

Electronic Supporting Information for

Hydrogen atoms in supramolecular chemistry: a structural perspective.

Where are they, and why does it matter?

Amber L. Thompson^{*,†} and Nicholas G. White^{*,‡}

[†]*Chemical Crystallography, Chemistry Research Laboratory, Mansfield Road,
University of Oxford, Oxford OX1 3TA, UK*

[‡]*Research School of Chemistry, The Australian National University,
Canberra 2601, ACT, Australia*

E-mail: amber.thompson@chem.ox.ac.uk; nicholas.white@anu.edu.au

Abstract

Hydrogen bonding interactions are ubiquitous across the biochemical and chemical sciences, and are of particular interest to supramolecular chemists. They have been used to assemble hydrogen bonded polymers, cages and frameworks and are the functional motif in many host-guest systems. Single crystal X-ray diffraction studies are often used as key support for proposed structures, although this presents challenges as hydrogen atoms interact only weakly with X-rays. In this *Tutorial Review*, we discuss the information that can be gleaned about hydrogen bonding interactions through crystallographic experiments, key limitations of the data, and emerging techniques to overcome these limitations.

Contents

1	D—H···O distances from neutron diffraction	3
2	The true value for X—H distances from the CSD	5
3	Details of crystal refinements with manually varied X—H bond lengths	6
3.1	Refinements of 1·Cl (CCDC 2050539)	6
3.2	Refinements of 2·12H₂O (CCDC 2169825)	7
3.3	Refinements of anhydrous 2 (CCDC 2169830)	9
	References	11

1 D—H···O distances and angles for structures determined using neutron diffraction

The Cambridge Structural Database (CSD)¹ V5.43 plus four updates (March 2022, June 2022, September 2022 & November 2022) was searched for structures determined using neutron diffraction experiments containing C—H···O, N—H···O and O—H···O contacts equal to or shorter than the sum of the van der Waals' radii² of hydrogen and oxygen, i.e. $\leq 2.72 \text{ \AA}$.

This was done by combining two searches: one search containing the desired hydrogen bond, *e.g.* a carbon atom singly bonded to a hydrogen atom with a non-bonded contact to an oxygen atom with a distance of less than or equal to 2.72 \AA , and the second search containing the word "neutron" in the "All Text" search field. Search filters were used restricting the data to single crystal structures with 3D coordinates determined and no errors or disorder. In the interest of obtaining relatively high quality data, only structures with R factors $\leq 7.5\%$ were considered.

These searches found 301 different structures containing 921 C—H···O contacts, 56 different structures containing 83 N—H···O contacts and 129 different structures containing 322 O—H···O contacts. These data are presented graphically in Figure 3 of the main manuscript. The data are also presented in Tables S1 and S2, and have been uploaded in Spreadsheet format.

Table S1: Structural data for D—H···O distances (in Å) determined by neutron diffraction in the Cambridge Structural Database.

Parameter	C—H···O	N—H···O	O—H···O
Number	921	83	322
Minimum distance	2.080	1.686	1.094
Maximum distance	2.720	2.704	2.716
Median distance	2.516	2.324	1.839
Mean distance	2.506	2.255	1.857
Standard error of mean ^a	0.0044	0.025	0.031

^a Estimated as the sample standard deviation divided by the square root of the number of datapoints.

Table S2: Structural data for D—H···O angles (in °) determined by neutron diffraction in the Cambridge Structural Database.

Parameter	C—H···O	N—H···O	O—H···O
Number	921	83	322
Minimum angle	90.02	90.33	90.03
Maximum angle	176.92	167.65	180.00
Median angle	99.06	109.25	163.07
Mean angle	105.39	116.90	146.78
Standard error of mean ^a	0.53	2.46	1.78

^a Estimated as the sample standard deviation divided by the square root of the number of datapoints.

2 Investigations into the true value for X—H distances from the Cambridge Structural Database

A simple “3D” search for a C—H distance for hydrogen atoms bound to a phenyl ring using CSD¹ V5.43 (November 2021) plus four updates (March 2022, June 2022, September 2022 & November 2022) specifying single crystal data collected at 100 K with X-ray radiation and no errors or disorder gave 74,400 structures comprising 896,666 hits in a predominantly bimodal distribution (Figure S1). The mode value, i.e. the most frequently occurring value was 0.9500 Å with 465,167 hits, i.e. 51.9%, or more than half. The next most frequent value was 0.9300 Å with 111,837 hits (i.e. 12.5%). These two values, precise to four decimal places (i.e. exactly 0.9500 and 0.9300), together comprise nearly two thirds of the values for this bond as recorded in the CSD. It seems highly unlikely that this is pure chance; it is far more probable that this statistic is heavily skewed by the software used.

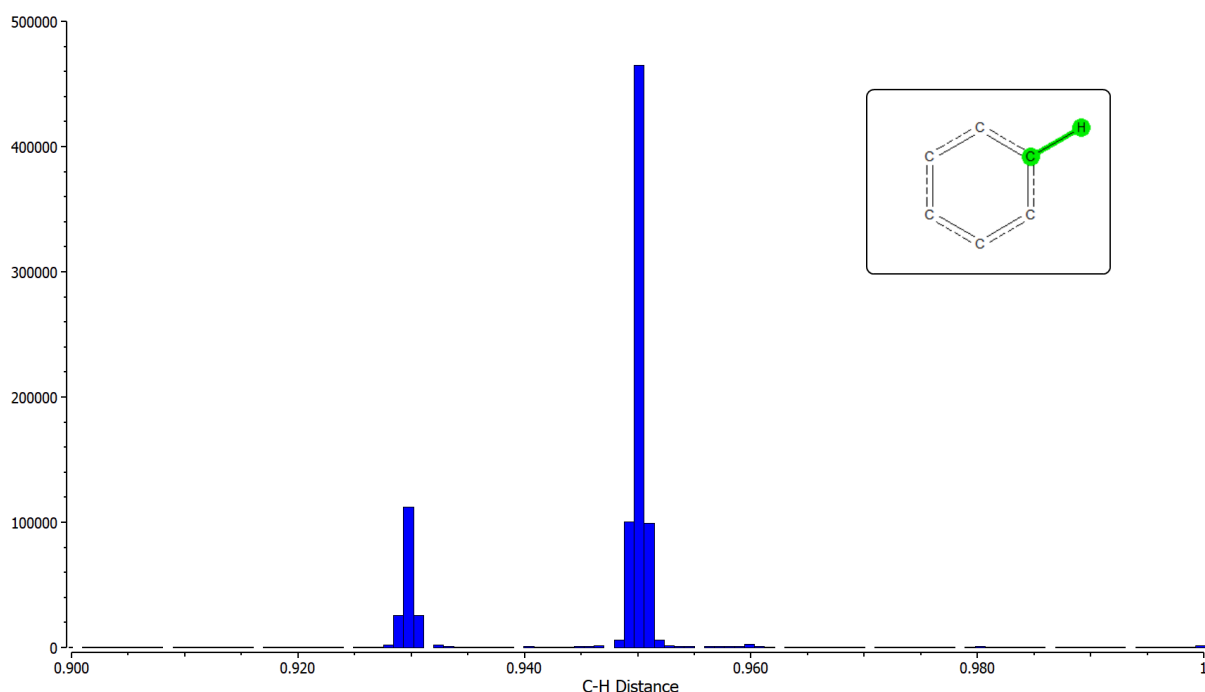


Figure S1: Histogram for distribution of C—H distances in a phenyl ring (with the search fragment inset). The range has been restricted to show values between 0.9 Å and 1.0 Å, however, there are few outliers outside this region.

3 Details of crystal refinements of 1·Cl, 2·12H₂O and anhydrous 2 with manually varied N—H and O—H bond lengths

In all cases, we started from the previously published structures of the compounds. These structures are available from the Cambridge Crystallographic Data Centre with more details given in the following section. In all cases, a significant effect on structural parameters (and ease of refinement) is only observed when extremely short or extremely long D—H values are used.

Refinements were conducted in one of three ways. In the first case, the refinements used the default hydrogen atom position restraints within the CRYSTALS suite,³⁻⁵ that is N—H bond lengths restrained to be 0.86(2) Å, C—N—H bond angles restrained to be 120(2)°, and O—H bond lengths restrained to be 0.82(2) Å. In the second case, no restraints were applied to hydrogen atoms, *i.e.* their positions were freely refined. In the third case, N—H and O—H distances were effectively fixed to values ranging from 0.50 to 1.50 Å by the use of very tight restraints (standard uncertainty = 0.001 Å).

3.1 Refinements of 1·Cl (CCDC 2050539)⁶

The chemical structure of this compound and asymmetric unit of its crystal structure are shown in Figure 8 of the main manuscript and in Figure S2. The structure contains a single O—H group, which was refined with the O—H values shown in Table S3. As can be seen, there is no detectable effect on the Flack parameter or mean C—C estimated standard deviation, and only minor effects on R_1 and wR_2 parameters, even when extreme O—H bond lengths are used.

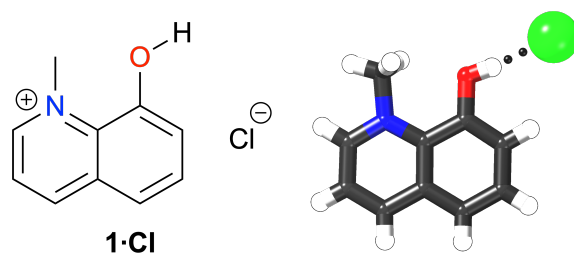


Figure S2: Chemical structure of 1·Cl and the asymmetric unit of its crystal structure (CCDC 20520539).

Table S3: Effect of O—H bond lengths on refinement parameters for the structure of **1·Cl**.

O—H distance (Å)	R ₁ [I > 2σ(I)]	wR ₂ (all)	Flack parameter	mean C—C esd (Å)
0.839(19) ^a	3.94	11.0	0.02(2)	0.0040
1.02(5) ^b	3.97	11.0	0.02(2)	0.0040
0.50 ^c	4.00	11.2	0.02(2)	0.0040
0.75 ^c	3.95	11.1	0.02(2)	0.0040
1.00 ^c	3.97	11.0	0.02(2)	0.0040
1.25 ^c	4.05	11.1	0.02(2)	0.0040
1.50 ^c	4.14	11.2	0.02(2)	0.0040

^a Hydrogen atom refined with default restraints within the CRYSTALS suite. ⁴ ^b Hydrogen atom refined with no restraints. ^c Hydrogen atom bond length fixed by restraining to the given value with heavily weighted restraints (standard uncertainty = 0.001 Å).

3.2 Refinements of **2·12H₂O** (CCDC 2169825)⁷

The chemical structure of this compound and asymmetric unit of its crystal structure are shown in Figure S3.

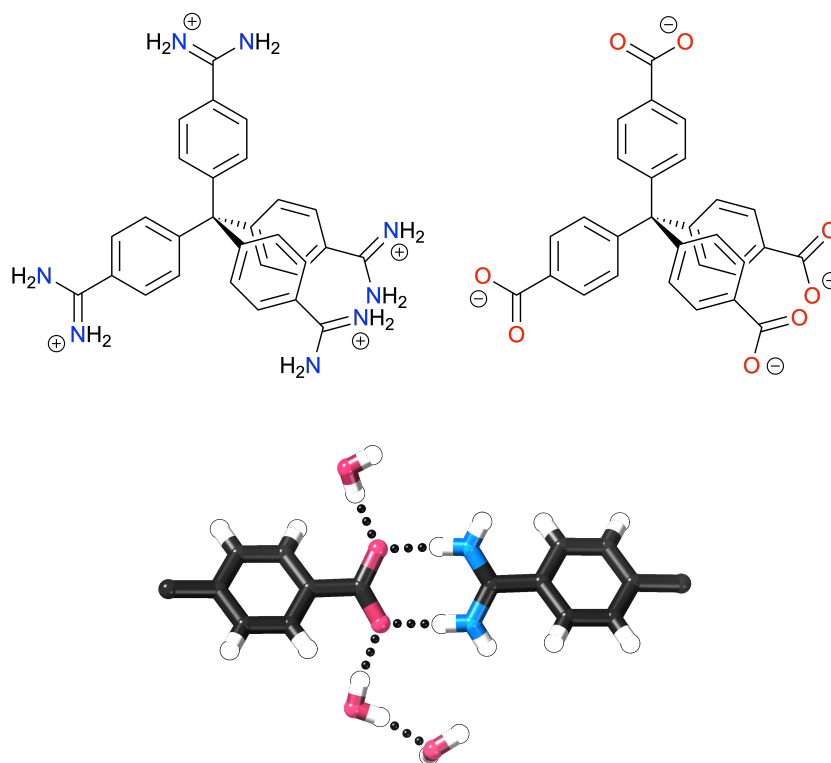


Figure S3: Chemical structure of **2·12H₂O** and the asymmetric unit of its crystal structure (CCDC 2169825). Note that this structure is a 3D hydrogen bonded framework, $Z' = 0.25$.

This structure is the hydrated phase of a hydrogen bonded framework assembled by charge-assisted hydrogen bonding between tetra-amidinium and tetra-carboxylate components, which contains twelve well-ordered water molecules in the framework channels per formula unit. Data

were collected at 100 K and are of good quality, at least for a supramolecular framework material.

We conducted two sets of refinements: one where only the four amidinium N—H bond lengths were varied (Table S4), and one where the four N—H bond lengths as well as the six water O—H bond lengths were varied (Table S5). In both cases, using no restraints for the hydrogen atoms gave better refinement metrics than when using the default parameters, although in the case of the O—H bonds this led to one of the water hydrogen atoms refining to an unreasonable position with an O—H length of 1.46 Å. As with **1·Cl** a significant deterioration in refinement metrics was only observed when D—H bond lengths were refined to extreme values.

Table S4: Effect of N—H bond lengths on refinement parameters for the structure of **2·12H₂O**.

N—H distances (Å)	R ₁ [I > 2σ(I)]	wR ₂ (all)	mean C—C esd (Å)
0.871(14) – 0.914(14) ^a	5.55	17.58	0.0019
0.89(2) – 1.02(2) ^b	5.47	17.33	0.0019
0.50 ^c	6.84	21.49	0.0027
0.75 ^c	5.76	18.32	0.0019
1.00 ^c	5.61	17.63	0.0019
1.25 ^c	6.34	19.59	0.0019
1.50 ^c	7.84	23.10	0.0028

^a Hydrogen atom refined with default restraints within the CRYSTALS suite.⁴ ^b Hydrogen atom refined with no restraints. ^c Hydrogen atom bond length fixed by restraining to the given value with heavily weighted restraints (standard uncertainty = 0.001 Å).

Table S5: Effect of N—H and O—H bond lengths on refinement parameters for the structure of **2·12H₂O**.

N—H and O—H distances (Å)	R ₁ [I > 2σ(I)]	wR ₂ (all)	mean C—C esd (Å)
0.839(15) – 0.914(16) ^a	5.55	17.58	0.0019
0.889(18) – 1.46(4) ^b	5.01	16.14	0.0018
0.50 ^c	7.89	23.39	0.0028
0.75 ^c	6.14	18.87	0.0019
1.00 ^c	5.44	17.15	0.0019
1.25 ^c	6.32	20.15	0.0020
1.50 ^c	8.94	25.83	0.0029

^a Hydrogen atom refined with default restraints within the CRYSTALS suite.⁴ ^b Hydrogen atom refined with no restraints. ^c Hydrogen atom bond length fixed by restraining to the given value with heavily weighted restraints (standard uncertainty = 0.001 Å).

3.3 Refinements of anhydrous **2** (CCDC 2169830)⁷

The chemical structure of this compound and asymmetric unit of its crystal structure are shown in Figure S4. This structure was recorded at 318 K after crystals of **2**·12H₂O had undergone a phase change. This phase change gives an anhydrous structure but results in a significant loss of crystal quality. The data are quite poor and diffraction could only be observed to a resolution of approximately 1.1 Å. While to many crystallographers, data like these are often considered unsatisfactory,⁸ this is quite typical of the data that can be obtained for many supramolecular/framework systems.

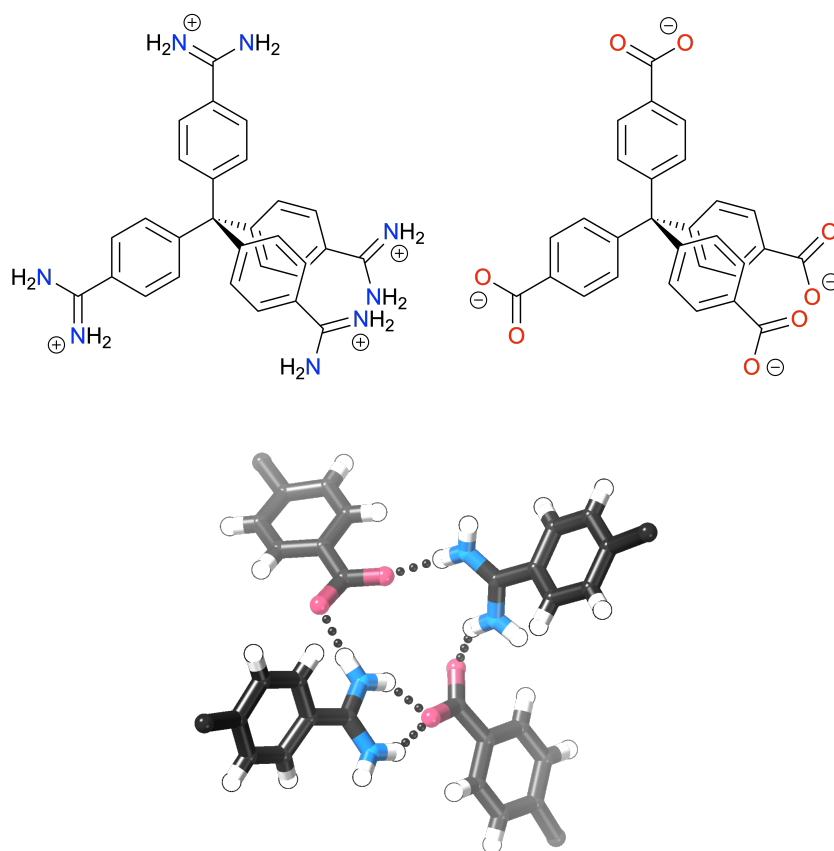


Figure S4: Chemical structure of anhydrous **2** and the asymmetric unit of its crystal structure (CCDC 2169830). Note that this structure is a 3D hydrogen bonded framework, $Z' = 0.5$.

The poor data for this structure mean that it was not possible to achieve a stable refinement without restraints on N—H bond lengths. It was also not possible to achieve a stable refinement with N—H bond lengths restrained to 0.50 or 0.75 Å and so hydrogen atoms were instead inserted at these positions and then these used as the basis for a riding model. With relatively poor data like this, changing the N—H distances to even extreme values has only a very small effect on refinement metrics (Table S6).

Table S6: Effect of N—H bond lengths on refinement parameters for the structure of anhydrous **2**.

N—H distances (Å)	R₁ [I >2σ(I)]	wR₂(all)	mean C—C esd (Å)
0.861(18) – 0.867(18) ^a	10.69	28.47	0.0140
0.50 ^b	11.02	28.92	0.0142
0.75 ^b	10.82	28.64	0.0140
1.00 ^c	10.68	28.52	0.0140
1.25 ^c	10.68	28.63	0.0140
1.50 ^c	10.80	28.89	0.0140

^a Hydrogen atom refined with default restraints within the CRYSTALS suite.⁴ ^b Hydrogen atom bond lengths fixed by inserting at positions with the given bond length (and idealised sp² bond angles) and using these positions as the basis for a riding model. ^c Hydrogen atom bond length fixed by restraining to the given value with heavily weighted restraints (standard uncertainty = 0.001 Å).

References

- (1) R. Taylor and P. A. Wood, *Chem. Rev.*, 2019, **119**, 9427–9477.
- (2) S. Alvarez, *Dalton Trans.*, 2013, **42**, 8617–8636.
- (3) P. W. Betteridge, J. R. Carruthers, R. I. Cooper, K. Prout and D. J. Watkin, *J. Appl. Cryst.*, 2003, **36**, 1487.
- (4) R. I. Cooper, A. L. Thompson and D. J. Watkin, *J. Appl. Cryst.*, 2010, **43**, 1100–110.
- (5) P. Parois, R. I. Cooper and A. L. Thompson, *Chem. Cent. J.*, 2015, **9**, Article no. 30.
- (6) R. J. Goodwin, M. T. Blyth, A. K. K. Fung, L. M. Smith, P. L. Norcott, S. Tanovic, M. L. Coote and N. G. White, *Org. Biomol. Chem.*, 2021, **19**, 2794–2803.
- (7) S. A. Boer, L. Conte, A. Tarzia, M. T. Huxley, M. G. Gardiner, D. R. T. Appadoo, C. Ennis, C. J. Doonan, C. Richardson and N. G. White, *Chem. Eur. J.*, 2022, **28**, e202201929.
- (8) R. L. Harlow, *J. Res. Natl. Inst. Stand. Technol.*, 1996, **101**, 327–339.