Electronic Supplementary Material (ESI) for CrystEngComm.

# Antielectrostatically hydrogen bonded anion dimers: counter-intuitive, common and consistent

Nicholas G. White

*Research School of Chemistry, The Australian National University, Canberra, ACT, Australia*

*Email: nicholas.white@anu.edu.au*

*URL: www.nwhitegroup.com*

# Search fragments

The Cambridge Structural Database (CSD)[S1] Version 5.40, November 2018 + 1 update was searched using Conquest.[S2] All structures that were found were manually checked to ensure that the structure was indeed the intended dimer, and that hydrogen atoms were pointing in the towards the other anion.

## $HCO_3^-$

The search fragment shown in Figure S1 returned 119 hits. Manual checking revealed that one of these (CSD refcode: BODGEG) was a neutral organic molecule, so this was manually removed to give 118 structures.

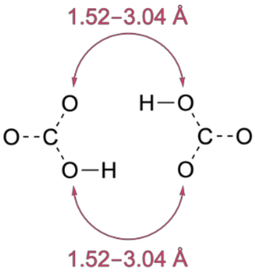


**Figure S1.** Search fragment used to find $HCO_3^-$ dimers.

## $H_2PO_4^-$

We initially searched with all P–O bonds set as "any" and no restrictions on the number of bonded species, but this returned a large number of structures including organophosphates, phosphoric acids *etc*. Instead we combined the two searches shown in Figure S2, where a * indicates the oxygen atom was limited to bonding to one group, and 4M refers to "any metal." The left search returned 289 structures, which were manually checked and all found to be genuine. The right search returned 43 structures: manual checking revealed that one (CSD refcode: RONBOL) contained a coordinated $H_2PO_4^-$ interacting with a $H_3PO_4$ molecule, so this was manually removed to give 42 structures. Searching for fragments where the metal group was coordinated to other oxygen atoms did not give any additional structures.

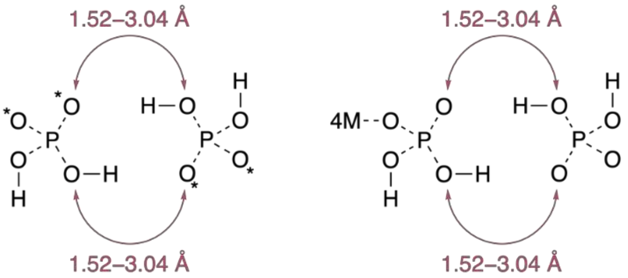


**Figure S2.** Search fragments used to find $H_2PO_4^-$ dimers.

### $HSO_4^-$

The search fragment shown in Figure S3 returned 74 hits. Manual checking revealed that one of these (CSD refcode: CUGMOG) involved a $HSO_4^-$ anion hydrogen bonding to a $H_2SO_4$ molecule, so this was manually removed to give 73 structures. We note that this is slightly fewer than the number reported by Flood who found 79 structures containing > 80 dimers (as some structures contain more than one dimer).[S3] A comparison of the structures indicates that Flood's data contain structures where the S–O–H hydrogen atom positions are not specified (but are still clearly $HSO_4^-\cdots HSO_4^-$ dimers), whereas we excluded these structures.

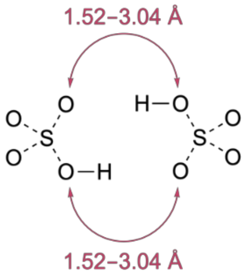


**Figure S3.** Search fragment used to find $HSO_4^-$ dimers.

### $H_2AsO_4^-$

The search fragment shown in Figure S4 returned 31 hits. Manual checking revealed that one of these (CSD refcode: ODOGUL) was a neutral organoarsenic compound, and four contained $H_3AsO_4$ (CSD refcodes: PASWIP, RIGFOD, RIGFUJ, XICNUS) so these were manually removed to give 26 structures.

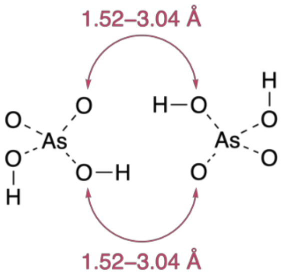


**Figure S4.** Search fragment used to find $H_2AsO_4^-$ dimers.

### $HSeO_4^-$

The search fragment shown in Figure S5 returned 5 hits. Manual checking revealed that all of these were genuine.

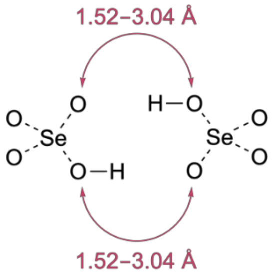


**Figure S5.** Search fragment used to find $HSeO_4^-$ dimers.

**butanoic acid**

The search fragment shown in Figure S6 returned 279 hits. Manual checking revealed that all of these were genuine.

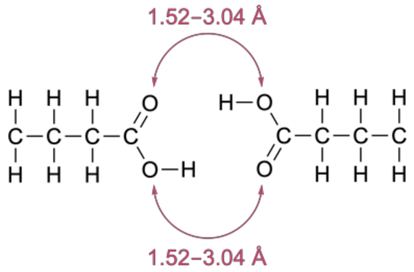


**Figure S6.** Search fragment used to find butanoic acid dimers.

**toluic acid**

The search fragment shown in Figure S7 returned 321 hits. Manual checking revealed that all of these were genuine.

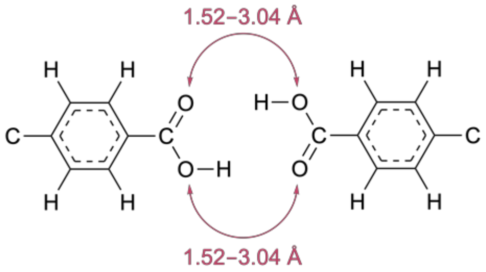


**Figure S7.** Search fragment used to find toluic acid dimers.

# Ridge plots of O⋯O distances in AEHB complexes

Ridge plots of the O⋯O distances in the AEHB dimers and carboxylic acid dimers are shown in Figure S8. These plots are scaled to have the same area, *i.e.* even rare dimers such as the $HSeO_4^-$ dimers (5 in the CSD) have the same area as common dimers.

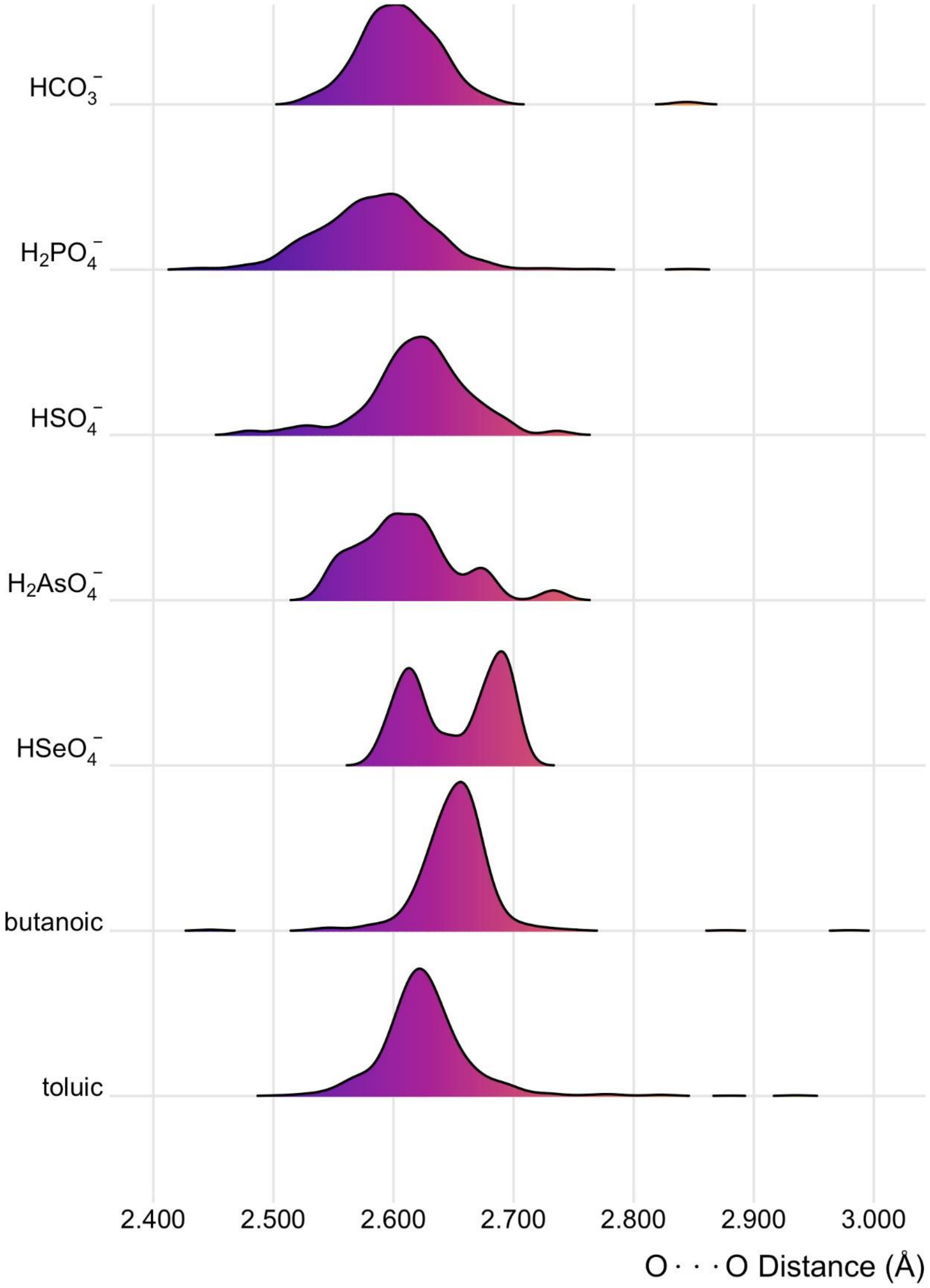


**Figure S8.** Ridge plot showing the O⋯O distances in the AEHB complexes and carboxylic acid dimers discussed in this work.

# Geometry of H-bonding interactions

The two O⋯O distances for each interaction are plotted against one another in Figure S9. Frequently, the interactions are the same or a very similar length, implying a very symmetrical dimer. Mean differences between the bond lengths are displayed on the Figure.

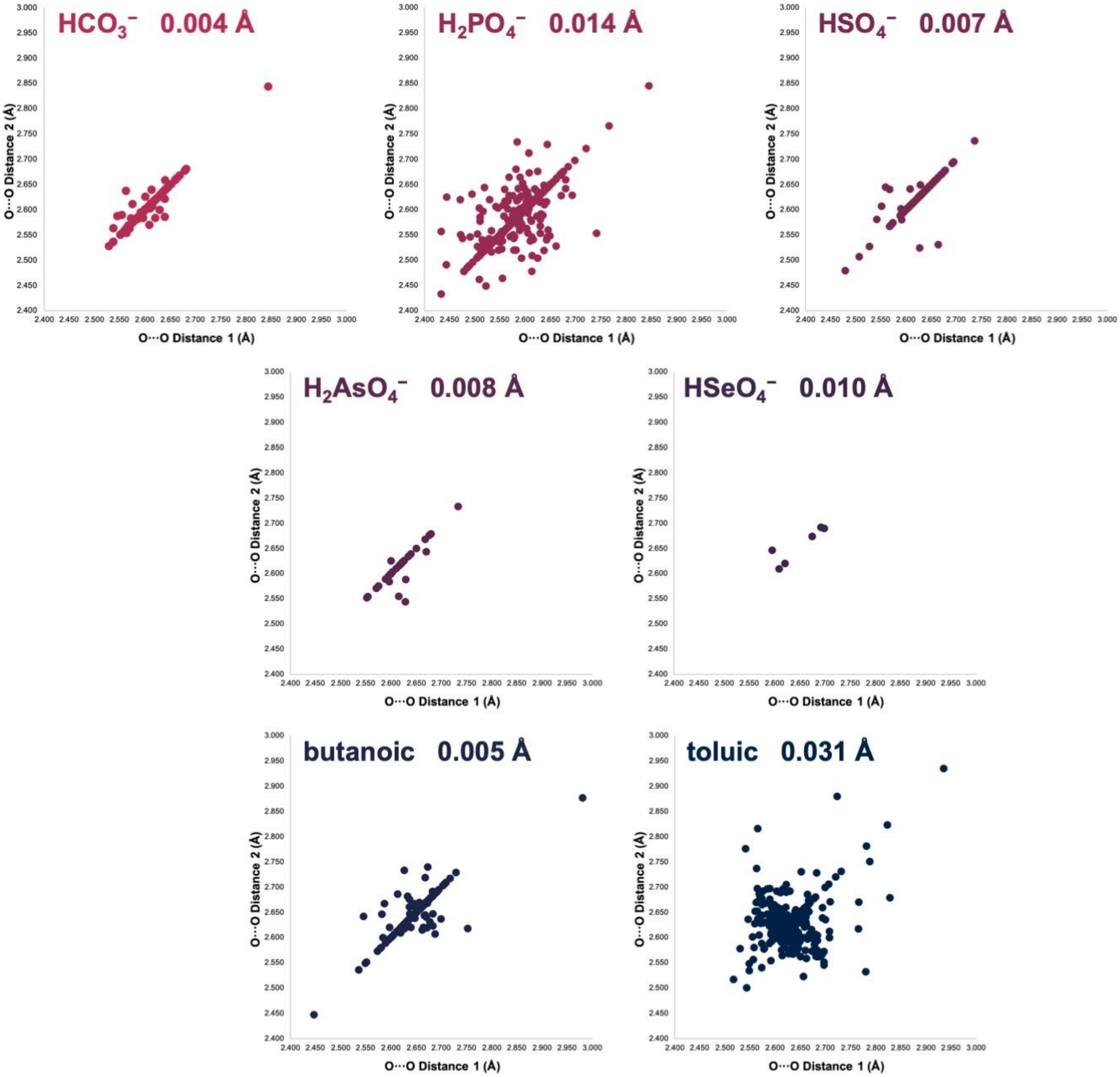


**Figure S9.** Analysis of similarity of distances in dimers, the mean difference in distance is given for each anion. Note, assignment of a distance as Distance 1 or 2 was made by Conquest search programme,[S2] and is arbitrary.

# Effects of crystal structure quality

Data in the main text refer to all crystal structures that fit a given search fragment, *i.e.* with no restrictions applied to data quality. To check if poor quality data were affecting the data analysis, we also performed searches but with restrictions applied that the structures must be "non-disordered" have "no errors" and with $R_1 < 5\%$ or $< 10\%$.

As shown in Tables S1–S3, these restrictions seemed to have little effect on the average values of the data. Therefore, in order to keep the number of data points as high as possible, no quality restrictions were applied to the data discussed in this work.

**Table S1.**[a] Comparison of interactions in AEHB anion complexes and carboxylic acid dimers considering all structures fitting a given search fragment.

| | $HCO_3^-$ | $H_2PO_4^-$ | $HSO_4^-$ | $H_2AsO_4^-$ | $HSeO_4^-$ | butanoic acid | toluic acid |
|---|---|---|---|---|---|---|---|
| no. of structures | 118 | 331 | 73 | 26 | 5 | 279 | 321 |
| no. of interactions | 254 | 942 | 156 | 64 | 12 | 632 | 956 |
| mean O···O distance (Å)[b] | 2.606(2) | 2.585(2) | 2.620(3) | 2.611(5) | 2.65(1) | 2.650(1) | 2.627(1) |
| median O···O distance (Å) | 2.605 | 2.585 | 2.624 | 2.607 | 2.660 | 2.653 | 2.623 |

[a] Note, these data are the same as Table 1 in the main manuscript. [b] Estimated standard error of mean given in parentheses.

**Table S2.** Comparison of interactions in AEHB anion complexes and carboxylic acid dimers considering structures that have no errors, are not disordered and have $R_1 < 10\%$.

| | $HCO_3^-$ | $H_2PO_4^-$ | $HSO_4^-$ | $H_2AsO_4^-$ | $HSeO_4^-$ | butanoic acid | toluic acid |
|---|---|---|---|---|---|---|---|
| no. of structures | 88 | 263 | 56 | 23 | 4 | 243 | 176 |
| no. of interactions | 184 | 718 | 118 | 56 | 10 | 552 | 408 |
| mean O···O distance (Å)[a] | 2.609(2) | 2.586(2) | 2.624(4) | 2.613(6) | 2.66(1) | 2.652(1) | 2.629(2) |
| median O···O distance (Å) | 2.605 | 2.586 | 2.625 | 2.610 | 2.674 | 2.654 | 2.625 |

[a] Estimated standard error of mean given in parentheses.

**Table S3.** Comparison of interactions in AEHB anion complexes and carboxylic acid dimers considering structures that have no errors, are not disordered and have $R_1 < 5\%$.

| | $HCO_3^-$ | $H_2PO_4^-$ | $HSO_4^-$ | $H_2AsO_4^-$ | $HSeO_4^-$ | butanoic acid | toluic acid |
|---|---|---|---|---|---|---|---|
| no. of structures | 52 | 191 | 38 | 21 | 3 | 116 | 87 |
| no. of interactions | 108 | 496 | 82 | 52 | 8 | 264 | 192 |
| mean O···O distance (Å)[a] | 2.616(3) | 2.587(2) | 2.628(4) | 2.611(6) | 2.67(1) | 2.650(2) | 2.632(3) |
| median O···O distance (Å) | 2.609 | 2.585 | 2.623 | 2.603 | 2.682 | 2.654 | 2.626 |

[a] Estimated standard error of mean given in parentheses.

# Analysis of factors affecting AEHB distances

## $H_2PO_4^-$ dimers, clusters and polymers

The CSD was searched using the fragment shown in Figure S10; after removing 10 $H_3PO_4$-containing structures, a total of 98 structures were returned. These were manually checked and all found to be genuine.

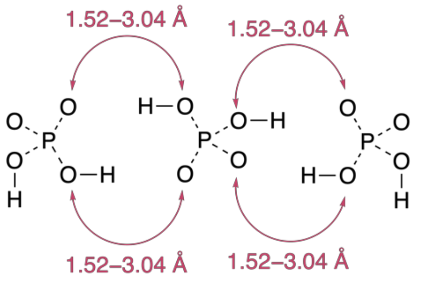
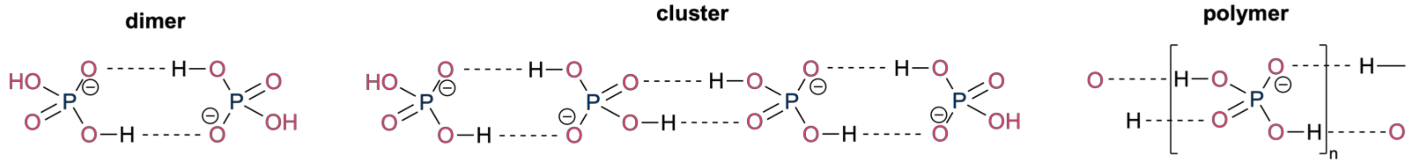


**Figure S10.** Search fragment used to find $H_2PO_4^-$ clusters and polymers.

A total of nine clusters (trimers to octamers) were found, as well as 89 polymeric structures (see Figure S11 for a definition of clusters and polymers).



dimer

cluster

polymer

**Figure S11.** Example structures of $H_2PO_4^-$ dimers, clusters (tetramer shown) and polymers. Note: some clusters were not linear. See He and Sessler,[S4] and Custelcean[S5] for a more detailed discussion.

As shown in Table S4, clusters and polymeric structures had slightly shorter O···O contacts than dimeric structures.

**Table S4.** Comparison of interactions in dimeric, cluster and polymeric $H_2PO_4^-$ complexes.

| | dimers | clusters | polymeric | polymers and clusters combined |
|---|---|---|---|---|
| no. of structures | 223 | 9 | 89 | 98 |
| no. of interactions[a] | 548 | 52 | 476 | 528 |
| mean O···O distance (Å)[b] | 2.616(3) | 2.555(7) | 2.573(2) | 2.571(2) |
| median O···O distance (Å) | 2.609 | 2.555 | 2.577 | 2.575 |

[a] Due to the way Conquest[S2] counts distances across symmetry positions, these numbers do not add up to the same number as those given in Tables 1 and S1. [b] Estimated standard error of mean given in parentheses.

## Effect of metal coordination

AEHB complexes where at least one anion is coordinated to a metal cation are compared with those that are not metal-coordinated in Table S5. There is no significant difference between the coordinated and non-coordinated structures.

Interestingly, the vast majority of coordinated structures have both anions coordinated to a metal cation.

**Table S5.** Comparison of interactions in metal-coordinated and non-coordinated anion dimers**.**

| | **$H_2PO_4^-$** | | **$HCO_3^-$** | | **$HSO_4^-$** | |
|---|---|---|---|---|---|---|
| | **metal-coordinated** | **non-coordinated** | **metal-coordinated** | **non-coordinated** | **metal-coordinated** | **non-coordinated** |
| no. of structures | 42 | 289 | 33 | 85 | 4 | 69 |
| no. of interactions | 102 | 840 | 74 | 180 | 14 | 142 |
| mean O···O distance (Å)[a] | 2.592(6) | 2.584(2) | 2.606(6) | 2.606(2) | 2.620(4) | 2.620(4) |
| median O···O distance (Å) | 2.592 | 2.585 | 2.600 | 2.607 | 2.624 | 2.624 |

[a] Estimated standard error of mean given in parentheses.

# DFT calculations

Gas phase equilibrium geometries for the $HCO_3^-$, $H_2PO_4^-$ and $HSO_4^-$ anions were calculated using Spartan ’18,[S6] and the electrostatic charges for the oxygen and hydrogen atoms compared (Table S6). As electrostatic charges are known to be strongly affected by choice of density functional,[S7] three different functionals that are known to deal well with non-covalent interactions were used;[S8–S10] the 6-311+G** basis set was used for all calculations.

While slightly different absolute values are obtained in each case, the trend is the same. In all cases, $HSO_4^-$ has the most highly positive hydrogen atom, followed by $H_2PO_4^-$ and then $HCO_3^-$ (*i.e.* consistent with p$K_a$). In all cases $H_2PO_4^-$ has the most negative oxygen atoms, suggesting that from an electrostatic perspective these would be the best hydrogen bond acceptors. Clearly, a fully electrostatic description of hydrogen bonding is a simplification.[S11]

**Table S6.** Electrostatic partial charges on atoms of $HCO_3^-$, $H_2PO_4^-$ and $HSO_4^-$ using different density functionals.[a]

| **Functional** | **ωB97X-D** [S8] | | | **B97M-V** [S9] | | | **M06-2X** [S10] | | |
|---|---|---|---|---|---|---|---|---|---|
| Anion | $HCO_3^-$ | $H_2PO_4^-$ | $HSO_4^-$ | $HCO_3^-$ | $H_2PO_4^-$ | $HSO_4^-$ | $HCO_3^-$ | $H_2PO_4^-$ | $HSO_4^-$ |
| H | +0.374 | +0.375 | +0.382 | +0.353 | +0.357 | +0.362 | +0.376 | +0.377 | +0.384 |
| H | – | +0.377 | – | – | +0.354 | – | – | +0.380 | – |
| **Mean H** | **+0.374** | **+0.376** | **+0.382** | **+0.353** | **+0.356** | **+0.362** | **+0.376** | **+0.379** | **+0.384** |
| O | –0.754 | –0.777 | –0.708 | –0.737 | –0.756 | –0.683 | –0.760 | –0.785 | –0.718 |
| O | –0.760 | –0.780 | –0.714 | –0.740 | –0.759 | –0.690 | –0.762 | –0.789 | –0.718 |
| O | –0.830 | –0.844 | –0.768 | –0.806 | –0.826 | –0.742 | –0.832 | –0.857 | –0.778 |
| O | – | –0.946 | –0.768 | – | –0.928 | –0.742 | – | –0.958 | –0.778 |
| **Mean O** | **–0.781** | **–0.837** | **–0.740** | **–0.761** | **–0.817** | **–0.714** | **–0.785** | **–0.847** | **–0.748** |

[a] All calculations were conducted using 6-311+G** basis set.

# References

[S1] C. R. Groom, I. J. Bruno, M. P. Lightfoot, S. C. Ward, The Cambridge Structural Database, *Acta. Crystallogr.* **2016**, *B72*, 171–179.

[S2] I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson, R. Taylor, New software for searching the Cambridge Structural Database and visualizing crystal structures, *Acta Crystallogr.* **2002**, *B58*, 389–397.

[S3] E. M. Fatila, E. B. Twum, A. Sengupta, M. Pink, J. A. Karty, K. Raghavachari, A. H. Flood, Anions stabilize each other inside macrocyclic hosts, *Angew. Chem., Int. Ed.* **2016**, *55*, 14057–14062.

[S4] Q. He, P. Tu, J. L. Sessler, Supramolecular chemistry of anionic dimers, trimers, tetramers, and clusters, *Chem*, **2018**, *4*, 46–93.

[S5] A. Rajbanshi, S. Wan, R. Custelcean, Dihydrogen phosphate clusters: trapping $H_2PO_4^-$ tetramers and hexamers in urea-functionalized molecular crystals, *Cryst. Growth Des.* **2013**, *13*, 2233–2237.

[S6] **Spartan '18**, *Wavefunction, Inc.* Irvine, CA.

[S7] C. P. Kelly, C. J. Cramer, D. G. Truhlar, Accurate partial atomic charges for high-energy molecules using class IV charge models with the MIDI! basis set, *Theor. Chem. Acc.* **2005**, *113*, 133–151.

[S8] J.-D. Chai, M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections, *Phys. Chem. Chem. Phys*. **2008**, *10*, 6615–6620.

[S9] N. Mardirossian, M. Head-Gordon, Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V, *J. Chem. Phys.* **2015**, *142*, 074111.

[S10] Y. Zhao, D. G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[S111] A. Sengupta, Y. Liu, A. H. Flood, K. Raghavachari, Anion-binding macrocycles operate beyond the electrostatic regime: interaction distances matter, *Chem. Eur. J.* **2018**, *24*, 14409–14417.