Aromatics Analyser 2

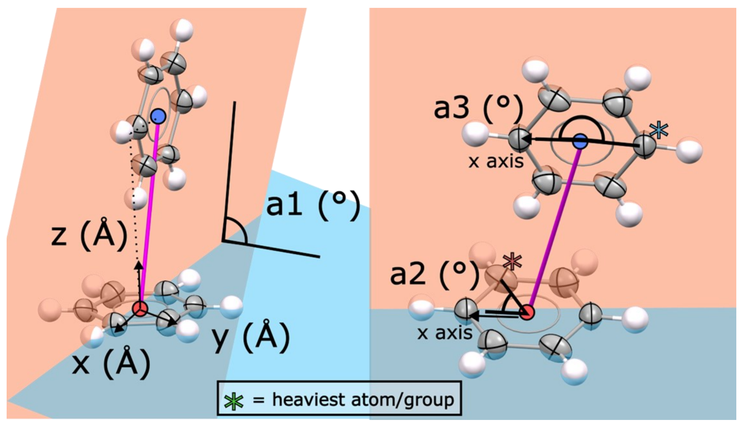
Christopher J. Kingsbury,* Patrick McCabe and Isaac Sugden; ckingsbury@ccdc.cam.ac.uk

The Cambridge Crystallographic Data Centre, Cambridge, CB2 1EZ, UK

Analysis of aromatic interactions is often necessary for understanding why a crystal has the form and properties it does. Aromatic groups in crystals mostly arrange into T-shape and offset stack arrangements; benzene, pyridine, thiophene and other rings tend to form geometrically similar arrangements. Aromatic interactions can be structure directing in the absence of other forces, and stable networks of these interactions can be formed. When multiple polymorphs exist, these interactions can be a hallmark of the first-formed structures. Across the CSD, aromatics can be observed to form preferential arrangements dependent on chemistry. We present “Aromatics Analyser 2” for analysis of these aromatic-specific interactions, and evidence from use of these subroutines to support the above assertions.

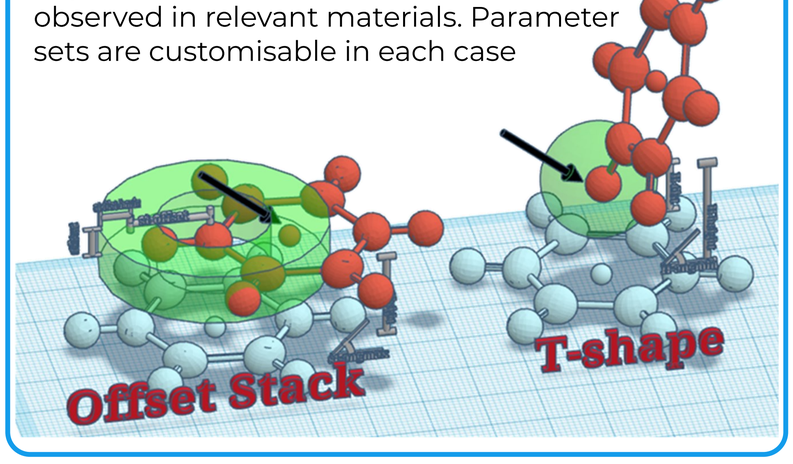
How it works—the short version

Rings are converted to a coordinate system with the six degrees of freedom of the relative orientation of two rigid bodies. The position of the second ring (x, y, z) relative to the first (at 0,0,0), the angle between these rings (a_1) and their rotation in-plane (a_2 and a_3 respectively)



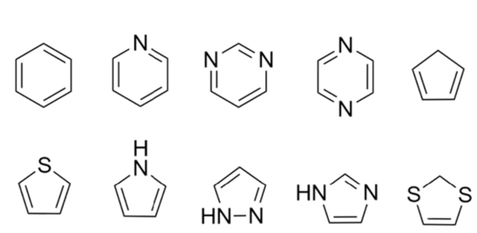
Interaction types

Interactions with expected geometry, generally the common **Offset Stack** and **T-shape**, are reported. Direct Stack, Thiophene-S···ϕ, tolyl-CH···ϕ and cation···ϕ interactions are observed in relevant materials. Parameter sets are customisable in each case



Scope

Many common rings and cation interaction partners are already included. Fused rings are calculated separately. New aromatic types can be added as shown in the scripts attached.



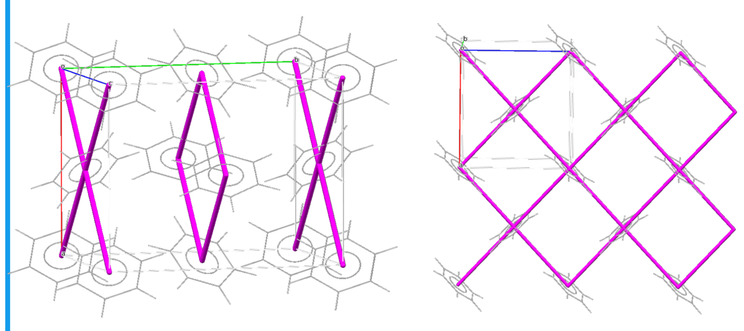
Data

Interaction type and distance, geometry, positions, symmetry, strength category, substitution, propensity, topology, ring names. All accessible in Mercury, notebook and command line interfaces.

Examples from the CSD

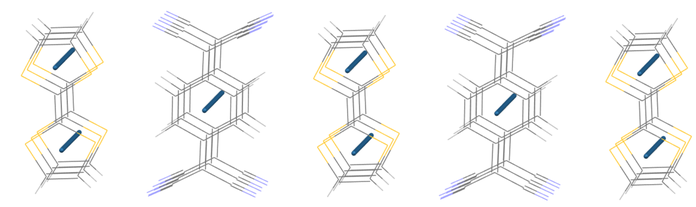
benzene (BENZEN06)

Forms a 2D grid of interactions



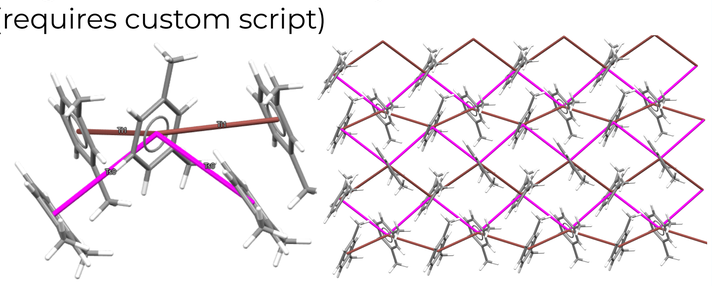
TTF-TCNQ (TTFTCQ04)

One-dimensional offset stacks



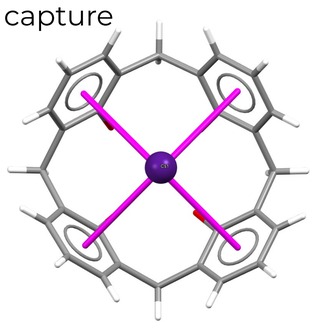
m-xylene (ZZZSPY01)

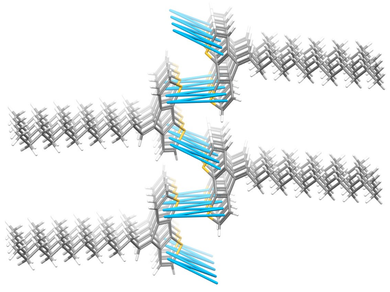
2D grid of T-shaped and tolyl interactions (requires custom script)



calix[4]arene-Cs (BASTOF)

Cation-pi interactions for radionuclide (^{137}Cs) capture



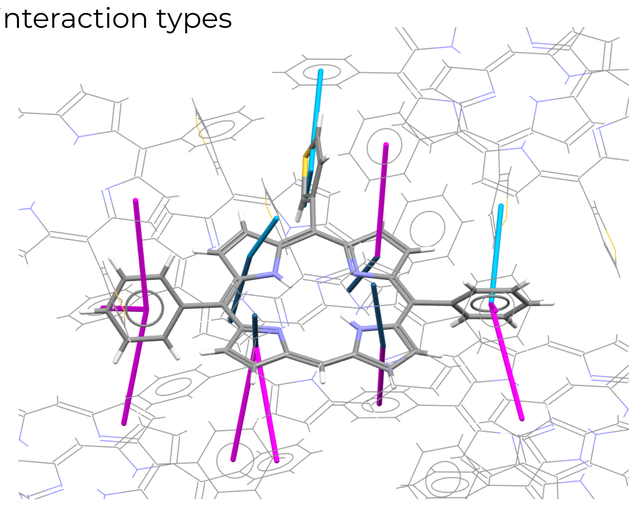


dodecyl-cyclopenta-dithiophene (ALAPIM)

Thiophene-S-aryl interactions

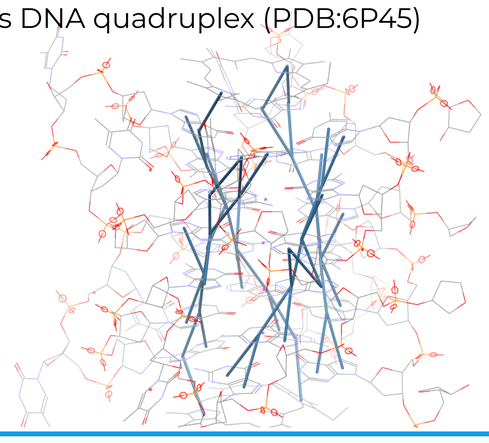
thiopheneyldiphenylporphyrin (TUSREH)

Visualising a network from many different interaction types



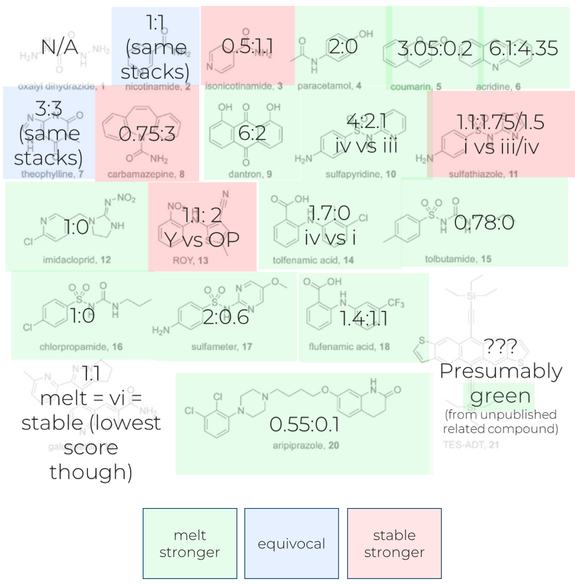
Analysis of biomolecules.

AA2 can be used in Hermes to understand intra– and inter-molecular interactions, such as in this DNA quadruplex (PDB:6P45)



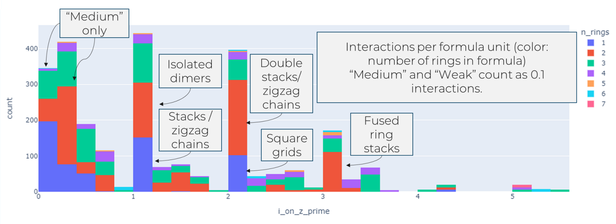
Comparing polymorphs —

“from melt” vs “stable” compounds can be compared by looking at the number of identified “strong” aromatic interactions per Z’



In the presence of hydrogen bonding pairs, metastable polymorphs can display more “strong” aromatic interactions than stable polymorphs **(12:2:4).** This behaviour is inverted in a sample of compounds without OH and NH moieties present **(3:7:8)**. Molecular compounds can optimise only a limited number of interactions; aromatic interactions may be a signifier of suboptimal hydrogen bonding

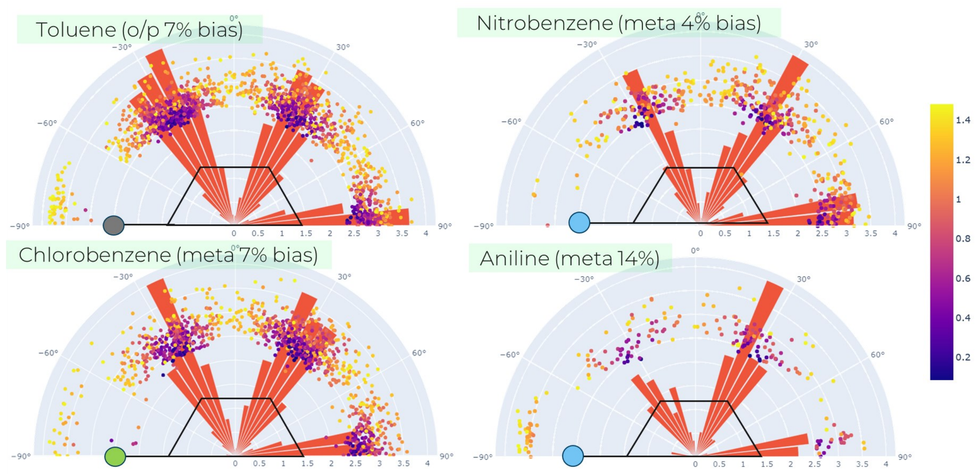
Interactions per Z’ in the CSD drug-single-component subset



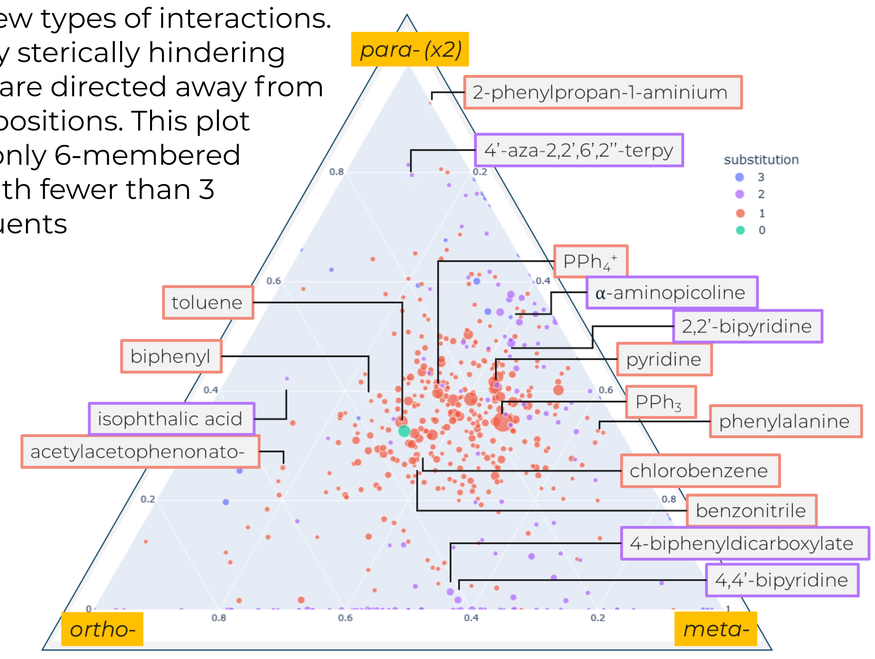
Most drugs have two or fewer strong interactions. Aromatic groups in drugs form only a few types of network topology, given most have fewer than three rings. Topology calculations can be performed using a vertex method with molecular centroids.

Predicting interactions —

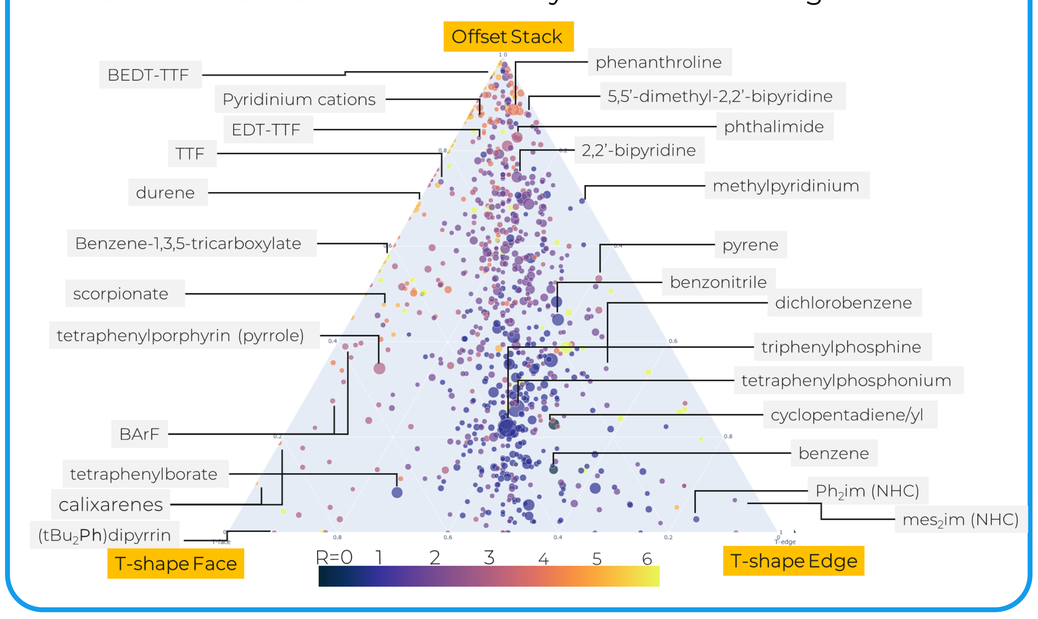
Substitution patterns may direct chemical behaviour, similar to S_NAr reactivity direction. In our data set, interactions appear around the best interaction profiles, and common single substitutions on benzene are not strongly *ortho-*, *meta-* or *para-* directions, with the percentages representing the excess of one interaction type over another



Relative propensity of T-shapes by mediating proton. Some substitution patterns are strongly directing — $C_6H_5NR_3^+$ directs to *meta-*, while pyridyl units coordinating metal centres can be activated to only a few types of interactions. Strongly sterically hindering groups are directed away from *ortho–* positions. This plot shows only 6-membered rings with fewer than 3 substituents



Relative propensity of different interaction types. Structures with more substituents form more offset-stack interactions. Tetraphenylborates are more likely to be the ‘face’ of a T-shape, pyridinium units and common solvents are more likely to interact as ‘edge’



Aromatics Analyser 2 is available through the CSD Python API menu in Mercury, through a command line interface and inside Jupyter notebooks. More information is available in the **readme.md** or by running ***./aa2.py --help***. Installation instructions will be available from the CCDC upon release.

For more information: **General:** E. Pidcock *et al*, 2022 https://doi.org/10.1021/acs.cgd.1c01293; Martinez and Iverson, 2012 https://doi.org/10.1039/C2SC20045G **Med Chem:** Bissantz, Kuhn, Stahl 2010 https://doi.org/10.1021/jm100112j; **Nucleation:** A. J. Cruz-Cabeza *et al*, 2017 https://doi.org/10.1039/C7CC02423A; **Coordinate system (abr.):** Huber et al, 2014 https://doi.org/10.1021/ci500183u