

Supplementary Materials

Figure S1: Graph of steric accessibility surface fraction (F_{AS}) against the average number of hydrogen bonds formed ($\langle n_{HB} \rangle$) for six different donor functional groups in the CSD. Functional groups shown are as follows; amido (green triangles), aromatic amino (orange diamonds), hydroxy aliphatic (blue plus signs), hydroxy aromatic (red squares), secondary amino (purple circles) and generic (black crosses).

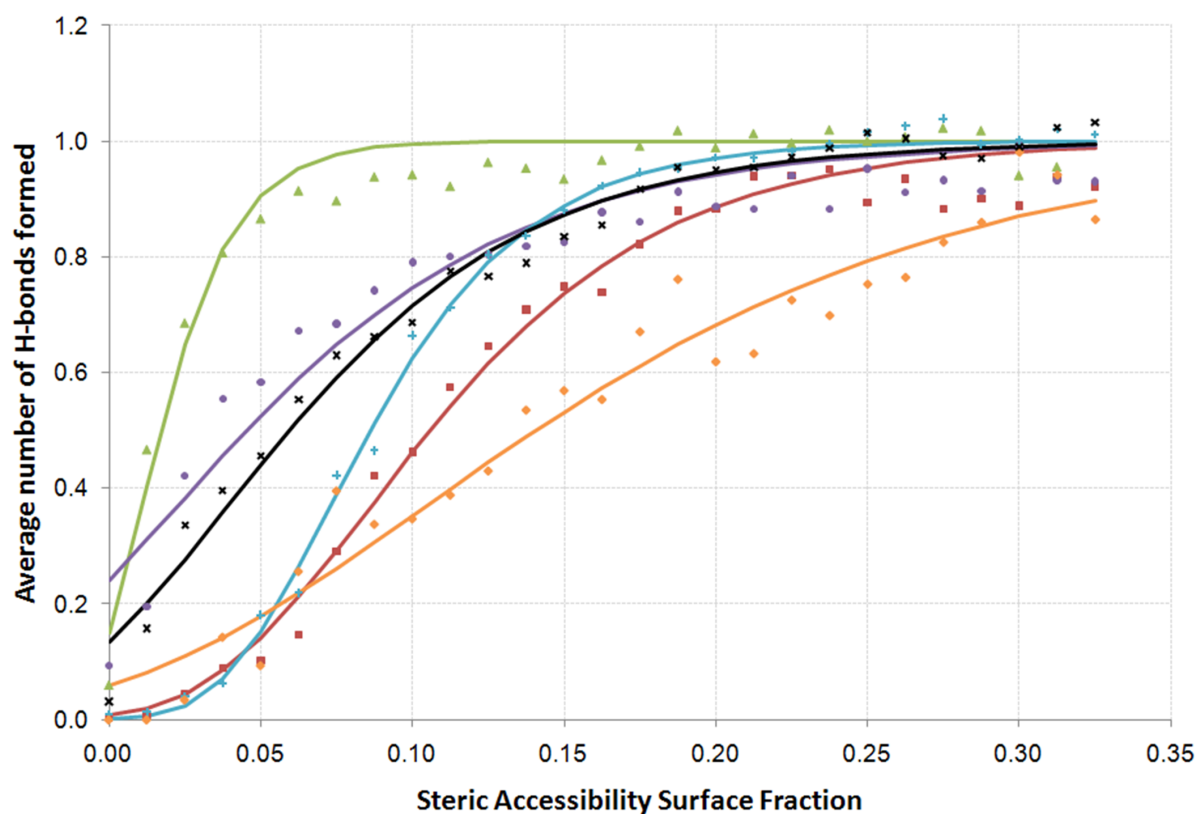
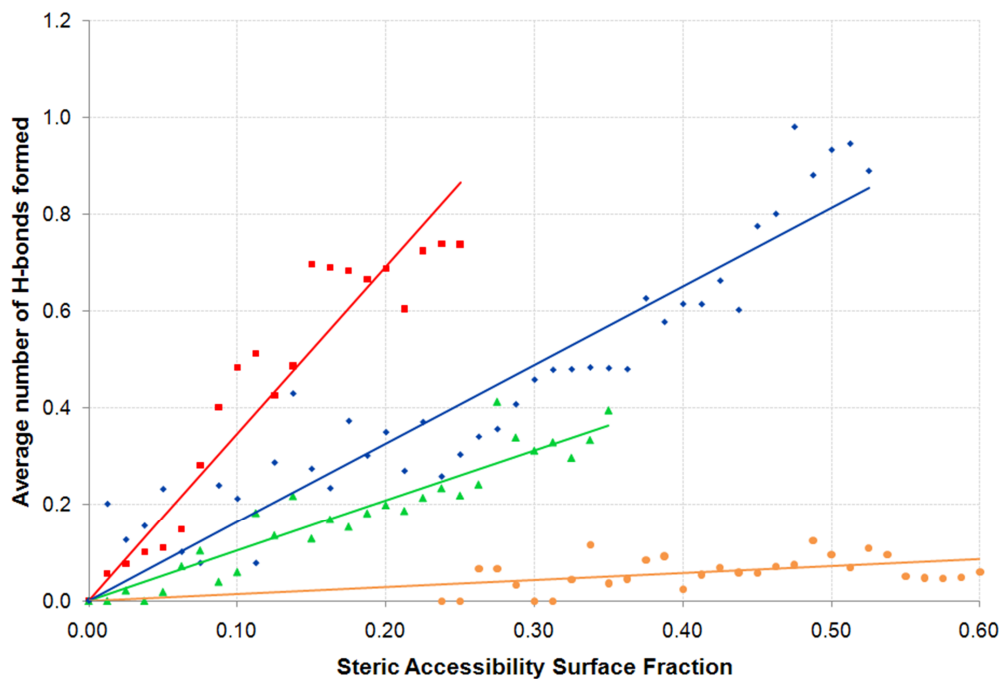
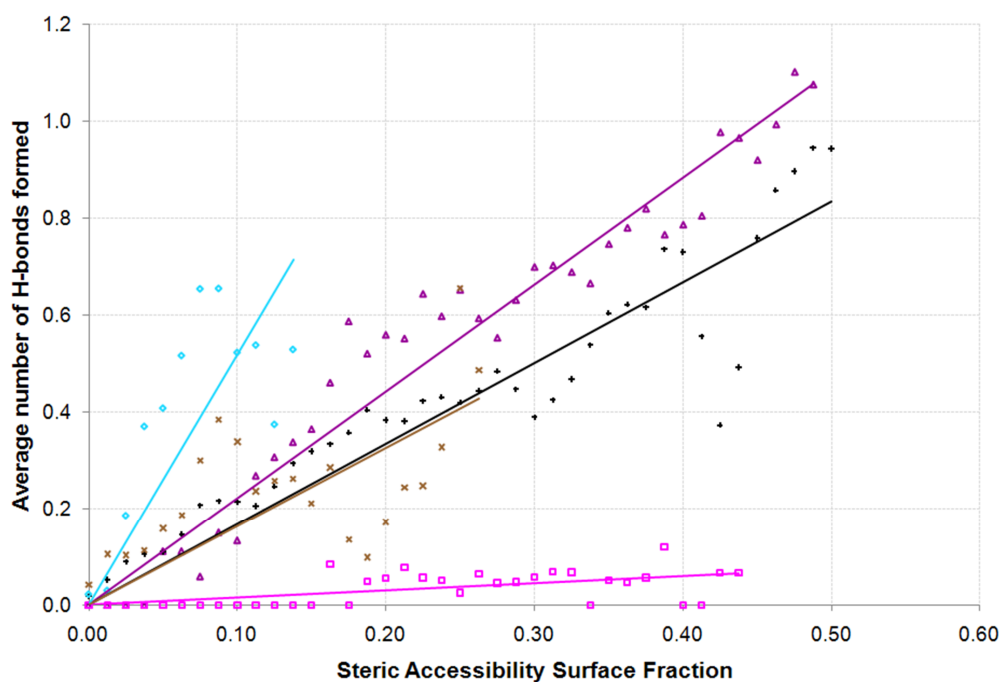


Figure S2: Graphs of steric accessibility surface fraction (F_{AS}) against the average number of hydrogen bonds formed ($\langle n_{HB} \rangle$) for nine different acceptor functional groups in the CSD. Functional groups are shown in two plots (on the same axes) for easier viewing. The functional groups are as follows; chlorine (orange filled circles), ketone (blue filled diamonds), phenol (green filled triangles), pyridyl (red filled squares) [graph *a*], ether (brown crosses), secondary amido (purple open triangles), secondary amino (cyan open diamonds), thioether (pink open squares) and generic (black plus signs) [graph *b*].



(a)



(b)

Table S1: Goodness-of-fit data for line-fitting in the plots of steric accessibility surface fraction (F_{AS}) against the average number of hydrogen bonds formed ($\langle n_{HB} \rangle$) for donor (left) and acceptor (right) groups as shown in Figs S1 & S2.

Donor Group	R^2
Amido	0.964
Aromatic amino	0.952
Carbamoyl	0.975
Generic	0.984
Hydroxyl, aliphatic	0.997
Phenol	0.988
Secondary amino	0.934

Acceptor Group	R^2
Chlorine	0.267
Ether	0.183
Generic	0.859
Keto	0.879
Phenol	0.893
Pyridyl	0.892
Secondary amido	0.947
Secondary amino	0.391
Thioether	0.349

Figure S3: Diagrams of the substructure definitions used for specific donor (*a*) and acceptor (*b*) functional groups in the study. QA is defined as N or O; QC is classified as an acceptor by CCDC toolkit code, but not identified as one of the other eight common acceptors.

