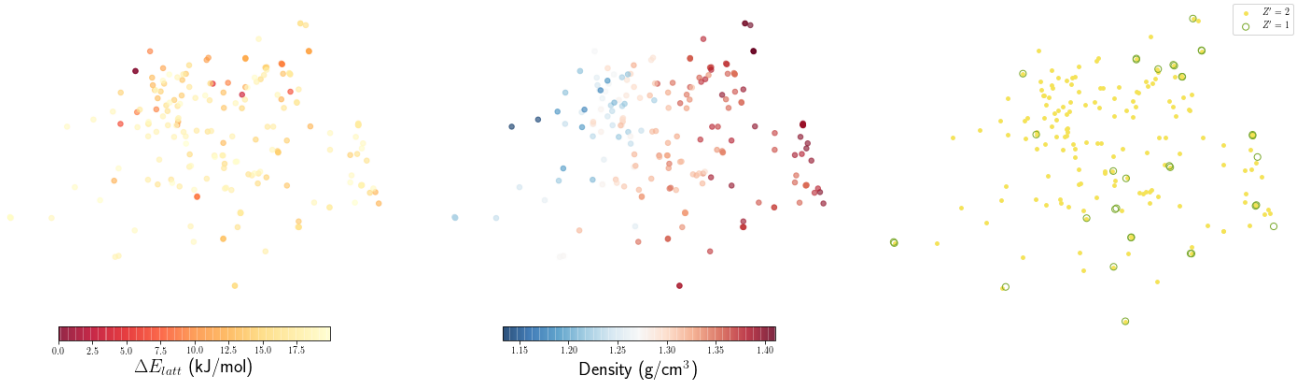

Figure 6: Sketch–map representation of the crystal packing landscapes for molecule **2**, with $r_c = 5$ Å

.

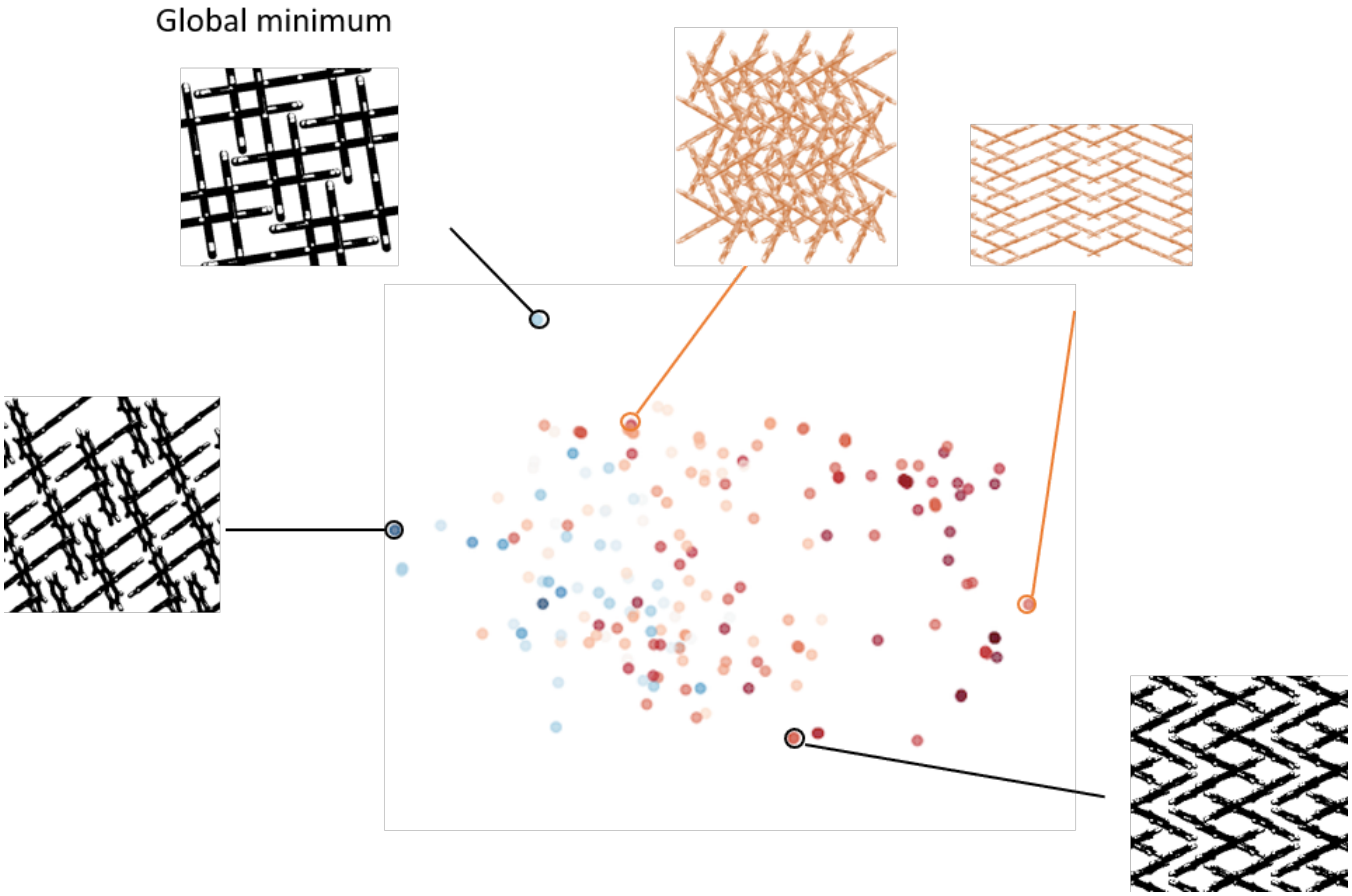

Figure 7: Sketch–map representation of the crystal packing landscapes for molecule **4** with SOAP–REMatch kernel calculated using $r_c = 3$ Å. Each point is colour–coded according to the density of the crystal [identical to Fig. 5 (middle)].

Molecule 3

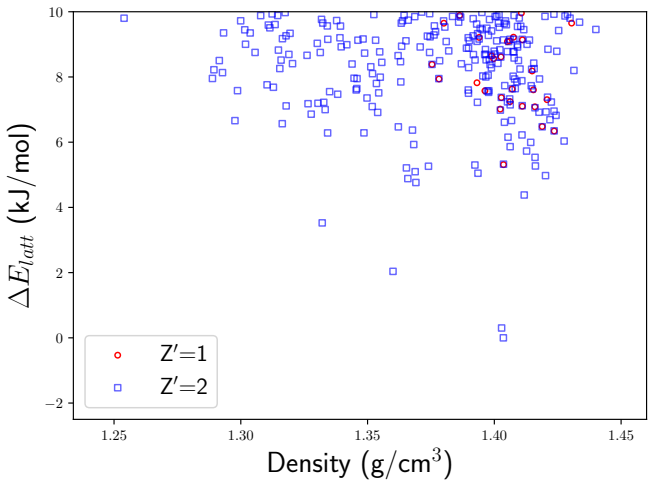

Figure 8: Energy–density plot for the predicted crystal structures for molecule **3** in both $Z' = 1$ and $Z' = 2$ settings.

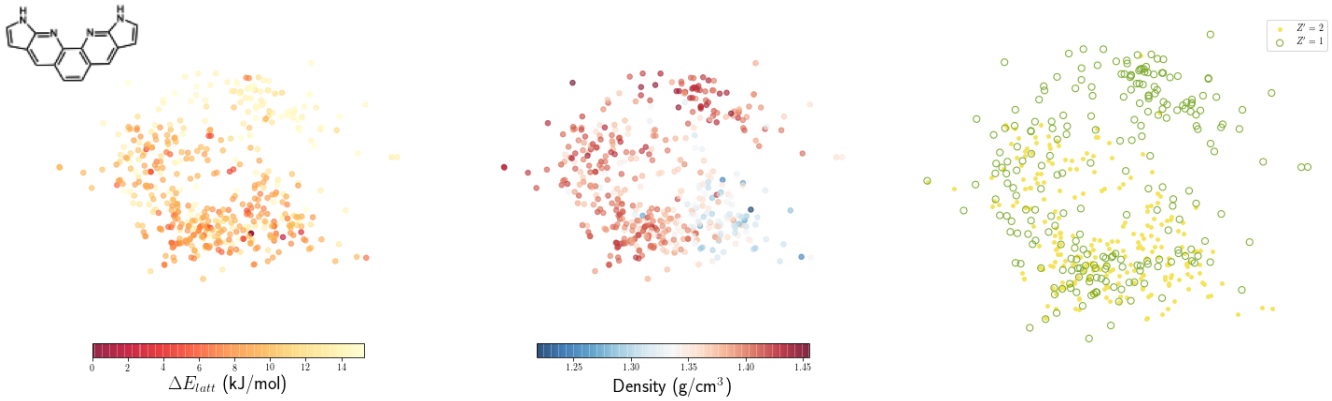

Figure 9: Sketch–map representation of the crystal packing landscapes for molecule **3** with $r_c = 3$ Å

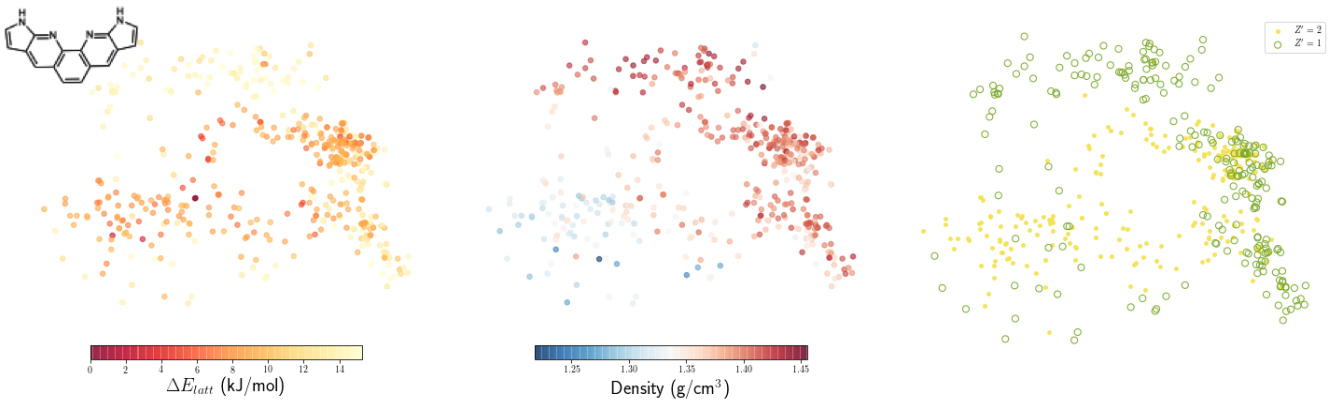

Figure 10: Sketch–map representation of the crystal packing landscapes for molecule **3** with $r_c = 5$ Å

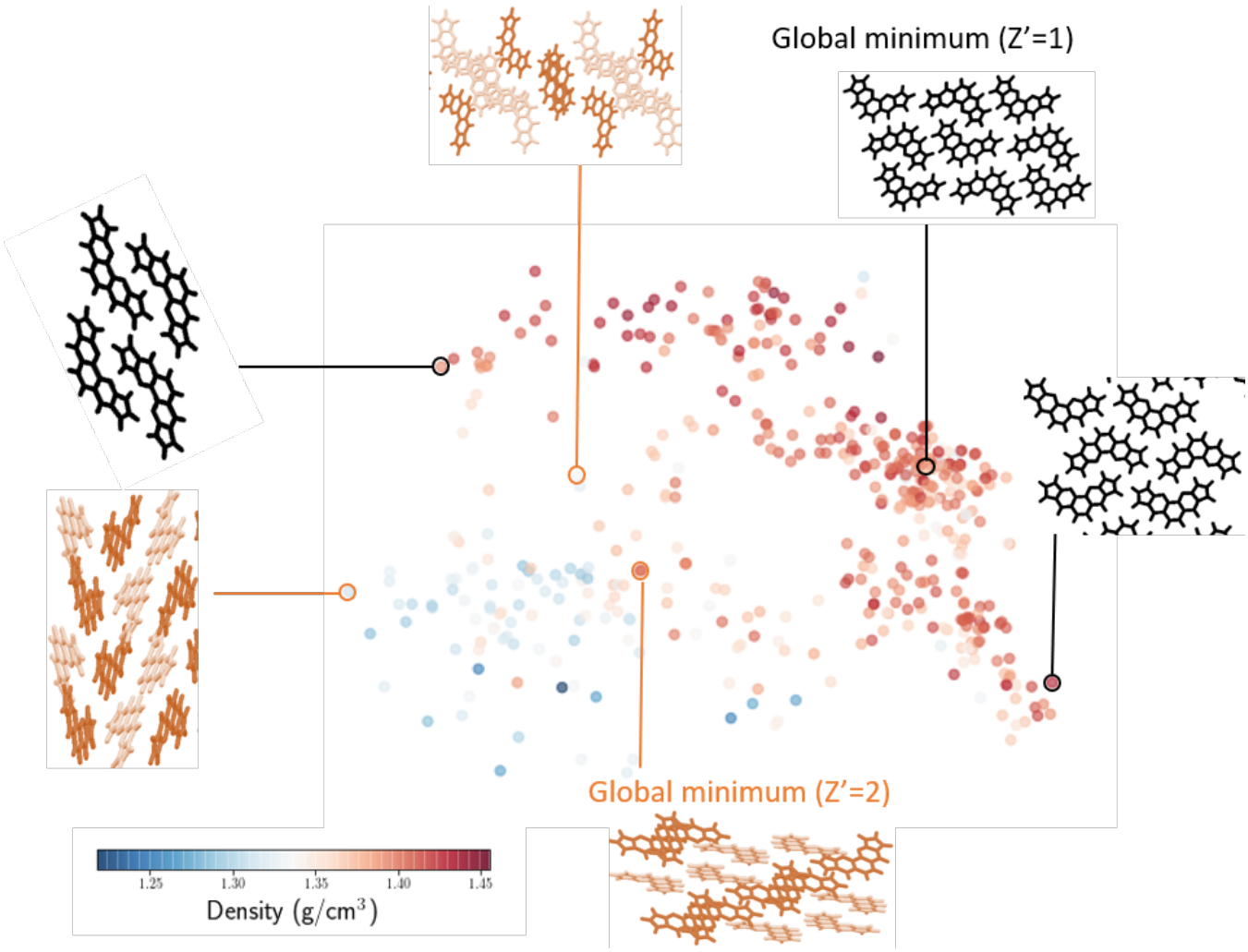

Figure 11: Sketch–map representation of the crystal packing landscapes for molecule **9** with SOAP–REMatch kernel calculated using $r_c = 5$ Å. Each point is colour–coded according to the density of the crystal [identical to Fig. 10 (middle)]. Structure(s) that are only found in the $Z' = 1$ (2) searches are coloured in black (orange), whereas structures coloured in blue are found from both $Z' = 1$ and 2 searches. For structures coloured in orange, the dark and light colouring corresponds to two symmetry unrelated molecules in the unit cell.

Molecule 4

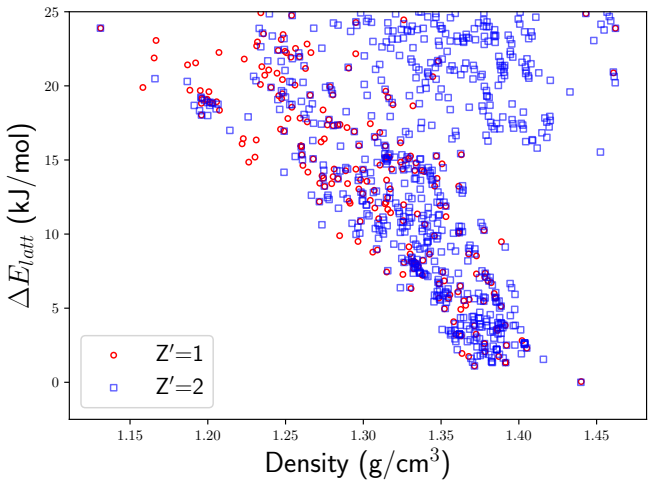

Figure 12: Energy–density plot for the predicted crystal structures for molecule **4** in both $Z' = 1$ and $Z' = 2$ settings.

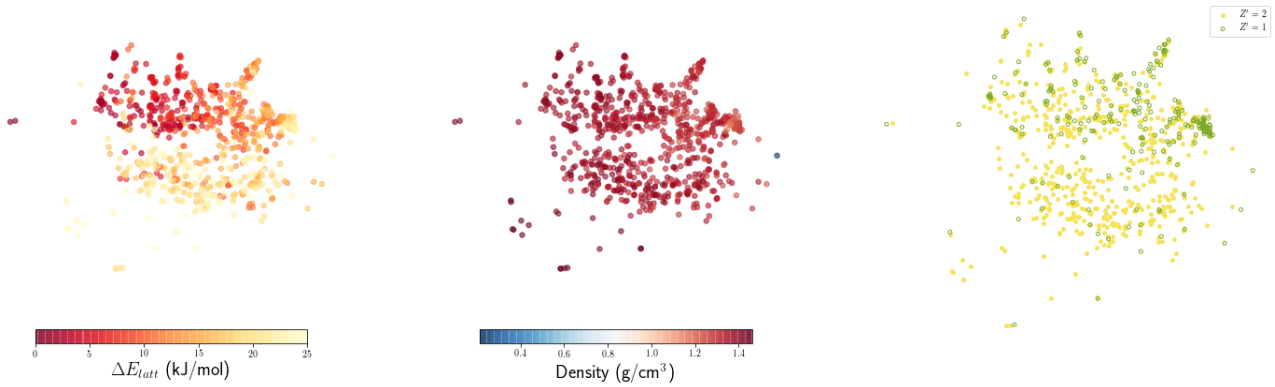

Figure 13: Sketch–map representation of the crystal packing landscapes for molecule **4** with $r_c = 3$ Å.

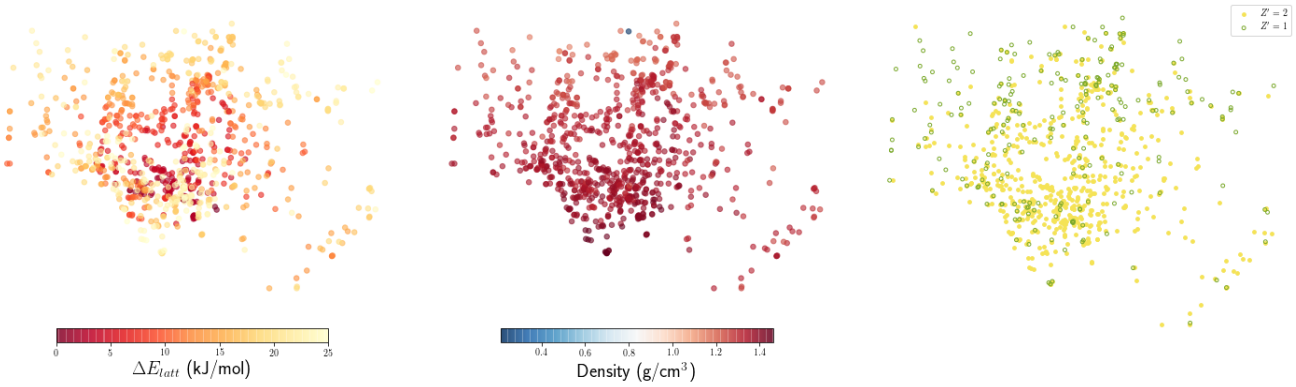

Figure 14: Sketch–map representation of the crystal packing landscapes for molecule **4** with $r_c = 5$ Å.

Molecule 5

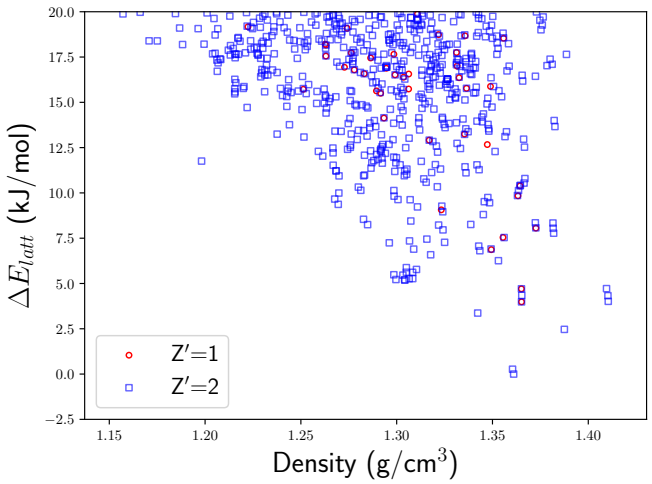

Figure 15: Energy–density plot for the predicted crystal structures for molecule **5** in both $Z' = 1$ and $Z' = 2$ settings.

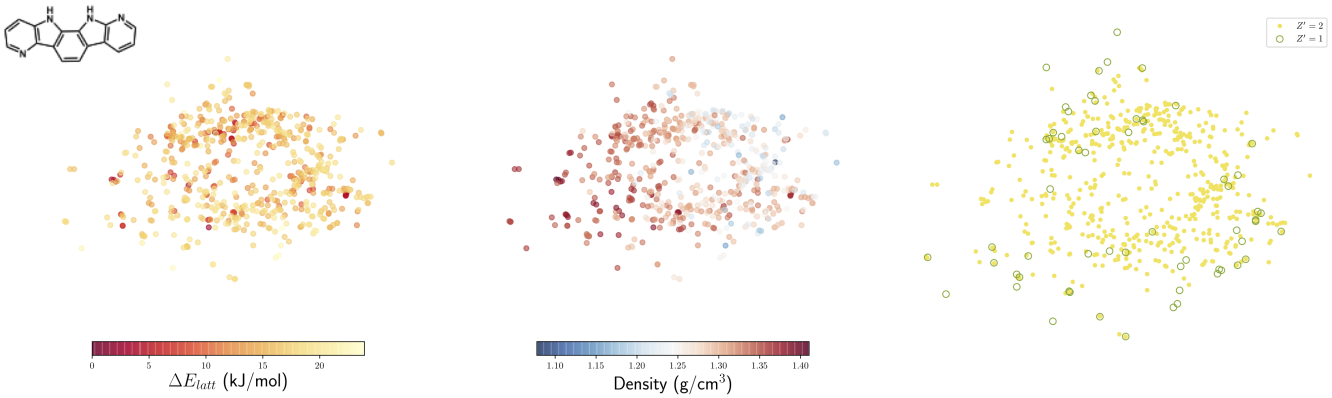

Figure 16: Sketch–map representation of the crystal packing landscapes for molecule **5** with $r_c = 3$ Å.

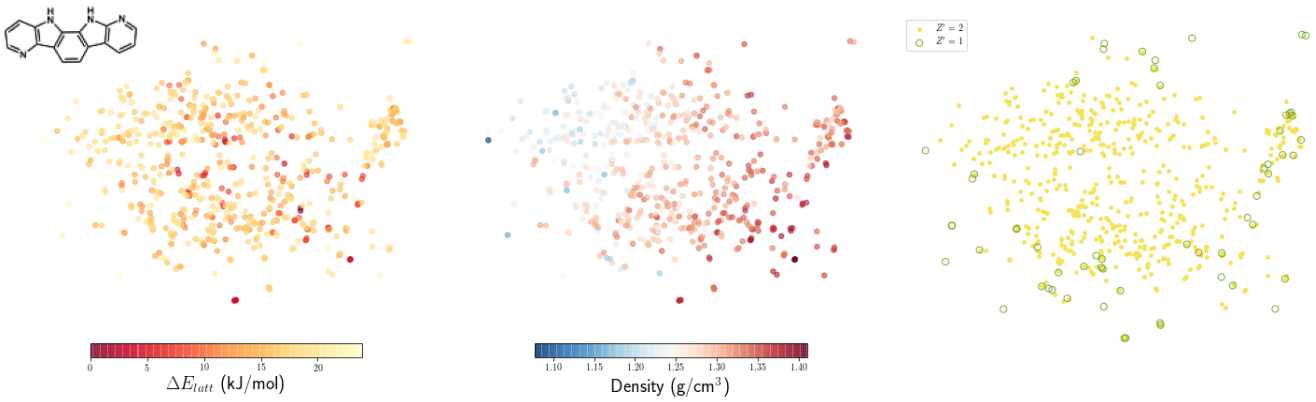

Figure 17: Sketch–map representation of the crystal packing landscapes for molecule **5** with $r_c = 5$ Å.

Molecule 6

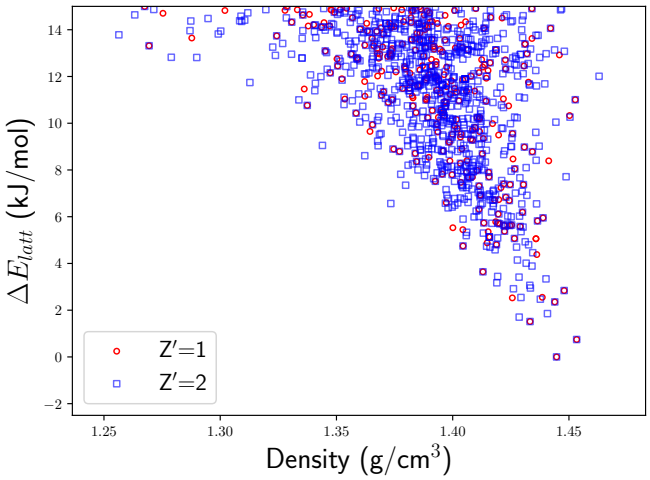

Figure 18: Energy–density plot for the predicted crystal structures for molecule **6** in both $Z' = 1$ and $Z' = 2$ settings.

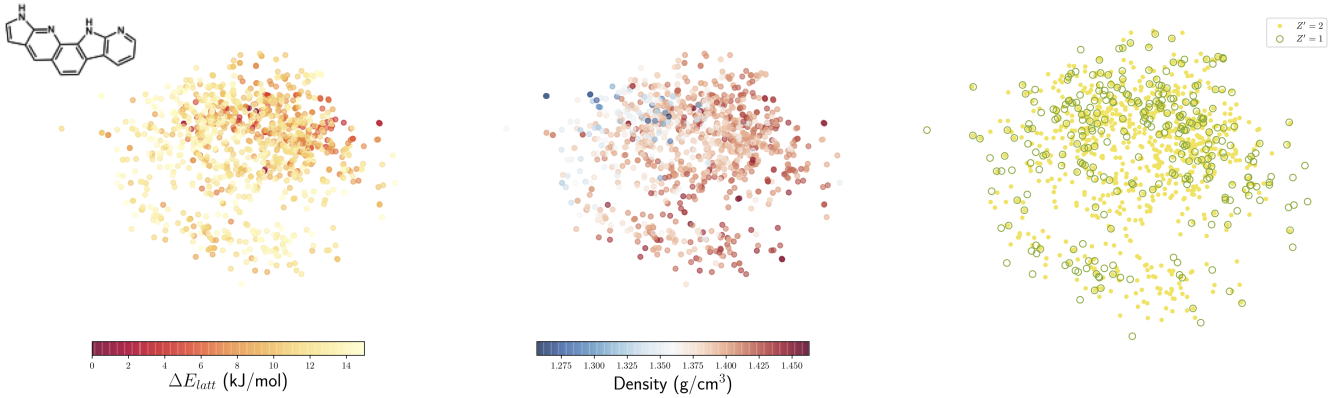

Figure 19: Sketch–map representation of the crystal packing landscapes for molecule **6** with $r_c = 3$ Å.

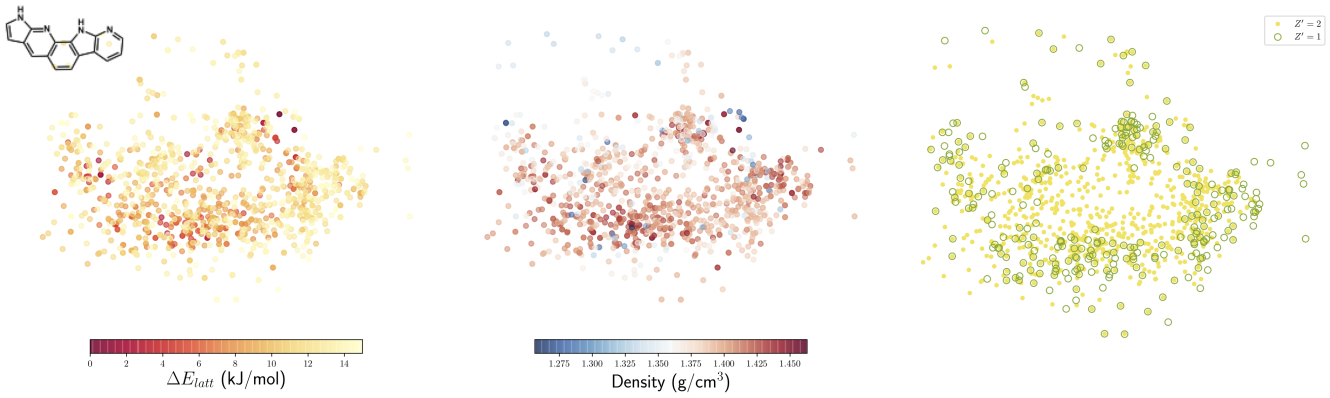

Figure 20: Sketch–map representation of the crystal packing landscapes for molecule **6** with $r_c = 5$ Å.

Molecule 7

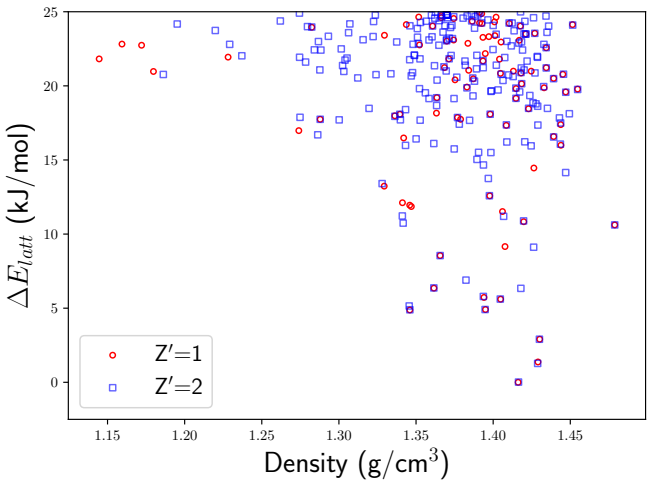

Figure 21: Energy–density plot for the predicted crystal structures for molecule **7** in both $Z' = 1$ and $Z' = 2$ settings.

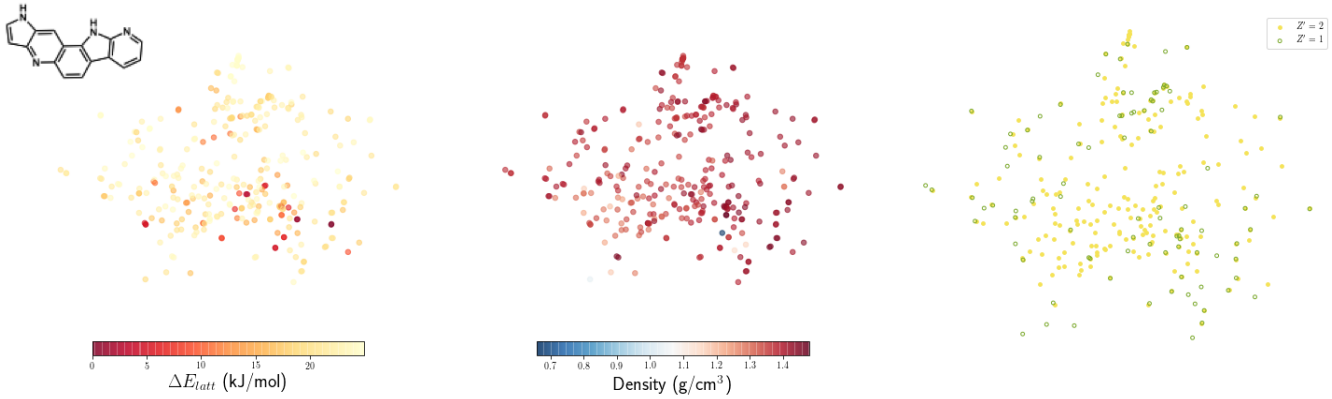

Figure 22: Sketch–map representation of the crystal packing landscapes for molecule **7** with $r_c = 3$ Å.

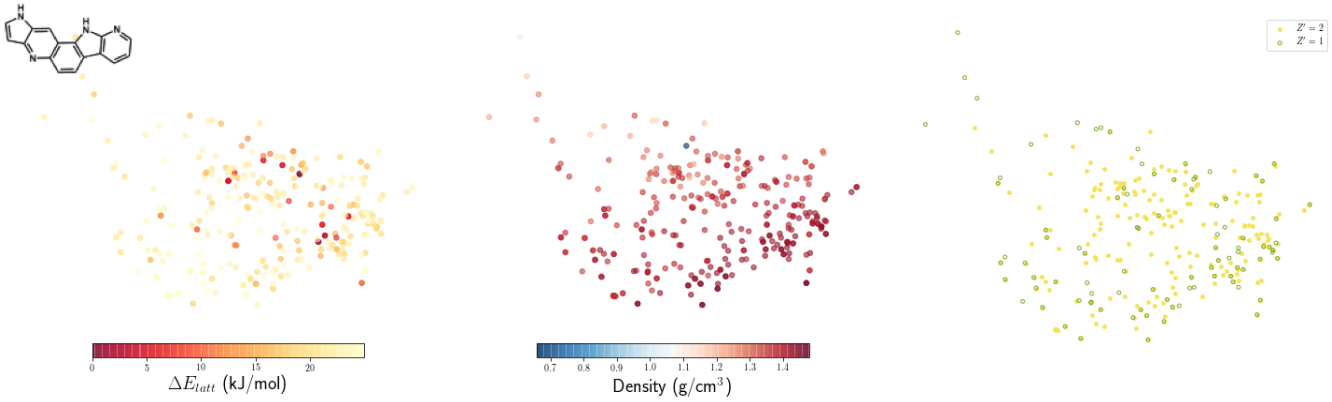

Figure 23: Sketch–map representation of the crystal packing landscapes for molecule **7** with $r_c = 5$ Å.

Molecule 8

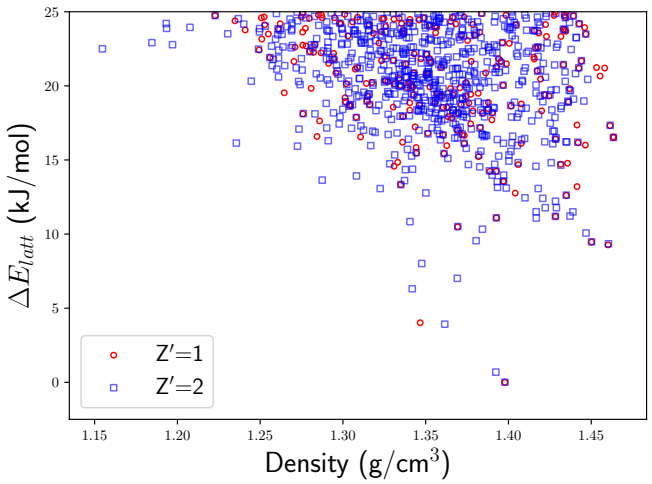

Figure 24: Energy–density plot for the predicted crystal structures for molecule **8** in both $Z' = 1$ and $Z' = 2$ settings.

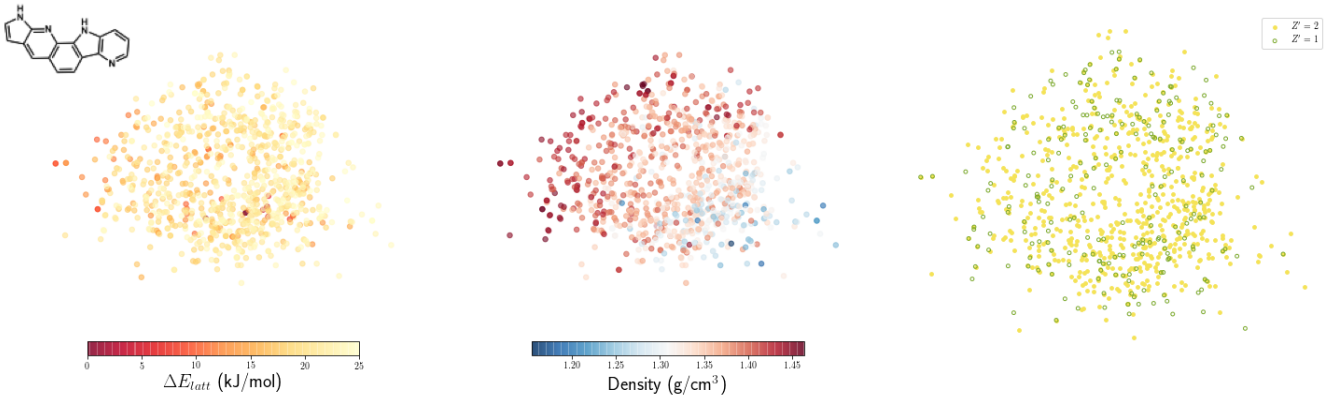

Figure 25: Sketch–map representation of the crystal packing landscapes for molecule **8** with $r_c = 3$ Å.

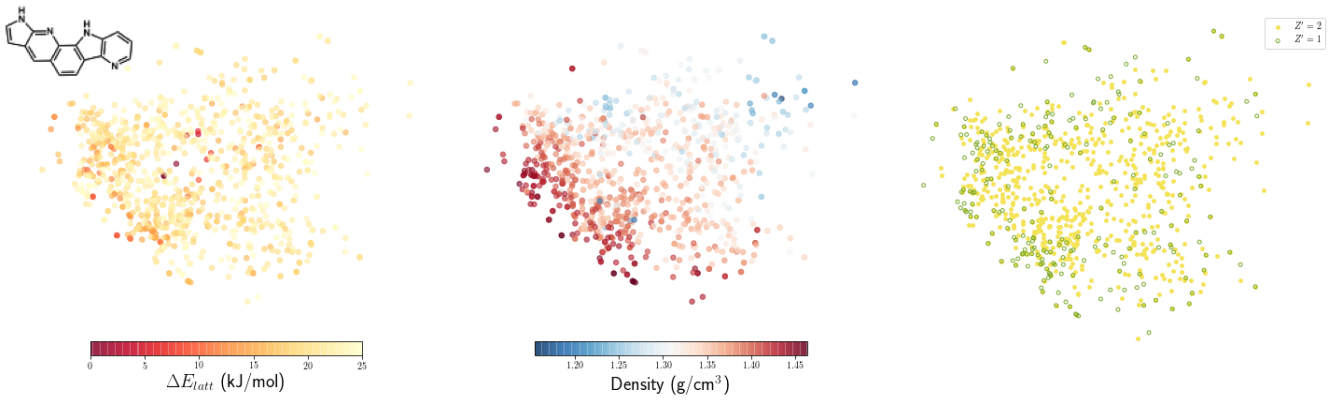

Figure 26: Sketch–map representation of the crystal packing landscapes for molecule **8** with $r_c = 5$ Å.

Halogen Substituted Pyrrole Azaphenacenes

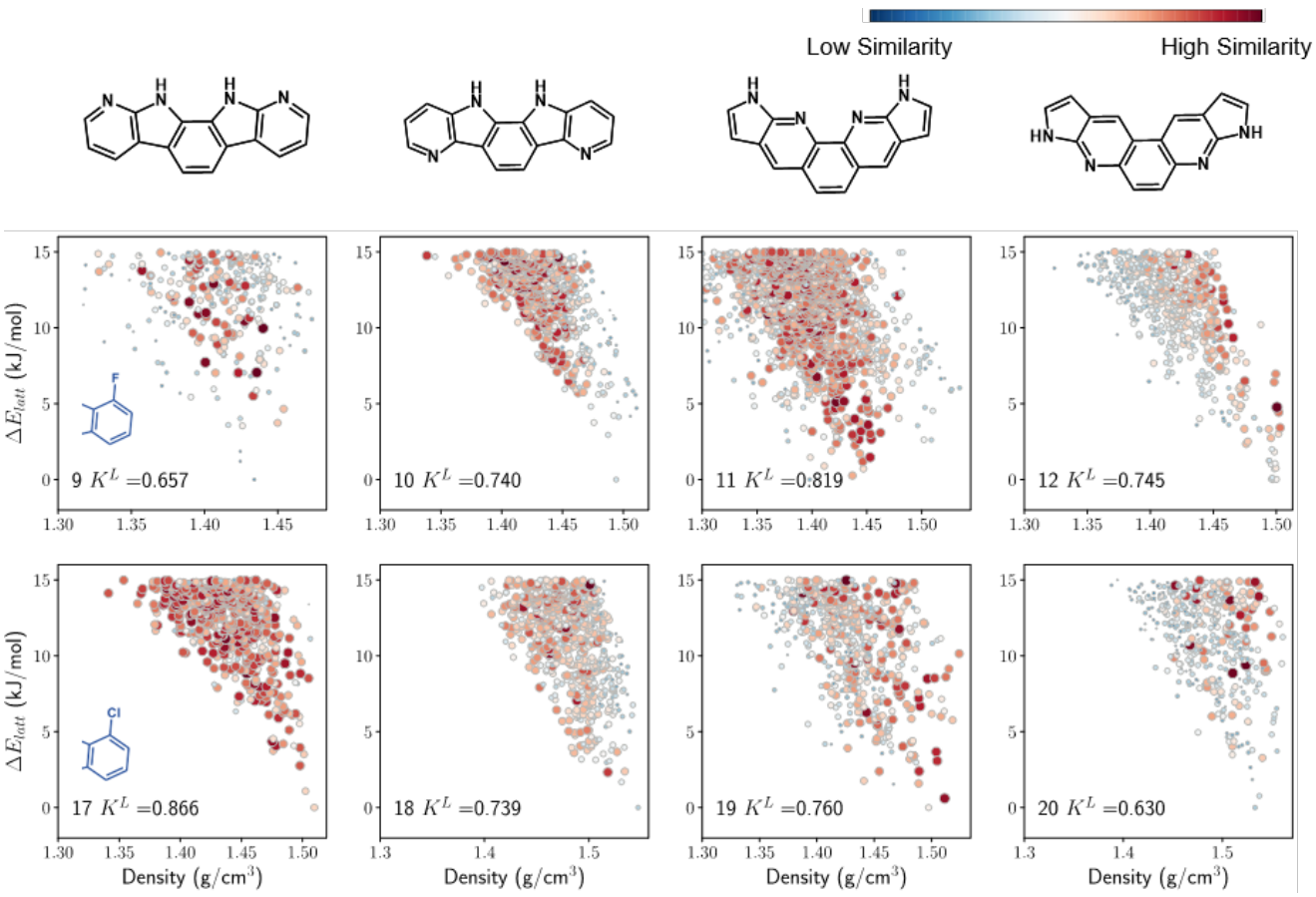

Figure 27: Energy–density plots of halogen–substituted symmetric pyrrole azaphenacenes, with each point sized and color–coded according to the normalized similarity scores with respect to the reference crystal structures [the global minimum of unsubstituted pyrrole azaphenacenes (top row) in $Z' = 1$ settings]. The SOAP–REMatch kernels are calculated with $r_c = 3$ Å and $\gamma = 1$ to highlight the similarities in short–range atomic environments.

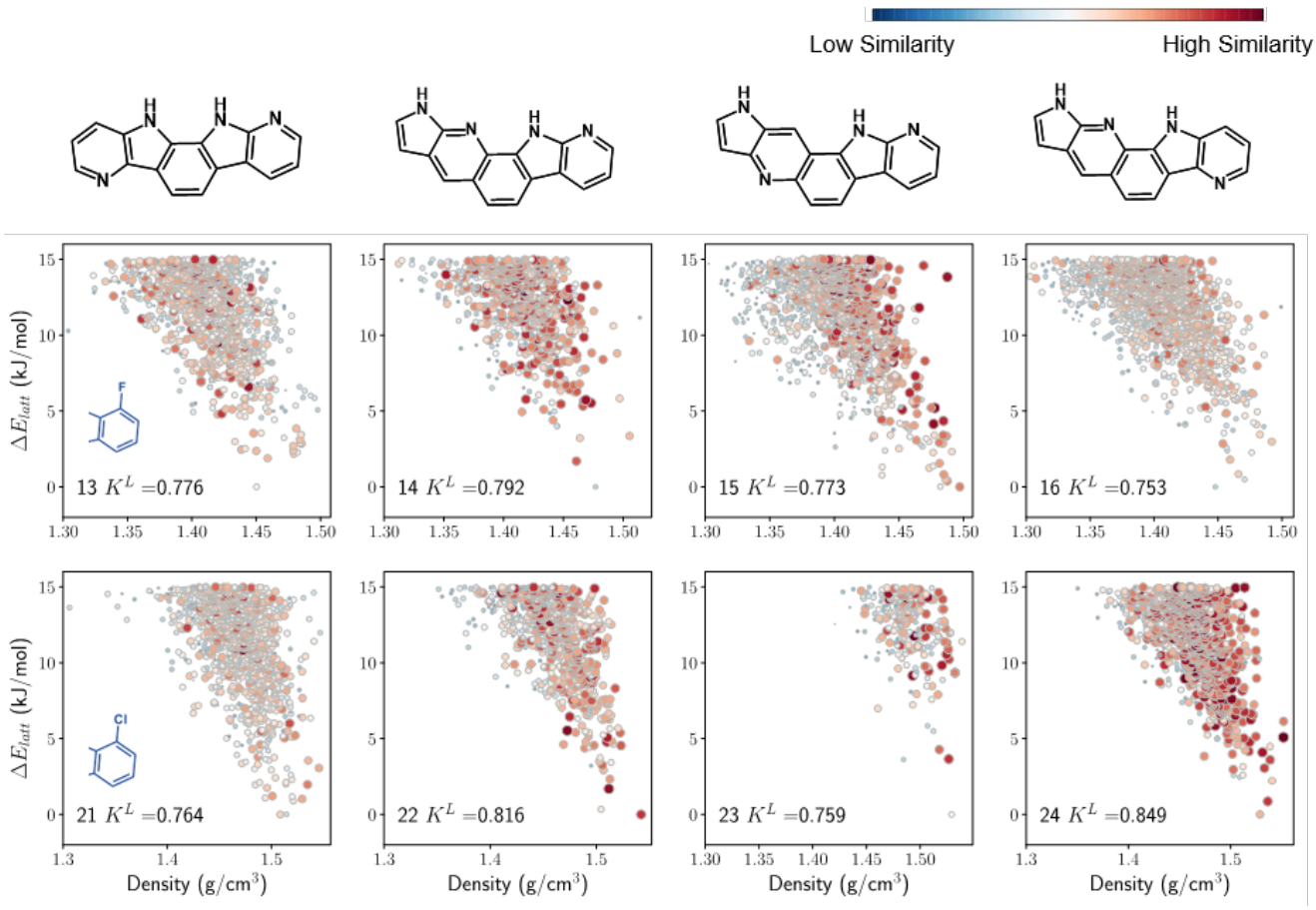

Figure 28: Energy–density plots of halogen–substituted asymmetric pyrrole azaphenacenes, with each point sized and color–coded according to the normalized similarity scores with respect to the reference crystal structures [the global minimum of unsubstituted pyrrole azaphenacenes (top row) in $Z' = 1$ settings]. The SOAP–REMatch kernels are calculated with $r_c = 3$ Å and $\gamma = 1$ to highlight the similarities in short–range atomic environments.