


 Cite this: *Phys. Chem. Chem. Phys.*, 2022, 24, 18124

# Quantification of secondary electrostatic interactions in H-bonded complexes†

 Maria Chiara Storer  and Christopher A. Hunter \*

The H-bonding properties of compounds that contain multiple functional groups are difficult to predict, because there are through-bond polarisation effects and long-range secondary electrostatic interactions that have significant effects on the interactions with solvents and other molecules. Here we use experimental measurements of association constants for formation of 1:1 H-bonded complexes that contain a single well-defined H-bond and a single well-defined secondary electrostatic interaction to quantify the magnitude of this effect. The results were used to develop a computational method for calculating functional group H-bond parameters that accurately reproduce the magnitudes of both primary H-bonding interaction and secondary electrostatic interactions. The effects of secondary electrostatic interactions are observed in calculations of *ab initio* Molecular Electrostatic Potential (MEP) values, but at the van der Waals surface, the magnitude of the effect is highly overestimated. MEP values calculated on electron density isosurfaces that lie closer to the nuclei provide a more accurate description of the experimental observations. H-bond parameters calculated using this approach successfully account for the properties of arrays of multiple H-bond donor and acceptor groups in different configurations. The results provide insight into the factors that govern the interaction properties of molecules that contain multiple functional groups and provide an accurate method for prediction of solution phase complexation free energies based on gas phase calculations of individual molecules.

 Received 1st July 2022,  
 Accepted 12th July 2022

DOI: 10.1039/d2cp03004g

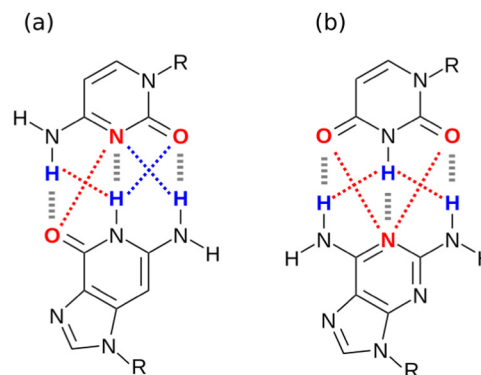
rsc.li/pccp

## 1 Introduction

H-bonding interactions are ubiquitous in nature and play an important role in materials and supramolecular chemistry. When a donor functional group comes into contact with an acceptor functional group, attractive non-covalent interactions lead to formation of a H-bond.<sup>1,2</sup> For molecules with single functional groups the donor and acceptor properties have been empirically characterised, and the resulting H-bonding scales can be used to reliably predict the stabilities of the complexes, solvation energies and partition coefficients.<sup>3,4</sup> However, using these scales to rationalise the behaviour of molecules containing multiple functional groups is less straightforward.

One factor that complicates the behaviour of molecules with closely packed arrays of polar groups is secondary electrostatic interactions, which are illustrated in Fig. 1. Although the two complexes are at first sight very similar, each with three intermolecular H-bonds, the association constant for formation

of the complex shown in Fig. 1(a) is two orders of magnitude higher than the association constant for formation of the complex shown in Fig. 1(b) ( $10^4$ – $10^5$  M<sup>-1</sup> compared with  $10^2$  M<sup>-1</sup> in chloroform).<sup>5,6</sup> In their classic paper,<sup>7</sup> Jorgensen and Pranata used molecular mechanics calculations to offer an explanation. The primary H-bonding interactions shown in



**Fig. 1** Secondary electrostatic interactions in multiply H-bonded complexes. (a) Three primary H-bonding interactions (grey), two repulsive secondary electrostatic interactions (red) and two attractive secondary electrostatic interactions (blue). (b) Three primary H-bonding interactions (grey) and four repulsive secondary electrostatic interactions (red).

Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK. E-mail: herchelsmith.orgchem@ch.cam.ac.uk

† Electronic supplementary information (ESI) available. CCDC 2170565 and 2170566. For ESI and crystallographic data in CIF or other electronic format see DOI: <https://doi.org/10.1039/d2cp03004g>



Fig. 1 bring neighbouring functional groups on the two molecules into close proximity. These functional groups are further apart than the H-bonded atoms, but they are close enough to make long-range secondary electrostatic interactions. Considering H-bond acceptors as atoms with partial negative charges and H-bond donors as atoms with partial positive charges provides a simple explanation for the experimental results for these systems. In Fig. 1(a), there two repulsive (red) and two attractive (blue) secondary interactions, whereas in Fig. 1(b) all the secondary interactions are repulsive. The concept of secondary electrostatic interactions has been successfully used to design exceptionally stable H-bonded arrays, which have applications in cosmetics, adhesives and surface films.<sup>8–10</sup>

There have been many theoretical investigations of the properties of complexes with arrays of multiple H-bonds, such as those shown in Fig. 1. A large number of different factors have been identified as contributors to the overall stability of the complexes. It is clear that interactions between all of the atoms in the two molecules should be considered, and it has been shown that Brønsted–Lowry acid/base properties,  $\pi$  resonance assistance, polarisation, Pauli repulsion, dispersion and charge accumulation all play a role.<sup>11–17</sup> These analyses highlight the difficulty of disentangling the properties of complexes, which contain multiple functional groups, where both through-bond polarisation and through-space effects contribute to interaction energies, and where there are multiple primary and secondary interactions that must be considered. Here we focus on simpler complexes with fewer interactions that allow more direct investigation of secondary electrostatic interactions. The approach is illustrated in Fig. 2. The complex in Fig. 2(a) provides a reference point with only one intermolecular H-bond and no significant secondary interactions. The complex in Fig. 2(b) introduces one additional secondary electrostatic interaction. Comparison of the stabilities of these two complexes therefore provides a simple quantification the secondary electrostatic interaction.

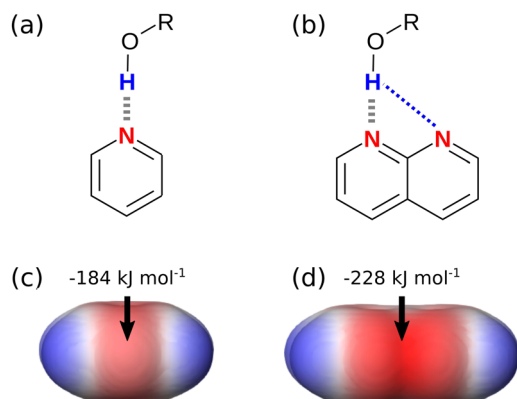


Fig. 2 (a) A complex with one intermolecular H-bond (grey). (b) A complex with one intermolecular H-bond and one attractive secondary electrostatic interaction (blue) (c) MEPS of pyridine (d) MEPS of naphthyridine. Calculations were carried out using DFT, B3LYP with a 6-31G\* basis set in NWChem, and the MEPS is shown on the 0.0020 e bohr<sup>-3</sup> electron density isosurface with the minimum value  $E_{\min}$  highlighted.

We have previously shown that the properties of complexes like the one in Fig. 2(a) which have a single intermolecular H-bond can be accurately predicted using eqn (1).

$$\Delta G^\circ/\text{kJ mol}^{-1} = -(\alpha - \alpha_s)(\beta - \beta_s) + 6 \quad (1)$$

where  $\alpha$  and  $\beta$  are H-bond donor and acceptor parameters that describe the solutes,  $\alpha_s$  and  $\beta_s$  are the corresponding H-bond parameters that describe the solvent, and 6 is an empirical constant.<sup>18</sup>

The H-bond parameters used in eqn (1) can be determined experimentally using measurements of association constants for formation of 1 : 1 complexes, or they can be calculated using *ab initio* methods. Fig. 2(c) shows the molecular electrostatic potential surface (MEPS) of pyridine calculated on the 0.0020 e bohr<sup>-3</sup> electron density isosurface, which is an approximation of the van der Waals surface.<sup>19</sup> The minimum on the MEPS,  $E_{\min}$ , is located over the nitrogen lone pair, and the value can be used in eqn (2) to calculate the H-bond acceptor parameter  $\beta$ .<sup>4</sup>

$$\beta = aE_{\min}^2 + bE_{\min} \quad (2)$$

Fig. 2(d) shows the MEPS of naphthyridine. The value of  $E_{\min}$  is significantly more negative than the value calculated for pyridine in Fig. 2(c). When these values of  $E_{\min}$  are used in eqn (2), the calculated values of  $\beta$  are 7.2 for pyridine and 10.1 for naphthyridine. Implementation of these parameters in eqn (1) would predict that the naphthyridine complex shown in Fig. 2(b) is significantly more stable than pyridine complex in Fig. 2(a). This example suggests that the MEPS of the individual molecules may provide a useful description of the effects of secondary electrostatic interactions on the stabilities of intermolecular complexes. Although it is known that through-space interactions with neighbouring functional groups affect the MEPS of a molecule, it is not clear whether the magnitude of this effect provides an accurate prediction of the magnitude of secondary electrostatic interactions.<sup>20,21</sup> In this paper, we use experimental measurements of association constants for formation of 1 : 1 complexes to establish a quantitative relationship between the *ab initio* calculated MEPS and the free energy contribution due to secondary electrostatic interactions to the stabilities of multiply H-bonded complexes.

## 2 Results

### Experimental measurement of secondary electrostatic interactions

Fig. 3 and 4 show the H-bond donors and acceptors chosen for experimental investigation. Compounds **A1** and **A2** were used as reference points, because the experimentally determined  $\beta$  parameters span a wide range (7.2 to 9.3), and there are minimal secondary electrostatic interactions. Compounds **A3**, **A4** and **A5** have two H-bond acceptor sites of different polarity arranged in different geometries, which should lead to measurable secondary electrostatic interactions and allow calibration of the magnitude of through-space effects on the calculated MEPS.



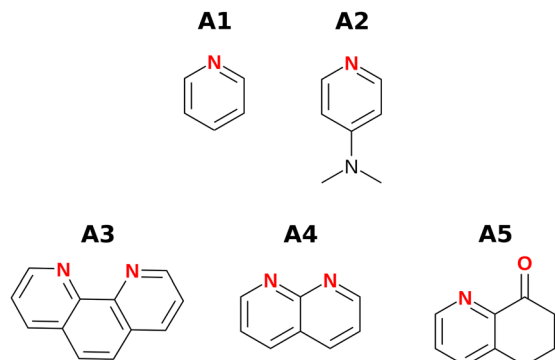


Fig. 3 H-bond acceptors.

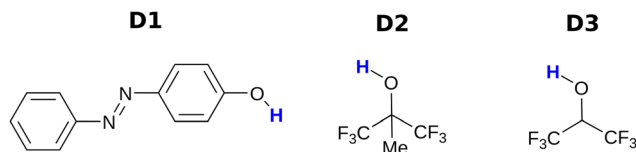


Fig. 4 H-bond donors.

The presence of two interaction sites in **A3** and **A4** is easily taken care of, because they are identical, which leads to a statistical correction of a factor of two in the measured association constant. There could be ambiguity in the site of interaction of a H-bond donor with **A5**, but pyridines are much stronger H-bond acceptors than ketones, so interaction with the oxygen is unlikely.

UV/vis absorption titrations were carried out for all combinations of H-bond donors and acceptors in chlorobenzene and dichloromethane. In four cases the overlap of the absorption spectrum of the host with the spectrum of the guest or the solvent made it impossible to collect reliable data. For the other 21 systems, the data fit well to a 1 : 1 binding isotherm (see ESI<sup>†</sup>). The resulting association constants are presented in Table 1.

The results in Table 1 were used in eqn (1) to determine experimental values of the H-bond acceptor parameters for compounds **A1–A5**,  $\beta_{\text{expt}}$ . The literature values shown in Table 2 for the H-bond parameters of **D1–D3** and the solvents allowed optimisation of the values of  $\beta_{\text{expt}}$  for **A1–A5** to minimise the RMSE between the experimental and calculated values

Table 1 Association constants ( $\text{M}^{-1}$ ) measured by UV/vis titrations at 298 K

	Solvent				
	Chlorobenzene			Dichloromethane	
	D1	D2	D3	D1	D3
A1	$99 \pm 6$	n.d. <sup>a</sup>	n.d. <sup>a</sup>	$42 \pm 3$	$94 \pm 4$
A2	$1070 \pm 20$	$496 \pm 4$	$3600 \pm 200$	$480 \pm 20$	$1000 \pm 300$
A3	$700 \pm 20$	$400 \pm 20$	$6300 \pm 100$	$71 \pm 8$	$380 \pm 30$
A4	$220 \pm 60$	$170 \pm 5$	$1630 \pm 70$	$53 \pm 5$	$210 \pm 10$
A5	n.d. <sup>a</sup>	$68 \pm 4$	$880 \pm 70$	n.d. <sup>a</sup>	$60 \pm 5$

<sup>a</sup> Not determined due to spectral overlap.

Table 2 H-bond parameters

	$\alpha_s$	$\beta_s$	$\alpha$	$\beta_{\text{expt}}$	$\beta_{\text{calc}}$
Chlorobenzene	1.4 <sup>a</sup>	1.1 <sup>a</sup>	—	—	—
Dichloromethane	1.6 <sup>b</sup>	1.5 <sup>a</sup>	—	—	—
<b>D1</b>	—	—	4.3 <sup>c</sup>	—	—
<b>D2</b>	—	—	4.0 <sup>d</sup>	—	—
<b>D3</b>	—	—	4.6 <sup>d</sup>	—	—
<b>A1</b>	—	—	—	7.2 <sup>e</sup>	7.2
<b>A2</b>	—	—	—	9.3 <sup>f</sup>	9.5
<b>A3</b>	—	—	—	7.9	14.6
<b>A4</b>	—	—	—	7.5	10.1
<b>A5</b>	—	—	—	7.7	12.0

<sup>a</sup> Ref. 3. <sup>b</sup> The literature value in ref. 3 is 1.7, but a value of 1.6 gave much better agreement with the measurements reported here. <sup>c</sup> Ref. 25. <sup>d</sup> Ref. 26. <sup>e</sup> Agrees with ref. 24. <sup>f</sup> 9.4 in ref. 24.

of  $\Delta G^\circ$ . For **A3** and **A4**, which have two equivalent H-bond acceptor sites, the value of  $\beta$  was determined by applying a statistical correction to the experimentally measured association constants in Table 1.<sup>22</sup> The optimised H-bond acceptor parameters for **A1–A5** are shown in Table 2, and the agreement between the values of  $\Delta G^\circ$  calculated using these parameters and the experimentally measured values is illustrated in Fig. 5. The calculated values of  $\Delta G^\circ$  are all within 2.5  $\text{kJ mol}^{-1}$  of the experimental values, with the exception of the **D3–A3** complex in chlorobenzene, which is highlighted in red. An X-ray crystal structure of this complex was obtained (see ESI<sup>†</sup>). In addition to the primary  $\text{OH} \cdots \text{N}$  H-bond, and there is a short contact between the CH group of the donor and the second nitrogen atom of the acceptor.<sup>23</sup> This  $\text{CH} \cdots \text{N}$  H-bond would lead to an anomalous increase the stability of this complex, and the effects would be larger in chlorobenzene, which is less polar than dichloromethane.

The values of  $\beta_{\text{expt}}$  determined for the reference compounds **A1** and **A2** agree exactly with the previously reported values, which gives confidence in the reliability of the measurements.<sup>24</sup> The other three acceptors all have values of  $\beta_{\text{expt}}$  which are

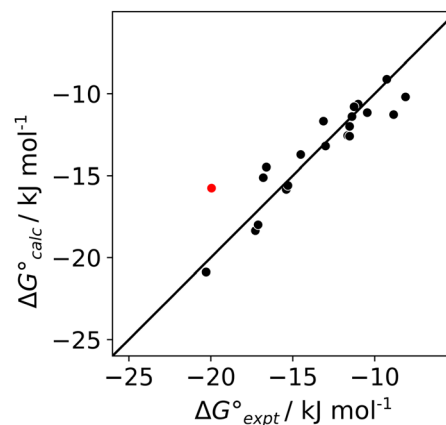


Fig. 5 Comparison of the experimental values of free energy changes for formation of 1:1 complexes ( $\Delta G_{\text{expt}}^\circ$ ) with calculated values ( $\Delta G_{\text{calc}}^\circ$ ) obtained from the H-bond parameters in Table 2 and eqn (1). The line corresponds to  $y = x$ , and the **D3–A3** complex in chlorobenzene is highlighted as an outlier (red point).



significantly larger than the value for pyridine (**A1**). This result suggests that the secondary electrostatic interactions illustrated in Fig. 2(b) do indeed enhance the stability of H-bonded complexes formed by **A3–A5**. Table 2 also shows values of  $\beta_{\text{calc}}$  which were obtained using eqn (2). For **A1** and **A2**, the calculated H-bond parameters agree with experimental values. For **A3–A5**, the calculated H-bond parameters are much larger than the experimental values, and the size of the error depends on the arrangement of H-bond acceptor sites: the discrepancy between calculation and experiment for **A3** is twice as large as the discrepancy for **A4**.

### Computational analysis of secondary electrostatic interactions

The experiments reported here show that although the calculated MEPS of an individual molecule can be used to identify the presence of secondary electrostatic interactions in a H-bonded complex, the magnitude of the effect is overestimated if the van der Waals surface is used to calculate the electrostatic potential. Through-space effects on the MEPS can be modulated by changing the isosurface used for the calculation.<sup>21,27,28</sup> Fig. 6 illustrates the principle for **A4**. Fig. 6(a) shows that the minimum on the MEPS calculated on the van der Waals surface (light grey) is close to both acceptor atoms. Therefore the value of  $E_{\text{min}}$  is determined by a large primary contribution from the closest acceptor as well as a large secondary contribution from the other acceptor. However, the magnitude of the secondary contribution can be reduced by using a higher electron density isosurface which is closer to the nuclei (dark grey in Fig. 6(b)). In Fig. 6(b), the minimum on the dark grey surface is much closer to one of the two acceptors, and so the secondary contribution to the value of  $E_{\text{min}}$  from the other acceptor is small. In other words, the magnitude of through-space effects on the MEPS can be tuned by varying the electron density of the surface used for the calculation. In this way, it should be possible to find an isosurface that provides an accurate representation of the magnitudes of both primary H-bonding interactions and secondary electrostatic interactions.

The experimentally determined H-bond acceptor parameters in Table 2 were used together with a collection of experimental parameters for compounds with  $\text{sp}^2$  nitrogen acceptors to investigate different electron density isosurfaces. For each

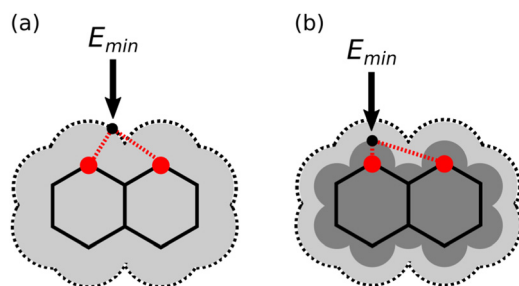


Fig. 6 Representations of electron density isosurfaces of naphthyridine. (a) On the van der Waals surface (light grey), the minimum on the MEPS (black dot) is close to both nitrogen atoms (red). (b) On a higher electron density isosurface (dark grey), the minimum on the MEPS is closer to one of the nitrogen atoms than the other.

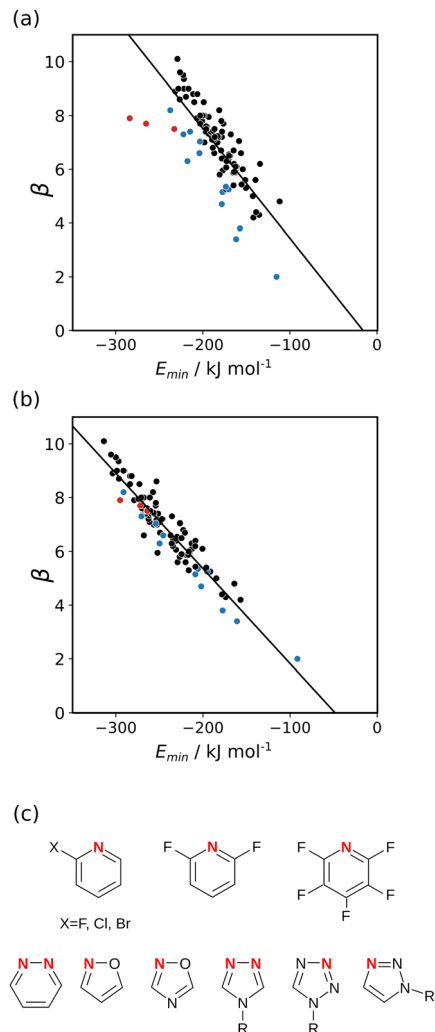


Fig. 7 Relationship between the experimentally measured H-bond parameter  $\beta$  for compounds with  $\text{sp}^2$  nitrogen acceptors and the value of  $E_{\text{min}}$  calculated using DFT, B3LYP with a 6-31G\* basis set at (a) the  $0.0020 \text{ e bohr}^{-3}$  electron density isosurface and (b) the  $0.0300 \text{ e bohr}^{-3}$  electron density isosurface. The lines are the best linear fit. The points corresponding to compounds **A3–A5** are highlighted in red. (c) The structures of the outliers highlighted in blue.

compound, the MEPS was calculated using DFT, B3LYP with a 6-31G\* basis set, at electron densities of 0.0020, 0.0050, 0.0104, 0.0200, 0.0300, 0.0400, 0.0500, 0.0600 and  $0.0700 \text{ e bohr}^{-3}$ . Fig. 7(a) shows the relationship between the experimental values of  $\beta$  and the values of  $E_{\text{min}}$  calculated at the van der Waals surface. Although there is some correlation, there is considerable scatter in the data, and the compounds from Table 2 are clear outliers (red points). The other outliers highlighted in blue in Fig. 7(a) are the compounds shown in Fig. 7(c). These compounds all have a second H-bond acceptor in close proximity to the  $\text{sp}^2$  nitrogen, and the calculated value of the MEPS overestimates the magnitude of the through-space effect on the H-bond acceptor properties. Fig. 7(b) shows the relationship between the experimental values of  $\beta$  and the values of  $E_{\text{min}}$  calculated at a higher density isosurface. The scatter is significantly reduced, and all of the outliers now



follow the same trend as the other compounds. This surface achieves the right balance of primary and secondary contributions to the value of  $E_{\min}$  to provide an accurate quantitative description of secondary electrostatic interactions.

Fig. 7(b) shows that relationship between the values  $\beta$  and  $E_{\min}$  is linear on the higher electron density isosurface. Therefore the values of  $E_{\min}$  calculated at different isosurfaces were fitted to eqn (3), and the resulting values of RMSE are compared in Fig. 8 (blue data). There is a clear minimum in RMSE for isosurface densities of 0.0200 e bohr<sup>-3</sup> to 0.0300 e bohr<sup>-3</sup> suggesting that these surfaces provide the best description of the H-bond properties of sp<sup>2</sup> nitrogen acceptors. The same analysis was carried out for amide carbonyl oxygen acceptors (see ESI†), and the resulting RMSE values are shown in Fig. 8 (red data). Although there is no well-defined minimum in this case, the higher density isosurfaces also provides a superior description of this class of H-bond acceptors. The 0.0300 e bohr<sup>-3</sup> isosurface was therefore selected for calculation of  $\beta$  parameters using eqn (3) (sp<sup>2</sup> nitrogen  $m_{\beta} = -0.036$  kJ mol<sup>-1</sup> and  $c_{\beta} = -1.7$ , amide oxygen  $m_{\beta} = -0.035$  kJ mol<sup>-1</sup> and  $c_{\beta} = 0.1$ ).

$$\beta = m_{\beta}E_{\min} + c_{\beta}. \quad (3)$$

A similar approach was then used for the H-bond donor parameter  $\alpha$ . The MEPS was calculated at different electron density isosurfaces for a collection of H-bond donors for which  $\alpha$  parameters have been experimentally determined.<sup>4</sup> Again higher electron density isosurfaces gave a superior description of the H-bond properties of these compounds compared with the van der Waals surface. Fig. 9 compares the experimental  $\alpha$  parameter for OH donors with the value of  $E_{\max}$  calculated at the 0.0104 e bohr<sup>-3</sup> isosurface. The relationship between  $\alpha$  and  $E_{\max}$  is clearly linear. Therefore the values of  $E_{\max}$  calculated at different isosurfaces were fit using eqn (4), and the resulting values of RMSE are compared in Fig. 10 (red data). There is a well-defined minimum in RMSE for an isosurface density of 0.0104 e bohr<sup>-3</sup> suggesting that this surface provides the best description of the H-bond properties of OH donors. A similar result was obtained for NH donors (blue data in Fig. 10). The

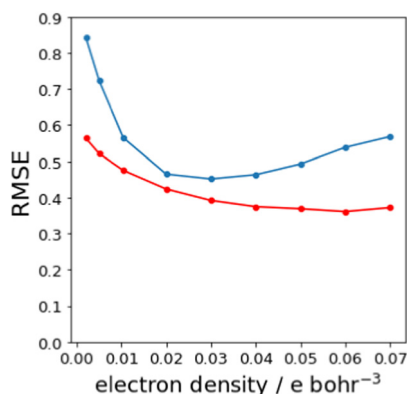


Fig. 8 RMSE between experimental values of  $\beta$  and values calculated using the best fit to eqn (3) for different electron density isosurfaces. Results for sp<sup>2</sup> nitrogen acceptors are shown in blue, and for amide carbonyl oxygen acceptors in red.

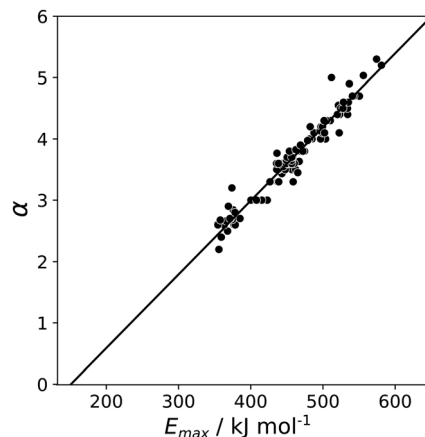


Fig. 9 Relationship between the experimentally measured H-bond parameter  $\alpha$  for compounds with OH donors and the value of  $E_{\max}$  calculated using DFT, B3LYP with a 6-31G\* basis set at the 0.0104 e bohr<sup>-3</sup> electron density isosurface. The line is the best linear fit.

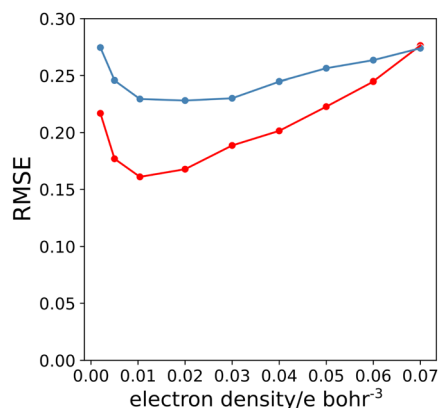


Fig. 10 RMSE between experimental values of  $\alpha$  and values calculated using the best fit to eqn (4) for different electron density isosurfaces. Results for NH donors are shown in blue, and for OH donors in red.

0.0104 e bohr<sup>-3</sup> isosurface was therefore selected for calculation of  $\alpha$  parameters using eqn (4) (OH  $m_{\alpha} = 0.015$  kJ mol<sup>-1</sup> and  $c_{\alpha} = -3.8$ , NH  $m_{\alpha} = 0.011$  kJ mol<sup>-1</sup> and  $c_{\alpha} = -2.8$ ).

$$\alpha = m_{\alpha}E_{\max} + c_{\alpha}. \quad (4)$$

Fig. 11 provides a rationale for the superior performance of the higher electron density isosurfaces in predicting H-bonding properties. Neutron diffraction experiments show that the distance between a H-bond acceptor and the hydrogen atom of the donor in a H-bond is shorter than the sum of the van der Waals radii by 0.5–1.0 Å.<sup>29</sup> On the optimal electron density isosurfaces identified above, the location of  $E_{\min}$  is 0.45 Å closer to the nucleus than the van der Waals surface and the location of  $E_{\max}$  is 0.25 Å closer to the nucleus. These locations would represent the point of contact between two H-bonded molecules that penetrate the van der Waals surfaces by 0.7 Å, which is consistent with the experimental measurements.



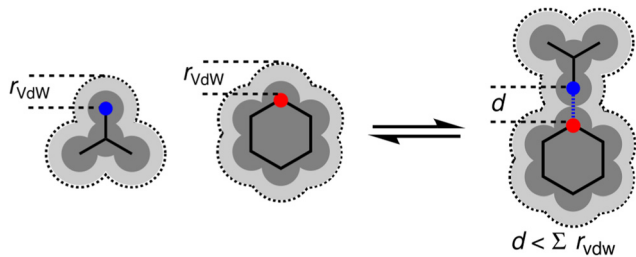


Fig. 11 Penetration of the van der Waals surface on formation of a H-bond.

### 3 Secondary electrostatic interactions in H-bond arrays

Experimentally determined association constants were collected from the literature for 37 different 1 : 1 complexes with multiple H-bonds measured in chloroform, dichloromethane or carbon tetrachloride.<sup>6,10,30–42</sup> The components of these complexes are all relatively rigid molecules, so that entropy changes due to conformational restriction make minimal contributions to the free energy changes associated with complexation. The geometries of the individual molecules were optimised using DFT. For each compound, the MEPS was calculated at the 0.0020, 0.0104 and 0.0300 e bohr<sup>-3</sup> electron density isosurfaces, and these surfaces were used to calculate H-bond parameters as explained below.

Two different methods were used to obtain H-bond parameters. In the first method, footprinting of the 0.0020 e bohr<sup>-3</sup> electron density isosurface was used to obtain a set of surface site interaction points (SSIP), and the SSIP closest in space to each H-bond donor and acceptor group was used as the H-bond parameter for that interaction site.<sup>4,43</sup> In the second method, eqn (3) and (4) were used in conjunction with the 0.0104 e bohr<sup>-3</sup> and the 0.0300 e bohr<sup>-3</sup> isosurfaces to determine H-bond parameters. Firstly, each point on the MEPS was assigned to the closest atom in space.<sup>44</sup> The maximum value on the MEPS assigned to each H-bond donor was then used to calculate the value of  $\alpha$  for that interaction site, and the minimum value assigned to each H-bond acceptor was used to calculate the value of  $\beta$ . For oxygen atoms with two well-defined minima on the MEPS, two values of  $E_{\min}$  corresponding to the two lone pair directions were considered.<sup>45</sup>

The association constant  $K$  for formation of a complex that makes  $N$  H-bonds was calculated using eqn (5). Each H-bond contributes a free energy change described by eqn (1), and an effective molarity (EM) term was added to account for the intramolecular nature of all interactions after the first intermolecular H-bond is made. The value of EM is not known, but we assume that a constant value can be used for all of these systems, because the molecules are relatively rigid and the H-bond arrays have similar geometries.

$$(-RT \ln K)/\text{kJ mol}^{-1} = -(N - 1)RT \ln \text{EM} + \sum_{i=1}^N [-(\alpha_i - \alpha_S)(\beta_i - \beta_S) + 6] \quad (5)$$

For some of the complexes used in this study, multiple binding modes are possible. For example, adenine can form

doubly H-bonded complexes of similar stability along the Watson–Crick or Hoogsteen faces. The overall association constant can be obtained by summing over all possible binding modes. In addition for complexes involving self-association, a statistical factor must be included to compare with the experimentally measured free energy change.<sup>46</sup> Eqn (6) was therefore used to obtain calculated values of  $\Delta G^\circ$  for comparison with experimental data.

$$\Delta G_{\text{calc}}^\circ/\text{kJ mol}^{-1} = -RT \ln \left( \sum K \right) + RT \ln \sigma \quad (6)$$

where  $\sigma$  is 2 for self-association and 1 for hetero-complexes.

For the two different sets of H-bond parameters described above, the value of EM was optimised to minimise the RMSE between the experimental values of  $\Delta G^\circ$  and the values calculated using eqn (6). The optimised value of EM did not show any significant dependence on the method used to calculate H-bond parameters: 93 M for H-bond parameters obtained from the van der Waals surface, and 125 M for H-bond parameters obtained from the higher density isosurfaces. However, there are dramatic differences in the results obtained using the two methods. Fig. 12(a) shows that H-bond parameters obtained from the van der Waals surface provide a poor description of the properties of multiply H-bonded complexes (RMSE = 9.8 kJ mol<sup>-1</sup>). Fig. 12(b) shows that H-bond parameters obtained from the higher density isosurfaces provide a good description of these systems (RMSE = 4.7 kJ mol<sup>-1</sup>). This method also performs better than direct calculation of complexation energies using *ab initio* methods (see ESI† for comparison with previous reports).<sup>47</sup>

The difference in the performance of the two methods can be directly related to the treatment of secondary electrostatic interactions. Fig. 13 compares the description of an ADAD-ADAD complex using H-bond parameters derived from different surfaces. The parameters for the two external H-bonds are similar in Fig. 13(a) and (b), but the parameters for the two internal H-bonds are quite different. Secondary electrostatic interactions lower the values of the H-bond parameters for the internal sites compared with the external sites, but in both cases the effects are much larger at the van der Waals surface (Fig. 13(a)). The result is an underestimate of the stability of the calculated values for all complexes containing alternating acceptors and donors in Fig. 12(a).

Similarly, Fig. 14 shows the H-bond parameters used to describe an AAA-DDD complex where the secondary electrostatic interactions are attractive. The H-bond donor and acceptor parameters that describe the central H-bond are elevated in both descriptions, but the effects are much larger at the van der Waals surface (Fig. 14(a)). This effect leads to an overestimate of the stability of the calculated values for all complexes containing attractive secondary electrostatic interactions in Fig. 12(a).

The method for determining H-bond parameters using higher electron density isosurfaces clearly provides an excellent quantitative description of the secondary electrostatic interactions that determine the properties of H-bond arrays. There is



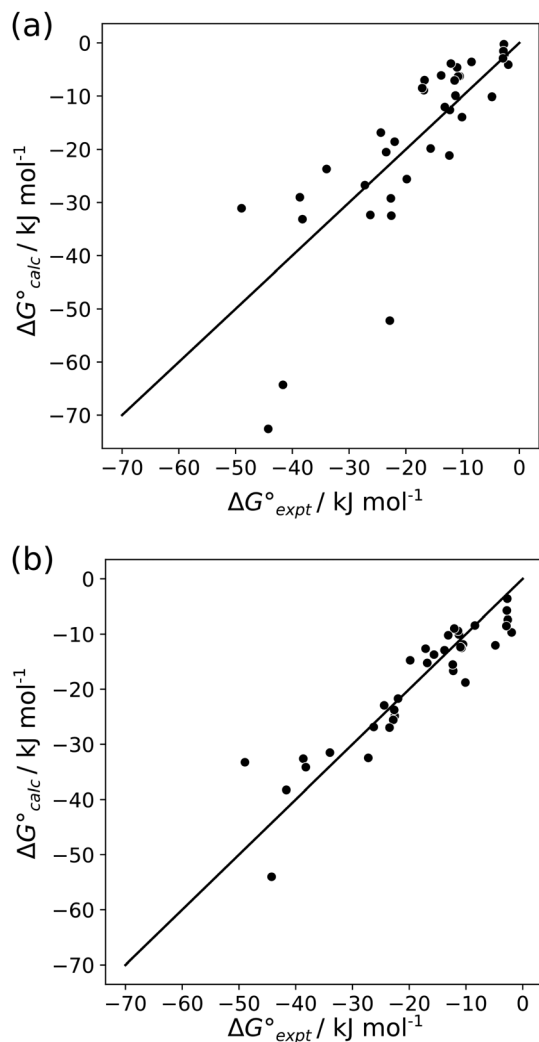


Fig. 12 Comparison of experimentally measured free energy changes for formation of multiply H-bonded complexes ( $\Delta G^{\circ}_{\text{expt}}$ ) with calculated values ( $\Delta G^{\circ}_{\text{calc}}$ ) obtained from eqn (5) and (6) and H-bond parameters calculated on (a) the van der Waals surface, and (b) at the higher electron density isosurfaces ( $0.0300 \text{ e bohr}^{-3}$  for  $\beta$  and  $0.0104 \text{ e bohr}^{-3}$  for  $\alpha$ ). The black lines are  $y = x$ .

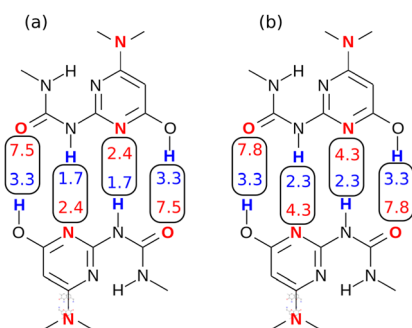


Fig. 13 H-bond parameters for an ADAD-ADAD complex calculated on (a) the van der Waals surface, and (b) the higher electron density isosurfaces ( $0.0300 \text{ e bohr}^{-3}$  for  $\beta$  and  $0.0104 \text{ e bohr}^{-3}$  for  $\alpha$ ).

one outlier in Fig. 12(b), which is the ADDA-DAAD complex shown in Fig. 15(a). This complex has a very similar structure to

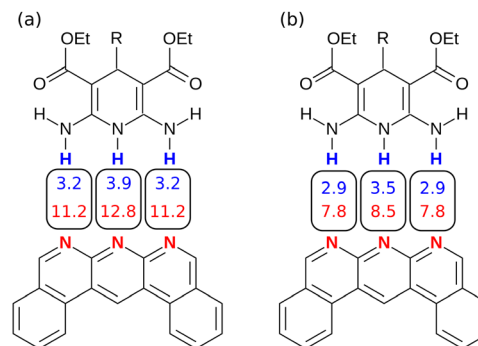


Fig. 14 H-bond parameters for an AAA-DDD complex calculated on (a) the van der Waals surface, and (b) the higher electron density isosurfaces ( $0.0300 \text{ e bohr}^{-3}$  for  $\beta$  and  $0.0104 \text{ e bohr}^{-3}$  for  $\alpha$ ).

another ADDA-DAAD complex, which is shown in Fig. 15(b). The reported association constants are  $3.8 \times 10^8 \text{ M}^{-1}$  and  $6 \times 10^6 \text{ M}^{-1}$ , but the values predicted using eqn (5) are  $7 \times 10^6 \text{ M}^{-1}$  for both complexes<sup>10,30</sup> so this outlier may be an artefact. Removing this outlier reduces the RMSE for the results shown in Fig. 12(b) from 4.7 to 4.0  $\text{kJ mol}^{-1}$ .

## 4 Discussion

The results shown in Fig. 12(b) demonstrate that H-bond parameters calculated from the MEPS of the individual molecules provide a quantitative description of the contributions of secondary electrostatic interactions to the stabilities of complexes that feature multiple interactions. An advantage of the approach is that these H-bond parameters include the effects of neighbouring groups due to through-bond polarisation as well as through-space electrostatic interactions. Secondary electrostatic interactions will perturb interactions with the solvent as well as interactions with the binding partner, and implementation of the H-bond parameters described here have the effects of desolvation implicitly built into the approach through eqn (5).

It is instructive to compare the treatment of secondary electrostatic interactions based on H-bond parameters with the original molecular mechanics interpretation. Fig. 16(a) illustrates an atom-centred point charge representation of secondary electrostatic interactions. The total Coulombic interaction  $E$  in Fig. 16(b) is the sum of the short range H-bonding interaction and the long-range secondary electrostatic

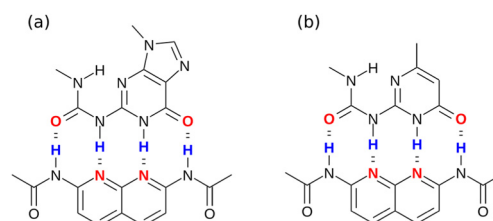


Fig. 15 ADDA-DAAD complexes. (a)  $K = 3.8 \times 10^8 \text{ M}^{-1}$  in chloroform<sup>10</sup> (b)  $K = 6 \times 10^6 \text{ M}^{-1}$  in chloroform.<sup>30</sup>



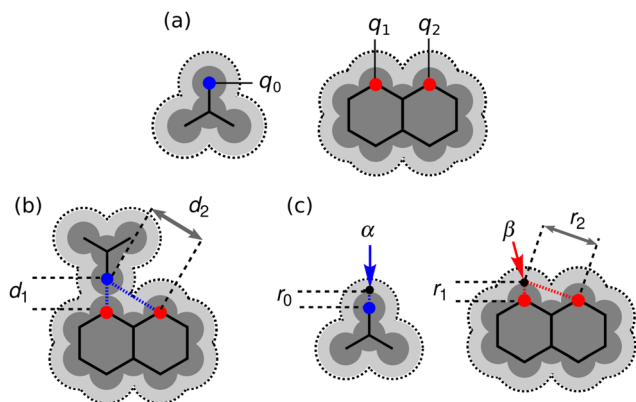


Fig. 16 (a) Representation of molecular charge distributions as point charges. (b) Coulombic interactions between a H-bond donor and a molecule with two H-bond acceptors. (c) The H-bond donor ( $\alpha$ ) and acceptor ( $\beta$ ) parameters for the two compounds are proportional to the molecular electrostatic potential calculated on surfaces defined by the dark grey regions. The pale grey surfaces represent the van der Waals surface.

interaction (eqn (7)).

$$E = \frac{q_0 q_1}{d_1} + \frac{q_0 q_2}{d_2} \quad (7)$$

Fig. 16(c) shows how the H-bond parameters for the two compounds would be calculated using an atom-centred point charge representation of the molecular charge distribution. Since the values of  $\alpha$  and  $\beta$  are linearly related to the electrostatic potential calculated at a distance  $r$  from the nucleus, which is defined by the choice of isosurface, the H-bond donor and acceptor parameters are given by eqn (8) and (9) respectively.

$$\alpha \propto \frac{q_0}{r_0} \quad (8)$$

$$\beta \propto \frac{q_1}{r_1} + \frac{q_2}{r_2} \quad (9)$$

The energy associated with formation a H-bond between the two molecules is proportional to the product  $\alpha\beta$ . By multiplying

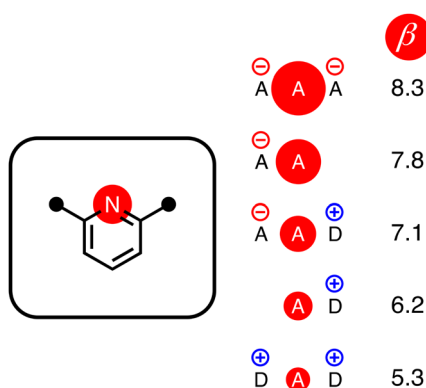


Fig. 17 The effect of secondary electrostatic interactions on the H-bond acceptor parameters of the nitrogen atoms of pyridines in different molecular environments. The value of  $\beta$  for unsubstituted pyridine is 7.2.

eqn (8) and (9), we obtain an expression with two terms: one proportional to  $q_0 q_1$ , which represents the primary H-bonding interaction, and one proportional to  $q_0 q_2$ , which represents the secondary electrostatic interaction. This analysis demonstrates that the approach developed here based on the MEPS of individual molecules is exactly analogous to the original interpretation based on intermolecular electrostatic interactions developed by Jorgensen and Pranata. Indeed the qualitative ideas that stem from the original secondary electrostatic approach can be applied equally well to an approach based on H-bond parameters. Fig. 17 illustrates how the properties of H-bond arrays can be understood based on the H-bond parameters of individual molecules. The average values of the H-bond acceptor parameter  $\beta$  calculated for pyridine-type acceptors in different molecular environments are shown in Fig. 17. A H-bond acceptor next to the nitrogen atom increases the  $\beta$  value, but a neighbouring H-bond donor group decreases the  $\beta$  value. The effects are roughly additive, so we can think about how secondary electrostatic interactions affect the local MEPS within a molecule in just the same way as we would think about secondary electrostatic interactions between two different molecules.

## 5 Conclusion

In this paper we have studied the through-space effects of neighbouring polar groups on the MEP at different electron density surfaces. We conclude that the effect is exaggerated at the van der Waals surface, but inner surfaces provide a much better description of the electrostatics of polar molecules. We selected the  $0.0300 \text{ e bohr}^{-3}$  surface for obtaining  $\beta$  parameters and the  $0.0104 \text{ e bohr}^{-3}$  surface for obtaining  $\alpha$  parameters. Moreover, we have developed a method to predict the complexation energies of multiply hydrogen bonded systems. We show that the through-space effects at the inner surfaces quantitatively predict secondary interactions. Ultimately we show that the secondary interaction term can be obtained as an inherent property of the MEP of an individual molecule, and does not depend on the identity of the functional groups on the other molecule.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank Dr A. Bond for collecting and analysing X-ray diffraction data. We thank Dr D. P. Reynolds for the stimulating discussions. We thank the Herchel Smith Fund for financial support.

## Notes and references

- W. M. Latimer and W. H. Rodebush, *J. Am. Chem. Soc.*, 1920, **42**, 1419–1433.



- 2 L. Pauling and L. O. Brockway, *Proc. Natl. Acad. Sci. U. S. A.*, 1934, **20**, 336–340.
- 3 R. Cabot, C. A. Hunter and L. M. Varley, *Org. Biomol. Chem.*, 2010, **8**, 1455–1462.
- 4 C. S. Calero, J. Farwer, E. J. Gardiner, C. A. Hunter, M. Mackey, S. Scuderi, S. Thompson and J. G. Vinter, *Phys. Chem. Chem. Phys.*, 2013, **15**, 18262–18273.
- 5 Y. Kyogoku, R. Lord and A. Rich, *Biochim. Biophys. Acta*, 1969, **179**, 10–17.
- 6 Y. Kyogoku, R. C. Lord and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1967, **57**, 250–257.
- 7 W. L. Jorgensen and J. Pranata, *J. Am. Chem. Soc.*, 1990, **112**, 2008–2010.
- 8 N. Mougin, A. Livoreil and J. Mondet, *EU Pat.*, EP1392222B1, 2006.
- 9 G. M. L. H. van Gemert, S. Chodorowski-Kimmes, H. M. Janssen, E. W. Meijer and A. W. Bosman, *Int. Pat.*, WO2010002262A1, 2010.
- 10 T. Park, S. C. Zimmerman and S. Nakashima, *J. Am. Chem. Soc.*, 2005, **127**, 6520–6521.
- 11 P. L. A. Popelier and L. Joubert, *J. Am. Chem. Soc.*, 2002, **124**, 8725–8729.
- 12 W. E. V. Narváez, E. I. Jiménez, E. Romero-Montalvo, A. S. de la Vega, B. Quiroz-García, M. Hernández-Rodríguez and T. Rocha-Rinza, *Chem. Sci.*, 2018, **9**, 4402–4413.
- 13 W. E. V. Narváez, E. I. Jiménez, M. Cantú-Reyes, A. K. Yatsimirsky, M. Hernández-Rodríguez and T. Rocha-Rinza, *Chem. Commun.*, 2019, **55**, 1556–1559.
- 14 C.-H. Wu, Y. Zhang, K. van Rickley and J. I. Wu, *Chem. Commun.*, 2018, **54**, 3512–3515.
- 15 A. Asensio, N. Kobko and J. J. Dannenberg, *J. Phys. Chem. A*, 2003, **107**, 6441–6443.
- 16 S. C. C. van der Lubbe and C. Fonseca Guerra, *Chem. – Eur. J.*, 2017, **23**, 10249–10253.
- 17 W. Hujo and S. Grimme, *J. Chem. Theory Comput.*, 2012, **9**, 308–315.
- 18 C. A. Hunter, *Angew. Chem., Int. Ed.*, 2004, **43**, 5310–5324.
- 19 R. F. W. Bader, W. H. Henneker and P. E. Cade, *J. Chem. Phys.*, 1967, **46**, 3341–3363.
- 20 S. C. C. van der Lubbe, F. Zaccaria, X. Sun and C. Fonseca Guerra, *J. Am. Chem. Soc.*, 2019, **141**, 4878–4885.
- 21 A. Oliver, C. A. Hunter, R. Prohens and J. L. Rosselló, *J. Comput. Chem.*, 2018, **39**, 2371–2377.
- 22 M. M. Cetin, R. T. Hodson, C. R. Hart, D. B. Cordes, M. Findlater, D. J. C. Jr., A. F. Cozzolino and M. F. Mayer, *Dalton Trans.*, 2017, **46**, 6553–6569.
- 23 D. J. Sutor, *Nature*, 1962, **195**, 68–69.
- 24 M. Berthelot, C. Laurence, M. Safar and F. Besseau, *J. Chem. Soc., Perkin Trans. 2*, 1998, 283–290.
- 25 S. J. Pike, J. J. Hutchinson and C. A. Hunter, *J. Am. Chem. Soc.*, 2017, **139**, 6700–6706.
- 26 R. Cabot and C. A. Hunter, *Org. Biomol. Chem.*, 2010, **8**, 1943–1950.
- 27 P. W. Kenny, *J. Chem. Inf. Model.*, 2009, **49**, 1234–1244.
- 28 P. W. Kenny, *J. Chem. Soc., Perkin Trans. 2*, 1994, 199.
- 29 Z. Rahim and B. N. Barman, *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.*, 1978, **34**, 761–764.
- 30 T. F. A. de Greef, G. Ercolani, G. B. W. L. Ligthart, E. W. Meijer and R. P. Sijbesma, *J. Am. Chem. Soc.*, 2008, **130**, 13755–13764.
- 31 F. H. Beijer, R. P. Sijbesma, J. A. J. M. Vekemans, E. W. Meijer, H. Kooijman and A. L. Spek, *J. Org. Chem.*, 1996, **61**, 6371–6380.
- 32 Y. Kyogoku, R. C. Lord and A. Rich, *Nature*, 1968, **218**, 69–72.
- 33 S. E. Krikorian, *J. Phys. Chem.*, 1982, **86**, 1875–1881.
- 34 T. J. Murray and S. C. Zimmerman, *J. Am. Chem. Soc.*, 1992, **114**, 4010–4011.
- 35 P. L. Wash, E. Maverick, J. Chiefari and D. A. Lightner, *J. Am. Chem. Soc.*, 1997, **119**, 3802–3806.
- 36 B. A. Blight, A. Camara-Campos, S. Djurdjevic, M. Kaller, D. A. Leigh, F. M. McMillan, H. McNab and A. M. Z. Slawin, *J. Am. Chem. Soc.*, 2009, **131**, 14116–14122.
- 37 V. Hegde, P. Madhukar, J. D. Madura and R. P. Thummel, *J. Am. Chem. Soc.*, 1990, **112**, 4549–4550.
- 38 R. P. Sijbesma and E. W. Meijer, *Chem. Commun.*, 2002, 5–16.
- 39 T. W. Bell and J. Liu, *J. Am. Chem. Soc.*, 1988, **110**, 3673–3674.
- 40 Y. Ducharme and J. D. Wuest, *J. Org. Chem.*, 1988, **53**, 5787–5789.
- 41 S. H. M. Söntjens, R. P. Sijbesma, M. H. P. van Genderen and E. W. Meijer, *J. Am. Chem. Soc.*, 2000, **122**, 7487–7493.
- 42 J. Hine, S. Hahn and J. Hwang, *J. Org. Chem.*, 1988, **53**, 884–887.
- 43 M. D. Driver, M. J. Williamson, N. D. Mitri, T. Nikolov and C. A. Hunter, *J. Chem. Inf. Model.*, 2021, **61**, 5331–5335.
- 44 J. L. Bentley, *Commun. ACM*, 1975, **18**, 509–517.
- 45 P. Murray-Rust and J. P. Glusker, *J. Am. Chem. Soc.*, 1984, **106**, 1018–1025.
- 46 C. Hunter and H. Anderson, *Angew. Chem., Int. Ed.*, 2009, **48**, 7488–7499.
- 47 O. Lukin and J. Leszczynski, *J. Phys. Chem. A*, 2002, **106**, 6775–6782.

