Charting Hydrogen Bond Anisotropy

Diogo Santos-Martins¹ and Stefano Forlì^{1,✉}

¹Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, 92037, California, USA

Hydrogen bond (HB) is an essential interaction in countless phenomena, and regulates the chemistry of life. HBs are characterized by two main features, strength and directionality, with a high degree of heterogeneity across different chemical groups. These characteristics are dependent on the electronic configuration of the atoms involved in the interaction, which, in turn, is influenced strongly by the molecular environment where they are found. Studies based on the analysis of HB in solid phase, such as X-ray crystallography, suffer from significant biases due to the packing forces. These will tend to better describe strong HBs at the expenses of weak ones, which are either distorted or under-represented. Using quantum mechanics (QM), we calculated interaction energies for about a hundred acceptor and donors, in a rigorously defined set of geometries. We performed about 180,000 independent QM calculations, covering all relevant angular components, and mapping strength and directionality in a context free from external biases, with both single-site and cooperative HBs. We show that by quantifying directionality, there is not correlation with strength, and therefore these two components need to be addressed separately. Results demonstrate that there are very strong HB acceptors (e.g., DMSO) with nearly isotropic interactions, and weak ones (e.g., thioacetone) with a sharp directional profile. Similarly, groups can have comparable directional propensity, but be very distant in the strength spectrum (e.g., thioacetone and pyridine). These findings have implications for biophysics and molecular recognition, providing new insight for chemical biology, protein engineering, and drug design. The results require rethinking the way directionality is described, with implications for the thermodynamics of HB.

hydrogen bond | drug design | force field | non-covalent interaction
Correspondence: forli@scripps.edu

Introduction

Hydrogen bond (HB) plays a role in countless phenomena and has been studied in several fields within chemistry and biology. Arguably, it is the main force that drives reversible interactions in biological contexts^{1,2}, being responsible for the physicochemical properties of water^{3,4}, orchestrating structural arrangement of proteins⁵, and stabilizing transition states in natural⁶ and engineered catalysts⁷. The hallmark of HB is its directionality: its strength can be heavily modulated by the geometries of acceptor and donor with respect to each other; the presence of directionality is used to distinguish HB from van der Waals interactions^{8,9}. Directionality has been studied by statistical analysis of the Cambridge Structural Database (CSD)¹⁰ and the Protein Data Bank (PDB)¹¹. These analyses are based on the assumption that HB strength and directionality can be inferred from structures because “interaction geometries are conserved with fidelities that correspond to their

strengths”^{9,12}. Statistical distributions of such observed geometries can be then translated into chemical potentials^{13,14}. Over the past decades, several chemical groups were studied using statistical approaches, including halogens^{15–17}, sulfur and oxygen^{18–20}, weak and strong donors⁸, among others^{17,21–25}.

Despite providing invaluable information, retrospective structural analysis has several issues that limit the coverage of HB properties. First, it is chemically inhomogeneous because several chemical groups are under- or overrepresented. Furthermore, structural patterns of HBs can be smeared by factors such as steric hindrance⁸, packing forces²⁶, or be biased by the refinement programs used (which often include specific constraints for HB)²⁷.

Quantum mechanics (QM) calculations, while requiring more resources than statistical surveys, have been used to characterize various HB properties^{26,28–35}, usually with good agreement with observations¹³. Among them, very few calculated interaction energies at non-equilibrium geometries (the ones that did are summarized in Supplementary Table S1). These few studies lack chemical diversity, as relevant chemical groups such as halogens, sulfur or phosphate oxyacids have never been characterized. Furthermore, with respect to the geometrical degrees of freedom, most studies performed one-dimensional scans along a given interaction angle, and only a few considered more than one angle (Supplementary Table S1).

This work is an effort towards a systematic and comprehensive characterization of HB directionality. We have performed about 180,000 energy QM calculations on a broad and diverse panel of chemical groups of relevance in biological and drug design contexts, and performed volumetric energy scans to address multiple angular components, as well as the distance at which the interaction is established. Strength and directionality have been measured and discussed.

Results

The dataset contains 34 donors (Fig.S1) and 67 acceptors (Fig.S2, Methods). For the entire dataset, calculations were performed using B3LYP density functional with the cc-pVDZ basis set and the counterpoise correction, while MP2/aug-cc-pVTZ with counterpoise correction was used for smaller subsets of complexes. The validation of these two methods, their differences and their effect on the results has been discussed in *Supplementary Methods*.

Strength. We measured the strength of HBs on the 34 donors (Fig.S1) and 67 acceptors, performing B3LYP/cc-pVDZ calculations of donor-acceptor dimers interacting in optimal ge-

ometries. Interaction energies (ΔE) between a probe P and the target T were calculated as:

$$\Delta E = E_{P,T} - [E_P + E_T] \quad (1)$$

where $E_{P,T}$, E_P , and E_T are the energies of the probe-target dimer, the probe, and the target, respectively. Interaction energies at optimal geometries indicate the maximum strength of a given HB pair.

For a subset of 20 donors and 44 acceptors, representative of the strength spectrum considered, we enumerated all pairwise combinations at optimal geometries (880 dimers), generating the panel in Fig.S9. Interaction energies present a systematic and continuous trend, showing that relative strength ranking of a monovalent HB group is independent of the molecule used to probe it.^{36–39} For this reason, we chose ΔE values calculated with hydrogen cyanide (HCN) to rank the strengths of all donors (as compound **61**) and acceptors (as compound **24**). HCN is a linear molecule, sterically unhindered and highly polarized, minimizing the relative contribution of dispersion interactions and steric clashes. Interaction energies calculated *in vacuo* allow a raw assessment of their competitive strength with respect to water, which was expressly included in the dataset (as compounds **15** and **75**). Overall, strengths of neutral acceptors and donors in the dataset ranges from 0.3 to 1.7 times that of water (Fig.S7 and Fig.S8). Charged groups show a 4- to 10-fold increase with respect to their corresponding neutral moieties.

Donor strength (Fig.S7) correlates with high electronegativity and low oxydation state of the heavy atom to which hydrogen is bound, and is further modulated by the groups attached to it. Sulfur and carbon donors are weaker donors than water (with the notable exception of HCN **61**), while nitrogen- and oxygen-containing donors are generally stronger. Acceptor strength (Fig.S8) depends on the availability of electrons to interact with the donor proton, therefore it is modulated by atoms directly bound to the acceptor atom, and the degree in which electrons are shared.

When the same group can act as both acceptor and donor, the respective strengths tend to be negatively correlated. For example, the hydroxyl in phenol is stronger than water when acting as a donor, but weaker when acting as acceptor (compounds **26** and **55**, respectively, Fig.S9). The same principle applies to the chemical environment of HB capable groups. Electron-withdrawing substituents decrease acceptor strength, but increase donor strength.

Amino acid ranking. The HB strength calculated for groups contained in amino acid side chains can be used to assess their relative strength. Even though in biological macromolecules their physico-chemical properties will depend on the local environment resulting from folding, these values can provide a rough but insightful gauge of their HB capabilities. According to our calculations, the relative strength of HB donors is: Lys⁺ (**34**), His⁺ (**32**), Arg⁺ (**31**) $\gg$ Glu/Asp (**27**), Tyr (**26**) $>$ His (**25**) $>$ Trp (**22**) $>$ Gln/Asn (**20**), backbone amide (**18**) $>$ Ser/Thr (**16**) $>$ water (**15**) $>$ Arg (**13**) $>$ Lys (**6**) $>$ Cys (**3**). For HB acceptors the rank is: Asp⁻/Glu⁻ (**101**) $\gg$ Arg (**95**) $>$ His (**92**), Lys (**91**), backbone amide

Fig. 1 Methodology to probe and visually represent anisotropy. **a** Definition of spherical coordinates (r , θ , ϕ) for sampling interaction energies. **b** Interaction dome. Each sphere corresponds to a position of the probe atom for which an interaction energy was calculated. **c** Lambert azimuthal equal-area projection (see details in the Methods).

oxygen (**90**) $>$ Asn/Gln (**85**) $>$ water (**75**) $>$ Asp/Glu carbonyl (**74**) $>$ Ser/Thr (**71**) $\gg$ Met (**56**), Tyr (**55**), Cys (**53**) $>$ Asp/Glu hydroxyl (**43**), disulfide (**42**), Phe (**41**).

Anisotropy. We studied the anisotropy of 9 donors and 39 acceptors representative of the chemical space and strength spectrum of the full dataset (see Fig. 2 and Supplementary Figs. S4, S5, S10, S11, S12, S13, S14, S15, S16, S17 and S18).

Starting from the optimized dimer coordinates, possible geometries around HB minima were explored for each target atom, mapping variations of ΔE as a function of the probe position (Fig. 1a). Interaction energies at non-optimal geometries reveal the loss of energy associated with each angular component, i.e. the anisotropy profile. For each target, which can be either a donor or an acceptor, we built a tridimensional map of the interactions, called “dome” (Fig. 1b). Domes were built by moving the probe in a spherical coordinate system (r , θ , ϕ) centered on the target atom (Fig. 1a). To visualize the relationship between interaction energy and angular components, we plotted ΔE as a function of θ and ϕ using the Lambert azimuthal equal-area projection (Fig. 1c). For each point in the surface defined by θ and ϕ the dome point associated with the optimal distance r (i.e., with the lowest ΔE) was chosen.

Anisotropy was quantified by defining the specific directionality index for a given probe as following:

$$D_{[probe]} = \frac{1}{N} \sum_p \frac{\Delta E_{opt} - \Delta E_p}{\Delta E_{opt}} \quad (2)$$

where N is the total number of accessible dome points (i.e., where interaction energy $\Delta E \leq 0$), ΔE_{opt} is the best energy in the dome (i.e., at the optimal geometry), and ΔE_p is the interaction energy with the probe at the point p . Very directional HBs have $D_{[probe]}$ values closer to 1, while isotropic ones (or dimers interacting mostly via dispersion interactions) have values approaching 0. To quantify the average energy loss at any non-optimal geometry point, we defined the isotropy error $\varepsilon_{[probe]}$ as the product of $D_{[probe]}$ with its optimal interaction energy with the probe:

$$\varepsilon_{[probe]} = D_{[probe]} \times \Delta E_{opt[probe]} \quad (3)$$

The anisotropy metrics $D_{[probe]}$ and $\varepsilon_{[probe]}$ provide complementary information. While $\varepsilon_{[probe]}$ is expressed in kcal/mol,

Fig. 2 Domes of a series of acceptor groups (water **75**, thioacetone **57**, *N*-methyl acetamide **90**, DMSO **94**, trimethylamine **78** and imidazole **92**). **a** Dome points for which ΔE is $< 80\%$ of ΔE_{opt} . **b** Interaction domes. **c** Equal area projections (θ , ϕ) at optimal r distance. Each dome point in **a** and **b** corresponds to a position sampled by the donor probe atom, and is colored according to the normalized $\Delta E/\Delta E_{opt}$ value. Points and areas with positive ΔE values (i.e. repulsions) are transparent (**a**, **b**) or white (**c**).

$D_{[probe]}$ is dimensionless and normalized, making it insensitive to the magnitude of ΔE .

Both metrics are independent from the absolute amount of surface area accessible to the probe. When two target atoms have similar $D_{[probe]}$ values, they share the same fraction of accessible surface showing anisotropy for a HB probe. Furthermore, these metrics are insensitive to the derivative of the interaction energy with respect to angular components. For example, methylamine (**91**), dimethylamine (**83**) and trimethylamine (**78**) show progressively smaller accessible surfaces due to the increasing hindrance of the methyl groups, as well as a steeper change in interaction energy with an increase in the ϕ angle (Supplementary Fig. S11), but $D_{[probe]}$ values remain almost identical (Table 1). This happens because the surface associated to ΔE_{opt} values is reduced as well, due to the diminished HB capable volume, which preserves the same ratio with respect to the total accessible surface.

The calculated values of $D_{[probe]}$ were then used to analyze the molecules in our dataset, and characterize their anisotropy. Similarly to what done for the strength analysis, we selected probes with the least hindered geometries, as well as the strongest HB interactions, in order to increase the magnitude of interaction energies and minimize the relative contribution of dispersion interactions. Therefore, HCN (**61**) was selected to probe donors, and HF (**30**) to probe acceptors. Donors show rather uniform anisotropies, with $D_{[HCN]}$ ranging between 0.3 and 0.5, which depend almost entirely on the ϕ angle (i.e., the D-H...A angle), with minor variations induced by the different molecular shapes (Supplementary Figs. S10 and S5). Acceptors, on the other hand, present a much larger spectrum of variations, from nearly isotropic to

very directional (Fig. 2). Variations in closely related chemical groups show that the broader molecular environment of the acceptor atom, up to two covalent bonds distance, must be considered (Table 1). For example, replacing one of the methyl groups of acetone **67** ($D_{[HF]} = 0.29$) with nitrogen in amides (**80**, **85**, **87**) reduces sp^2 oxygen $D_{[HF]}$ to 0.21–0.24 (Fig. 2). Secondary amide **90** is the least directional carbonyl (Table 1), and the most tolerant to angular deviations, retaining 80% of ΔE_{opt} at $\phi = 0$ while acetone **67** retains less than 75% (Fig. 2). Variations of θ (i.e., out-of-plane) are also more tolerated in amides than in acetone, with strengths of 50% and 30% of those at optimal geometries, respectively, when both θ and ϕ approach $\pm 90^\circ$ (Fig. 2). Such overall increased tolerance is likely responsible of the high propensity of proteins backbone to form stable HBs that are co-linear with the C=O bond such as in α -helices^{5,25}.

When oxygen is bound directly to sulfur, as in DMSO **94**, the partial triple bond characteristics⁴⁰ increases acceptor strength to about 140% that of water, and further reduce directionality ($D_{[HF]} = 0.19$, Fig. 2) due to the lone pair delocalization. Similar effect is observed in sulfate **58**, sulfonamide **65**, sulfones (**76**, **77**) and phosphorous **89**, with directionality ranging between 0.20 and 0.16 (Table 1, Fig. S13), and strength between 74% to 120% of that of water. Anisotropy profiles of sulfur and phosphate oxyacids (Supplementary Fig. S13), appear compatible with hydration patterns identified experimentally^{41,42}.

Sulfur itself is a weaker acceptor than oxygen but shows the same anisotropy change observed between acetone and amides. From thioacetone **57** to *N*-methylthioacetamide **62** (Table 1, Fig. S12), the increase in favorable out-of-plane interactions (i.e., more isotropic interaction) is accompanied by

Table 1 Directionality of HB acceptors probed with HF ranked by their $D_{[HF]}$ values, their best interaction energy ($\Delta E_{opt}(HF)$), and quantitative directionality error $\varepsilon_{[probe]}$

	$D_{[HF]}$	$\Delta E_{opt}(HF)$	$\varepsilon_{[HF]}$	element	hybridization
hydrogen fluoride (52)	0.08	-5.0	-0.39	F	sp^3
fluoromethane (50)	0.10	-5.0	-0.48	C	sp^3
acetate (101)	0.11	-26.3	-2.74	O	ionic
methanethiolate (100)	0.13	-25.2	-3.25	S	ionic
methanesulfonamide (65)	0.16	-8.0	-1.20	O	$sp^{2(*)}$
dimethyl disulfone (77)	0.16	-8.8	-1.32	O	$sp^{2(*)}$
dimethyl sulfoxide (94)	0.18	-12.3	-2.27	O	$sp^{2(*)}$
trimethyl phosphate (89)	0.19	-10.6	-1.93	O	$sp^{2(*)}$
dimethyl sulfate (58)	0.20	-7.1	-1.30	O	$sp^{2(*)}$
fluorodimethylborane (49)	0.20	-4.6	-0.90	F	sp^3
N-methylacetamide (90)	0.21	-10.6	-2.15	O	sp^2
acetamide (85)	0.24	-10.9	-2.66	O	sp^2
N,N-dimethylacetamide (88)	0.25	-10.2	-2.60	O	sp^2
methyl acetate (72)	0.25	-8.2	-2.17	O	sp^2
TMAO (97)	0.25	-17.2	-4.28	O	sp^2/sp^3
benzamide (80)	0.27	-9.9	-2.77	O	sp^2
nitrobenzene (59)	0.29	-6.4	-1.80	O	sp^2/sp^3
acetone (67)	0.29	-9.0	-2.56	O	sp^2
hydrogen cyanide (61)	0.29	-6.5	-1.89	N	sp
water (75)	0.34	-8.6	-2.98	O	sp^3
methyl isocyanide (73)	0.36	-8.5	-3.02	C	sp
N-methylthioacetamide (62)	0.36	-8.3	-2.95	S	sp^2
methanol (71)	0.36	-8.4	-3.07	O	sp^3
dimethyl ether (63)	0.41	-8.1	-3.29	O	sp^3
phenol (55)	0.42	-6.2	-2.58	O	sp^3
acetic acid (43)	0.44	-5.3	-2.27	O	sp^3
imidazole (92)	0.45	-11.6	-5.20	N	aromatic
furan (45)	0.46	-4.7	-2.15	O	aromatic
dimethyl disulfide (42)	0.47	-5.2	-2.12	S	sp^3
methyl acetate (46)	0.48	-5.7	-2.66	O	sp^3
dimethyl sulfide (56)	0.48	-6.5	-3.04	S	sp^3
pyridine (82)	0.53	-10.6	-5.64	N	aromatic
methanethiol (53)	0.50	-5.7	-2.82	S	sp^3
thioacetone (57)	0.50	-7.2	-3.49	S	sp^2
N-methylacetoneimine (86)	0.51	-12.3	-6.26	N	sp^2
methanesulfonamide (48)	0.53	-7.0	-3.68	N	sp^2/sp^3
dimethylamine (83)	0.54	-11.6	-6.24	N	sp^3
trimethylamine (78)	0.55	-11.0	-6.04	N	sp^3
methylamine (91)	0.55	-12.0	-6.60	N	sp^3

(*): partial sp bond

an increase in the magnitude of ΔE_{opt} .

Overall, nitrogen has a sharper directional character than oxygen as its optimal interaction regions more precisely defined (Figs. 3 and S11). Its relatively limited range of variations across the dataset suggests less dependency on hybridization or molecular environment. Aromatic and sp^2 nitrogens show minor broadening of the optimal interaction regions, above and below planes of the ring (see imidazole **92** in Fig. 2 and pyridine **82**, *N*-methylacetoneimine **86** and guanidine **95** in Supplementary Fig. S16). Interestingly, the nitrogen in sulfonamide **48** shows a very high $D_{[HF]}$ value similar to amines (**91**, **83**, **78**) despite being a weaker HB acceptor (< 75% the strength of amines).

In charged species, such as thiolate **100** and carboxylate **101** (Supplementary Fig. S17), the increased acceptor strength is associated with a sharp $D_{[HF]}$ reduction with respect to their corresponding neutral forms, and independently from the initial anisotropy: from methanethiol **53** to thiolate **100**,

$D_{[HF]}$ drops from 0.5 to 0.13. from carboxylic acid **74** to carboxylate **101**, $D_{[HF]}$ drops from 0.24 to 0.11. Additional acceptor directionality patterns are shown in Supplementary Figs. S14, S15 and S16.

Anisotropy in cooperative HBs. Molecules in the condensed phase tend to be more polarized than in the gas phase⁴³ because of the electrostatic stabilization provided by the adjacent molecules. This cooperative effect plays an important role for HB networks where molecules act as donors and acceptors simultaneously⁴⁴.

To study the impact of cooperativity on the directionality profile of donor and acceptor sites, we considered three molecules capable of forming multiple HBs: water (**15** and **75**), imidazole (**25** and **92**) and *N*-methylacetamide (**18** and **90**). To each target molecule (*T*) we added a water molecule (*W*) to one donor or acceptor site (Fig. S18), then we adapted eq.1 to calculate the interaction energy ΔE_{wat} with the probe

P at the free site:

$$\Delta E_{wat} = E_{PTW} - [E_P + E_{TW}] - \Delta E_{PW} \quad (4)$$

where E_{PTW} , E_P , and E_{TW} are, respectively, the absolute energies of the probe-target-water trimer, the probe alone, and the target-water dimer. ΔE_{PT} is the interaction energy between the probe and the water, which can be expanded as $E_{PW} - [E_P + E_W]$. The polarization induced by the water molecule⁴⁴ affects strengths of both probe-target and target-water interactions. The result is a strength increase for all six sites considered, with ΔE_{wat} approximately 3% to 25% larger than the corresponding ΔE_{opt} of dimers without the added waters (Supplementary Tables S3 and S4). On the other hand, directionality mapped using the same scheme applied for dimers in the gas phase showed no significant changes, with profiles that were virtually indistinguishable from the non-cooperative ones (Fig. S18). The results suggest that anisotropy on the gas phase is transferable to the condensed phase.

Discussion

Anisotropy is one of the key features of HB^{45,46} but its directional preferences are often overlooked in recent literature. While many optimal geometry configurations have been described^{47–49}, there is still considerable uncertainty on how strength decays when deviating from these conditions, especially for acceptors. For example, the interaction angle at the acceptor atom is not considered as a criterion for HB identification, while the donor angle is⁹. While it is true that some acceptors are nearly isotropic, others show strong directional preferences.

Retrospective statistical analyses on structural databases are arguably more biased against weak HB interactions, which are either under-⁴⁶ or mis-represented by the influence of stronger forces. This is particularly relevant in the constrained conditions of solid state^{9,46}.

The focus of this work was to characterize the HB properties of wide range of chemical groups capable of HB interactions, especially of the interplay between strength and directionality. In order to do that, we designed a method to quantify directionality, which allowed us to compare such properties between different chemical groups. QM calculations were used to characterize variations of HB interaction energies at high level of detail, while minimizing interferences from non-HB interactions, and free from the packing bias of the condensed phase state.

The initial strength assessment showed a broad and robust trend spanning the whole dataset, which confirmed the additive nature of the interaction in the gas phase and highlighted the transferability properties of single, non-cooperative HBs. That is, while overall strength is the result of respective donor and acceptor strengths, their individual strengths are independent of each other. Calculations for most of the molecules were performed *in vacuo*, but water itself was part of the dataset to provide a reference value for the assessment of strengths of the other molecules with respect to the biological solvent. Whether or not two solute molecules can form

a stable HB in the presence of solvent depends on the relationship between relative strengths of solute-solute, solute-solvent and solvent-solvent interactions. For that, gas phase interaction energies normalized with respect to water can be used to provide insight on single HBs between solute and solvent, and assess their competitive roles. For a subset of them, we analyzed the effect of cooperative HBs, showing how they affect the interaction strength with respect to energies calculated *in vacuo*.

Next, uniform geometry sampling was used to build detailed directionality maps probing the volume around HB atoms. To have a consistent metric to quantify anisotropy, we defined a directionality index $D_{[probe]}$, which makes possible to quantify interactions with a probe to compare and rank different groups.

Results confirmed that all donors present rather similar directional propensity, with minor variations across the set (Supplementary Fig. S5) and in agreement with previous reports^{35,50,51}. Acceptors, on the other hand, were found to be anywhere from nearly isotropic to very directional (Fig. 3). In very isotropic acceptors, such as DMSO **94** in which $D_{[HF]}$ values are lower, most of the accessible surface can be engaged by a donor hydrogen without significant loss of interaction energy relatively to the optimal interaction geometry. Conversely, highly directional acceptors show larger $D_{[HF]}$ values in the 0.4 - 0.5 range, which are close to those of HB donors.

Whereas strength variations can be rationalized basing on properties such as electronegativity, hybridization and acidity, directionality showed to be more complex and difficult to reconcile than previously hypothesized. Clusters in Fig. 3a suggests that the element at an acceptor site is an important factor regulating the overall general strength and directionality characteristics of HB acceptors. Nitrogens are typically strong anisotropic acceptors, occupying the top right corner of the plot. Notably, among acceptors considered here, no other element is simultaneously as strong and anisotropic as nitrogen. Fluorine is simultaneously weak and isotropic (Fig. S16), occupying the opposite corner of the plot with respect to nitrogen. Sulfur is a weak acceptor but displays large anisotropy, occupying the top left corner of the plot. Oxygen displays the widest trend, ranging from weak to strong and from isotropic to directional, but it is neither simultaneously strong and directional like nitrogen, nor weak and isotropic like fluorine. In fact, in oxygen acceptors, strength and $D_{[HF]}$ are anti-correlated such that the stronger the acceptors, the least they manifest directional propensity.

Overall, these results challenge the established consensus according to which strength and directionality are correlated, with the latter depending on the former. Therefore, strong acceptors are expected to be more directional^{35,45,52–54}, and weak ones less so^{35,45,54,55}. When quantifying directionality of acceptors and donors separately, we found no such correlation with acceptor strength. For example, in the same molecular context, sulfur shows more directional character than oxygen, despite the former being a weaker acceptor (Fig. 2). When interacting with HF **30** probe, thioacetone **57** is 70%

Fig. 3 Directionality and energy errors of molecules in the dataset; points are colored by element type (black carbon; blue nitrogen; red oxygen; cyan fluorine; yellow sulfur) (a) $D_{[HF]}$ values over acceptor strength normalized with respect to water; higher $D_{[HF]}$ values correspond to more directional acceptors; strong and directional nitrogens tend to cluster on the *top-right* corner, weak and isotropic fluorines on the *bottom-left* corner, and weak and isotropic sulfur are on the *top-left* corner; oxygen atoms are distributed across strengths and directionality (b) quantitative directionality error $\epsilon_{[HF]}$ versus acceptor strength normalized with respect to water; lower values correspond to higher average energy loss when not establishing optimal HBs

more directional than acetone **67** ($D_{[HF]}$ 0.50 and 0.29, respectively; Table 1) and 20% weaker (ΔE_{opt} -6.95 and -8.81 kcal/mol, respectively; Fig. S9).

However, one might argue that high anisotropy of weak acceptors would not translate into effective directional propensity, given the small energies involved. Quantitative isotropy error ($\epsilon_{[HF]}$) can be used to estimate and compare these differences. In absence of external factors, the estimated energy loss at non-optimal geometries would be 0.9 kcal/mol higher for thioacetone than for acetone ($\epsilon_{[HF]}$ values -3.49 kcal/mol and -2.56 kcal/mol, respectively), revealing the more anisotropic character of the weaker acceptor, even when considering the strength of the interaction. The lack of correlation between $\epsilon_{[HF]}$ and acceptor strength is the prominent observation in Fig. 3b. Notably, acceptor nitrogens are the only element to reach $\epsilon_{[HF]}$ below -3.0 kcal/mol, a conse-

Table 2 Variations of directionality values calculated with MP2 for imidazole, acetone, and DMSO using acetylene **12** ($D_{[acetylene]}$), water **15** ($D_{[water]}$), HCN **24** ($D_{[HCN]}$), and HF **20** probes ($D_{[HF]}$)

	$D_{[acetylene]}$	$D_{[water]}$	$D_{[HCN]}$	$D_{[HF]}$
imidazole 92	0.32	0.35	0.34	0.43
acetone 67	0.18	0.23	0.19	0.27
DMSO 94	0.16	0.24	0.14	0.19

Table 3 Interaction energies ΔE_{opt} between imidazole, acetone and DMSO and four donors: acetylene **12**, water **15**, HCN **24** and HF **20**. The level of theory is MP2/aug-cc-pVTZ with counterpoise correction.

	$\Delta E_{opt}[acetylene]$	$\Delta E_{opt}[water]$	$\Delta E_{opt}[HCN]$	$\Delta E_{opt}[HF]$
imidazole 92	-4.3	-7.2	-8.1	-13.4
acetone 67	-3.4	-5.9	-6.1	-10.0
DMSO 94	-4.6	-8.4	-8.0	-13.0

quence of being simultaneously strong and anisotropic.

We then studied the influence of the probe strength on the measure of directionality, probing three acceptors (imidazole **92**, acetone **67**, DMSO **94**) using acetylene **12**, water **15**, HCN **24** and HF **20**, and calculated the resulting $D_{[probe]}$ values (Tables 2 and 3, Fig. 4).

Contrarily to what we found for cooperative HBs, where the presence of an extra HB with a water molecule affected the overall interaction strength but not directionality, the increase in the probe strength itself induces a systematic directionality increase. The results show that the relative directionality rank between two acceptors is conserved even when sampled with different probes, hinting at an intrinsic directionality property that indeed depends on the sole target characteristics, but gets also modulated by the probe. Notably, a directional acceptor such as imidazole **92** interacting with a very weak HB probe (acetylene **12**) maintains more directional character than the more isotropic DMSO **94** interacting with a very strong HB (HF **20**). Deviations from the trend were found when probing with a less linear molecule, such as water, due to interactions with flanking atoms (i.e., water hydrogens and acetone and DMSO methyls, Fig. 4a). Another interesting aspect is the effect of the HCN probe on the profiles, suggesting that the C-H pair in the nitrile induces similar HB anisotropy to the ethyne in acetylene despite being stronger (Fig. 4b and c), possibly due to the association of its strong dipole and medium HB strength.⁵⁶

The results require rethinking the way directionality is described, with implications for different aspects in which HBs are involved, and where directionality shapes the energy funnels around acceptors. From a thermodynamic perspective, isotropic acceptors are expected to be potentially more stable in water due to the larger number of microstates with energetically favorable HBs with the solvent, which shift the thermodynamic equilibrium towards the unbound state. Conversely, anisotropic groups (such as acceptor nitrogen) will interact less extensively and in a more ordered fashion with water, hence will pay smaller desolvation penalties upon binding. In fact, introducing a nitrogen acceptor site in a compound is often associated with an increase in potency⁵⁷. Drug design efforts toward improved binders should take into account that the directionality of acceptors can have an impact beyond the direct interaction between ligand and receptor, affecting also

Fig. 4 Directionality profiles and indices ($D_{[probe]}$) of imidazole **92**, acetone **67**, and DMSO **94** using acetylene **12**, water **15**, HCN **24**, and HF **20** probes; a) directionality indices versus water-relative probe strengths; b) *in plane* directionality profiles ($\theta = 0$, $\phi = [-90, 90]$); c) equal area projections of domes.

their individual interactions with the solvent. In this regard, not all HB groups might be counted equally when defining properties like drug-likeness⁵⁷. From a modeling perspective, the isotropy error $\varepsilon_{[probe]}$ is a useful metric for quantifying the estimate inaccuracy occurring when fully isotropic description (i.e., a spherical model) is used to describe a given acceptor-donor interaction. Generally, more detailed models tend to require more computational resources, therefore having a way to estimate the trade-off between accuracy and speed provides a guidance in the selection of appropriate approximations.⁵⁰

Far from being comprehensive, our dataset represents a compendium of the spectrum of HBs in the context of structural biology and supramolecular chemistry. We must note that in the condensed phase, polarization effects and networks of cooperative HBs modulate the strengths of HB capable molecules, therefore limiting the predictive power of our *in vacuo* calculations. Nevertheless, the chemical groups considered here provide an overview on how different donors and acceptors compare with each other, and with respect to water. This is particularly useful for chemical biology, protein engineering, drug design, and all cases in which solvation effects temper reversible interactions.

Finally, we believe that the geometric scheme described here, as well as the quantification of the directionality, have a more general interest. The possibility of performing such quantifications and comparisons in a consistent framework allowed us to identify a number of novel patterns. The method is suitable for generating high quality parameters for molecular mechanics force fields used in molecular dynamics and docking simulations, and can be easily adapted and extended to study other anisotropic interactions.

Methods

Dataset. Donors (**1-34**, Fig. S1) and acceptors (**35-101**, Fig. S2) were selected to provide a representative and diverse sample of classical^{9,58,59} and non-conventional^{55,59} HB groups, with emphasis on those common in biological and drug design contexts. With the exception of the cooperativ-

ity analysis, only single-site HBs have been considered to facilitate strength comparisons and directionality analysis, hence excluding multi-furcated interactions. In molecules capable of multiple HBs, each site was considered separately, e.g.: water is compound **15** as donor, and compound **75** as acceptor.

Dimer and trimer geometries and strength. All calculations were performed using Gaussian 09⁶⁰. Ideal acceptor-donor geometries were constructed and minimized with the B3LYP density functional⁶¹⁻⁶⁴ and the cc-pVDZ basis set⁶⁵ and counterpoise correction^{66,67}. Calculations were performed *in vacuo*, while water was considered part of the dataset and used to provide a reference value for the other molecules considered. Optimized structures were tested for absence of imaginary frequencies, and visually inspected to exclude interactions other than HBs, such as the formation of Lewis adducts or geometries that are clearly far from an ideal HB geometry. (i.e., interactions involving atoms other than the target/probe pairs, such as DMSO (**92**) methyls with HF (**30**) fluorine). Dimers and trimers that had optimized to transition states or to multi-furcated interactions were re-optimized from different starting geometries until issues were resolved. Electronic interaction energy ΔE of the minimized dimers calculated with eq. 1 was used to assess HB strength. The absolute energy of the isolated (unbound) monomers corresponds to their optimized geometries in gas-phase (as opposed to their optimized geometries in the dimer). The impact of the choice of the density functional is discussed in the *Supplementary Information*.

Directionality. The reference frame system defined the position of the probe (Fig. 1a) was defined as following: a) the reference center is set at the target atom, such that spherical coordinate r represents its distance with the probe; b) the z -axis is aligned along the mean of the bond vectors between neighbor atoms of the target and the target atom itself; c) the x axis is defined as the vector orthogonal to the z axis and oriented such that the bond to the bulkiest group in the molecule lies in positive x direction. Directional preferences of donors and acceptors were derived by sampling angular components (ϕ , θ) and distance (r), keeping the probe aligned to the r vector. Angular components were sampled at intervals of 10 degrees, from 0 to 90 for ϕ (10 intervals), and 0 to 360 for θ (37 intervals). Distance r was sampled at intervals of 0.1 Å from optimized dimer geometry up to 1 Å away (11 intervals), to accommodate for possible repulsion when leaving the HB well. Whenever possible, molecular symmetry was used to reduce the number of dome points to calculate. For most molecules, QM calculations were performed for θ values in the range 0–180, then equal area projections were generated by applying symmetry transformations. Exceptions were dimethyl sulfoxide **94** and trimethylphosphate **89**, for which QM calculations were performed covering the entire 0–360 range. During the geometry scanning, internal geometries (i.e.: bond lengths and angles) of target and probe were kept fixed, in order to preserve the intended (r , ϕ , θ) values. The valida-

tion of the level of theory to measure directionality, and the use of internal geometry restraints are discussed in *Supplementary Methods*.

Lambert azimuthal equal-area projection. The Lambert azimuthal equal-area projection was used to represent interaction energies as a function of the angular components. This projection has been used in prior work¹⁴. For each θ and ϕ pair, the lowest ΔE energy at the optimal r distance was represented in the projection. To generate the equal area projections, the position of probe atoms was converted from cartesian coordinates $\hat{r} = (x, y, z)$ into projection coordinates (X, Y) :

$$X = \frac{x}{|\hat{r}|} \sqrt{\frac{2}{1 + \frac{z}{|\hat{r}|}}} \quad (5)$$

$$Y = \frac{y}{|\hat{r}|} \sqrt{\frac{2}{1 + \frac{z}{|\hat{r}|}}} \quad (6)$$

ΔE is then represented as a function of projection coordinates (X, Y) . Values used for the graphical representation of equal-area projections were obtained with bilinear interpolation of ΔE values in (X, Y) coordinates. ΔE values interpolated from equal-area projection coordinates were used to calculate directionality indices D_i , to obtain a uniform sampling of the exposed target surface and compensate for variations in r due to different equilibrium distances, and the angular density variations in dome points.

ACKNOWLEDGEMENTS

The authors thank David S. Goodsell for insightful discussions, and Lisa Dong and J.C. Ducom at the TSRI HPC core for computational support. This work was supported by the National Institutes of Health (GM069832). This is manuscript number 29667 of the Scripps Research Institute. The authors declare that they have no competing financial interests.

Bibliography

- Bissantz, C., Kuhn, B. & Stahl, M. A medicinal chemist's guide to molecular interactions. *Journal of Medicinal Chemistry* **53**, 5061–5084 (2010).
- Snyder, P. W., Lockett, M. R., Moustakas, D. T. & Whitesides, G. M. Is it the shape of the cavity, or the shape of the water in the cavity? *The European Physical Journal Special Topics* **223**, 853–891 (2013).
- Smith, J. D. Energetics of hydrogen bond network rearrangements in liquid water. *Science* **306**, 851–853 (2004).
- Brini, E. *et al.* How water's properties are encoded in its molecular structure and energies. *Chemical Reviews* **117**, 12385–12414 (2017).
- Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences* **37**, 205–211 (1951).
- Shan, S.-o. & Herschlag, D. Energetic effects of multiple hydrogen bonds. implications for enzymatic catalysis. *Journal of the American Chemical Society* **118**, 5515–5518 (1996).
- Alegre-Requena, J. V., Marqués-López, E., Herrera, R. P. & Díaz, D. D. Metal-organic frameworks (mofs) bring new life to hydrogen-bonding organocatalysts in confined spaces. *CrystEngComm* **18**, 3985–3995 (2016).
- Steiner, T. & Desiraju, G. R. Distinction between the weak hydrogen bond and the van der Waals interaction. *Chemical Communications* 891–892 (1998).
- Arunan, E. *et al.* Definition of the hydrogen bond (IUPAC Recommendations 2011). *Pure and applied chemistry* **83**, 1637–1641 (2011).
- Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge structural database. *Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials* **72**, 171–179 (2016).
- Berman, H. M. The Protein Data Bank. *Nucleic Acids Research* **28**, 235–242 (2000).
- Desiraju, G. R. Iupac definition of the hydrogen bond. terminology and nomenclature. *Foundations of Crystallography* **70**, C308 (2017).
- Morozov, A. V., Kortemme, T., Tsemekhman, K. & Baker, D. Close agreement between the orientation dependence of hydrogen bonds observed in protein structures and quantum mechanical calculations. *Proceedings of the National Academy of Sciences* **101**, 6946–6951 (2004).
- O'Meara, M. J. *et al.* Combined covalent-electrostatic model of hydrogen bonding improves structure prediction with Rosetta. *Journal of chemical theory and computation* **11**, 609–622 (2015).
- Saha, B. K., Saha, A., Sharada, D. & Rather, S. A. F or O, which one is the better hydrogen bond (is it?) acceptor in C–H...X–C (X = F, O) interactions? *Crystal Growth & Design* (2017).
- Zhou, P., Tian, F., Lv, F. & Shang, Z. Geometric characteristics of hydrogen bonds involving sulfur atoms in proteins. *Proteins: Structure, Function, and Bioinformatics* **76**, 151–163 (2009).
- Brammer, L., Bruton, E. A. & Sherwood, P. Understanding the behavior of halogens as hydrogen bond acceptors. *Crystal Growth & Design* **1**, 277–290 (2001).
- Allen, F. H., Bird, C. M., Rowland, R. S. & Raithby, P. R. Resonance-induced hydrogen bonding at sulfur acceptors in R₁R₂C=S and R₁CS₂[–] systems. *Acta Crystallographica Section B Structural Science* **53**, 680–695 (1997).
- Allen, F., Bird, C., Rowland, R. & Raithby, P. Hydrogen-bond acceptor and donor properties of divalent sulfur (YSZ and RSH). *Acta Crystallographica Section B: Structural Science* **53**, 696–701 (1997).
- Wood, P. A., Pidcock, E. & Allen, F. H. Interaction geometries and energies of hydrogen bonds to C=O and C=S acceptors: a comparative study. *Acta Crystallographica Section B Structural Science* **64**, 491–496 (2008).
- Bruno, I. J. *et al.* Isostar: A library of information about nonbonded interactions. *Journal of Computer-Aided Molecular Design* **11**, 525–537 (1997).
- Steiner, T. The hydrogen bond in the solid state. *Angewandte Chemie International Edition* **41**, 48–76 (2002).
- Zhou, P., Zou, J., Tian, F. & Shang, Z. Fluorine bonding - how does it work in protein-ligand interactions? *Journal of Chemical Information and Modeling* **49**, 2344–2355 (2009).
- Voth, A. R., Khuu, P., Oishi, K. & Ho, P. S. Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nature Chemistry* **1**, 74–79 (2009).
- Nittinger, E. *et al.* Large-scale analysis of hydrogen bond interaction patterns in protein-ligand interfaces. *Journal of Medicinal Chemistry* **60**, 4245–4257 (2017).
- Allen, F. H., Harris, S. E. & Taylor, R. Comparison of conformer distributions in the crystalline state with conformational energies calculated by ab initio techniques. *Journal of Computer-Aided Molecular Design* **10**, 247–254 (1996).
- Hendrickson, W. A. Stereochemically restrained refinement of macromolecular structures. In *Methods in enzymology*, vol. 115, 252–270 (1985).
- Wendler, K., Thar, J., Zahn, S. & Kirchner, B. Estimating the hydrogen bond energy. *The Journal of Physical Chemistry A* **114**, 9529–9536 (2010).
- Kirchner, B. & Reiher, M. The secret of dimethyl sulfoxide- water mixtures. a quantum chemical study of 1dms- n water clusters. *Journal of the American Chemical Society* **124**, 6206–6215 (2002).
- Li, X.-Z., Walker, B. & Michaelides, A. Quantum nature of the hydrogen bond. *Proceedings of the National Academy of Sciences* **108**, 6369–6373 (2011).
- Lehmann, S., Spickermann, C. & Kirchner, B. Quantum cluster equilibrium theory applied in hydrogen bond number studies of water. 1. assessment of the quantum cluster equilibrium model for liquid water. *Journal of chemical theory and computation* **5**, 1640–1649 (2009).
- Green, A. J. & Popelier, P. L. Theoretical prediction of hydrogen-bond basicity p_k b_{hx} using quantum chemical topology descriptors. *Journal of chemical information and modeling* **54**, 553–561 (2014).
- Dannenberg, J., Haskamp, L. & Masunov, A. Are hydrogen bonds covalent or electrostatic? a molecular orbital comparison of molecules in electric fields and h-bonding environments. *The Journal of Physical Chemistry A* **103**, 7083–7086 (1999).
- Umeyama, H. & Morokuma, K. The origin of hydrogen bonding. an energy decomposition study. *Journal of the American Chemical Society* **99**, 1316–1332 (1977).
- Shahi, A. & Arunan, E. Why are hydrogen bonds directional? *Journal of Chemical Sciences* **128**, 1571–1577 (2016).
- Abraham, M. H., Grellier, P. L., Prior, D. V., Morris, J. J. & Taylor, P. J. Hydrogen bonding. Part 10. a scale of solute hydrogen-bond basicity using log *k* values for complexation in tetrachloromethane. *Journal of the Chemical Society, Perkin Transactions 2* 521 (1990).
- Laurence, C., Brameld, K. A., Graton, J., Le Questel, J.-Y. & Renault, E. The p_k_{bhx} database: Toward a better understanding of hydrogen-bond basicity for medicinal chemists. *Journal of Medicinal Chemistry* **52**, 4073–4086 (2009).
- Klamt, A., Reinisch, J., Eckert, F., Hellweg, A. & Diedenhofen, M. Polarization charge densities provide a predictive quantification of hydrogen bond energies. *Phys. Chem. Chem. Phys.* **14**, 955–963 (2012).
- Zheng, S. *et al.* Proposed hydrogen-bonding index of donor or acceptor reflecting its intrinsic contribution to hydrogen-bonding strength. *Journal of Chemical Information and Modeling* (2017).
- Kutzelnigg, W. Chemical bonding in higher main group elements. *Angewandte Chemie International Edition* **23**, 272–295 (1984).
- Śmiechowski, M., Gojlo, E. & Stangret, J. Systematic study of hydration patterns of phosphoric (V) acid and its mono-, di-, and tripotassium salts in aqueous solution. *The Journal of Physical Chemistry B* **113**, 7650–7661 (2009).
- Ruben, E. A., Chapman, M. S. & Evansek, J. D. Hydrogen bonding mediated by key orbital interactions determines hydration enthalpy differences of phosphate water clusters. *The Journal of Physical Chemistry A* **111**, 10804–10814 (2007).
- Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *The Journal of Physical Chemistry* **97**, 10269–10280 (1993).
- Nochebuena, J., Cuatli, C. & Ireta, J. Origin of cooperativity in hydrogen bonding. *Physical Chemistry Chemical Physics* **19**, 15256–15263 (2017).
- Pimentel, G. C. & McClellan, A. L. *The hydrogen bond* (Freeman, 1960).
- Steiner, T. & Desiraju, G. R. Distinction between the weak hydrogen bond and the van der Waals interaction. *Chemical Communications* 891–892 (1998).
- Taylor, R., Kennard, O. & Versichel, W. Geometry of the imino-carbonyl (N–H...O=C) hydrogen bond. 1. lone-pair directionality. *Journal of the American Chemical Society* **105**, 5761–5766 (1983).
- Desiraju, G. R. C–H...O and other weak hydrogen bonds. from crystal engineering to virtual screening. *Chemical Communications* 2995–3001 (2005).
- Wood, P. A., Allen, F. H. & Pidcock, E. Hydrogen-bond directionality at the donor H atom: analysis of interaction energies and database statistics. *CrystEngComm* **11**, 1563–1571 (2009).
- Lu, Z., Zhou, N., Wu, Q. & Zhang, Y. Directional dependence of hydrogen bonds: A density-based energy decomposition analysis and its implications on force field development. *Journal of chemical theory and computation* **7**, 4038–4049 (2011).
- Ireta, J., Neugebauer, J. & Scheffler, M. On the accuracy of dft for describing hydrogen bonds: Dependence on the bond directionality. *The Journal of Physical Chemistry A* **108**, 5692–5698 (2004).
- Kollman, P. A. Theory of hydrogen bond directionality. *Journal of the American Chemical Society* **94**, 1837–1842 (1972).
- Jeffrey, G. A. & Saenger, W. *Hydrogen bonding in biological structures* (Springer Science &

- Business Media, 1991).
54. Jeffrey, G. A. & Jeffrey, G. A. *An introduction to hydrogen bonding*, vol. 32 (Oxford university press New York, 1997).
 55. Desiraju, G. R. & Steiner, T. *The weak hydrogen bond: in structural chemistry and biology*, vol. 9 (International Union of Crystallography, Monographs on Crystallography, 2001).
 56. Jucks, K. & Miller, R. Infrared spectroscopy of the hydrogen cyanide dimer. *The Journal of chemical physics* **88**, 6059–6067 (1988).
 57. "Pennington, L. D. & Moustakas, D. T. . The Necessary Nitrogen Atom: A Versatile High-Impact Design Element for Multiparameter Optimization. *J. Med. Chem.* **60**, 3552–3579 (2017).
 58. Arunan, E. *et al.* Defining the hydrogen bond: An account (iupac technical report). *Pure and Applied Chemistry* **83**, 1619–1636 (2011).
 59. Alkorta, I. & Elguero, J. Non-conventional hydrogen bonds. *Chemical Society Reviews* **27**, 163–170 (1998).
 60. Frisch, M. J. *et al.* Gaussian 09 Revision A.02.
 61. Becke, A. D. Density-functional thermochemistry. III. the role of exact exchange. *The Journal of Chemical Physics* **98**, 5648–5652 (1993).
 62. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **37**, 785–789 (1988).
 63. Vosko, S. H., Wilk, L. & Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics* **58**, 1200–1211 (1980).
 64. Stephens, P. J., Devlin, F. J., Chabalowski, C. F. & Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry* **98**, 11623–11627 (1994).
 65. Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *The Journal of Chemical Physics* **90**, 1007–1023 (1989).
 66. Boys, S. & Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics* **19**, 553–566 (1970).
 67. Simon, S., Duran, M. & Dannenberg, J. J. How does basis set superposition error change the potential surfaces for hydrogen-bonded dimers? *The Journal of Chemical Physics* **105**, 11024–11031 (1996).
 68. Morozov, A. V. & Kortemme, T. Potential functions for hydrogen bonds in protein structure prediction and design. *Advances in Protein Chemistry* 1–38 (2005).
 69. Rezáč, J., Riley, K. E. & Hobza, P. Extensions of the S66 data set: More accurate interaction energies and angular-displaced nonequilibrium geometries. *Journal of Chemical Theory and Computation* **7**, 3466–3470 (2011).
 70. Llamas-Saiz, A. L., Foces-Foces, C., Mo, O., Yanez, M. & Elguero, J. Nature of the hydrogen bond: crystallographic versus theoretical description of the OH...N (sp^2) hydrogen bond. *Acta Crystallographica Section B: Structural Science* **48**, 700–713 (1992).
 71. Platts, J. A., Howard, S. T. & Bracke, B. R. F. Directionality of hydrogen bonds to sulfur and oxygen. *Journal of the American Chemical Society* **118**, 2726–2733 (1996).
 72. Allen, F. H., Baalham, C. A., Lommerse, J. P. M., Raithby, P. R. & Sparr, E. Hydrogen-bond acceptor properties of nitro-o atoms: A combined crystallographic database and ab initio molecular orbital study. *Acta Crystallographica Section B: Structural Science* **53**, 1017–1024 (1997).
 73. Nanda, V. & Schmiedeknecht, A. Are aromatic carbon donor hydrogen bonds linear in proteins? *Proteins: Structure, Function, and Bioinformatics* **70**, 489–497 (2007).
 74. Wood, P. A., Allen, F. H. & Pidcock, E. Hydrogen-bond directionality at the donor h atom: analysis of interaction energies and database statistics. *CrystEngComm* **11**, 1563 (2009).
 75. Choi, H., Kang, H. & Park, H. New angle-dependent potential energy function for backbone-backbone hydrogen bond in protein-protein interactions. *Journal of Computational Chemistry* **30**, 1000–1009 (2009).
 76. Choi, H., Kang, H. & Park, H. Extended morse function model for angle-dependent hydrogen bond in protein-protein interactions. *The Journal of Physical Chemistry B* **114**, 2980–2987 (2010).
 77. Lu, Z., Zhou, N., Wu, Q. & Zhang, Y. Directional dependence of hydrogen bonds: A density-based energy decomposition analysis and its implications on force field development. *Journal of Chemical Theory and Computation* **7**, 4038–4049 (2011).
 78. Tatipolsky, M. & Ansorg, K. Toward a physically motivated force field: Hydrogen bond directionality from a symmetry-adapted perturbation theory perspective. *Journal of Chemical Theory and Computation* **12**, 1267–1279 (2016).
 79. Mondal, A. & Datta, S. Quantum mechanical electronic structure calculation reveals orientation dependence of hydrogen bond energy in proteins. *Proteins: Structure, Function, and Bioinformatics* **85**, 1046–1055 (2017).
 80. Rao, L., Ke, H., Fu, G., Xu, X. & Yan, Y. Performance of several density functional theory methods on describing hydrogen-bond interactions. *Journal of Chemical Theory and Computation* **5**, 86–96 (2009).
 81. Koné, M., Illien, B., Graton, J. & Laurence, C. B3LYP and MP2 calculations of the enthalpies of hydrogen-bonded complexes of methanol with neutral bases and anions: Comparison with experimental data. *The Journal of Physical Chemistry A* **109**, 11907–11913 (2005).
 82. El Kerdawy, A., Tautermann, C. S., Clark, T. & Fox, T. Economical and accurate protocol for calculating hydrogen-bond-acceptor strengths. *Journal of Chemical Information and Modeling* **53**, 3262–3272 (2013).
 83. Hao, M.-H. Theoretical calculation of hydrogen-bonding strength for drug molecules. *Journal of Chemical Theory and Computation* **2**, 863–872 (2006).
 84. Graton, J., Le Questel, J.-Y., Maxwell, P. & Popelier, P. Hydrogen-bond accepting properties of new heteroaromatic ring chemical motifs: A theoretical study. *Journal of chemical information and modeling* **56**, 322–334 (2016).
 85. Kim, K. & Jordan, K. Comparison of density functional and mp2 calculations on the water monomer and dimer. *The Journal of Physical Chemistry* **98**, 10089–10094 (1994).
 86. Murashov, V. & Leszczynski, J. A comparison of the b3lyp and mp2 methods in the calculation of phosphate complexes. *Journal of Molecular Structure: THEOCHEM* **529**, 1–14 (2000).
 87. Goerigk, L. & Grimme, S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Physical Chemistry Chemical Physics* **13**, 6670–6688 (2011).
 88. Rezac, J., Huang, Y., Hobza, P. & Beran, G. J. Benchmark calculations of three-body intermolecular interactions and the performance of low-cost electronic structure methods. *Journal of chemical theory and computation* **11**, 3065–3079 (2015).
 89. Medders, G. R., Babin, V. & Paesani, F. A critical assessment of two-body and three-body interactions in water. *Journal of chemical theory and computation* **9**, 1103–1114 (2013).

Supplementary Information

Previous QM works. Supplementary Table S1 summarizes previous QM studies published to date, describing HB dimer interactions at non-optimal geometries, with the aim of characterizing variation in ΔE with angular components. Most of the previous studies considered only one angular component. Others addressed multiple angular components separately (Morozov et al.^{13,68} and Řezáč et al.⁶⁹). We could find only a single study in which two angular components were addressed simultaneously for the same target site (Lamas et al.⁷⁰).

Supplementary Table S1 Previous QM studies of HB directionality which considered interaction energies at non-optimal geometries (ranked by publication year)

Ref	Year	Acceptors	Donors	Level of theory	Geometric parameters ^a
Lamas et al. ⁷⁰	1992	N(sp^2)	-	HF/3-21G	$f(\theta, \phi)$
Platts et al. ⁷¹	1996	O(sp^2, sp^3) S(sp^2, sp^3)	-	multipole electrostatics ^b	$f(\phi)$
Allen et al. ⁷²	1997	NO ₂	-	IMPT/6-31G**	$f(\phi)$
Morozov et al. ^{13,68}	2004	formamide	formamide	PBE96/aug-cc-pVDZ	$f(\theta), f(\phi)$
Nanda et al. ⁷³	2007	-	CH(aromatic)	B3LYP/6-11++G**	$f(\phi)$
Wood et al. ²⁰	2008	O(sp^2) S(sp^2)	-	IMPT/6-31G**	$f(\phi)$
Wood et al. ⁷⁴	2009	-	OH(sp^3) CH(sp, sp^2)	IMPT/6-31G**	$f(\phi)$
Choi et al. ⁷⁵	2009	N-methylacetamide	N-methylacetamide	B3LYP/6-311G**	$f(\phi)$
Choi et al. ⁷⁶	2010	O(sp^2)	NH(sp^2) OH(sp^3)	B3LYP/6-311G**	$f(\phi)$
Řezáč et al. ⁶⁹	2011	10	10	CCSD(t)/CBS	$f(\phi), f(\theta)$
Lu et al. ⁷⁷	2011	O(sp^2, sp^3) N(sp, sp^2, sp^3) S(sp^2, sp^3)	-	B3LYP-D3/aug-cc-pVDZ	$f(\phi)$
Tafipolsky et al. ⁷⁸	2016	water	-	DFT-SAPT/aug-cc-pVQZ	$f(\phi)$
Mondal et al. ⁷⁹	2017	O(sp^2)	-	M06-2X/6-311+G*	$f(\phi)$

^a θ and ϕ values as depicted in Fig. 1

^b multipoles were derived from MP2 level of theory.

Dataset molecules. The molecules we selected for our dataset are depicted in Supplementary Figs. S1 (acceptors) and S2 (donors). The dataset includes conventional HB groups commonly found in drugs and biological molecules (e.g., hydroxyl, carbonyl, amines, etc.) as well as some inorganic ones (hydrogen fluoride and chloride, fluoroborane, etc.), spanning a wide range of strengths. Only single-site HBs have been considered to facilitate strength comparisons and directionality analysis, thus excluding multi-furcated interactions. Molecules capable of multiple HBs were considered as separate entries in the dataset, e.g.: water is compound **15** as donor, and compound **75** as acceptor, and each HB site was considered separately.

Level of theory validation. Given the large number of independent computations to be performed, we employed the B3LYP density functional with the cc-pVDZ basis set. B3LYP is known to systematically underestimate absolute interaction energies⁸⁰, but numerous previously reported works show that it provides accurate ranking of relative HB strengths^{81–86}, which is adequate for the scope of this work. As a validation, we performed multiple tests using MP2/aug-cc-pVTZ on a small sample of compounds (using the counterpoise correction), and compared the results obtained with the two levels of theory to assess the error in both HB strengths and anisotropy profiles.

Accuracy of HB strengths. We compared the values of ΔE_{opt} of dimer geometries fully minimized with B3LYP/cc-pVDZ and MP2/aug-cc-pVTZ levels of theory. A small subset of 33 dimers was used for this comparison: dimethylamine **9** – HCN **61**, HCN **24** – dimethylamine **83**, HCN **24** – dimethyl ether **63**, HCN **24** – dimethyl sulfide **56**, HCN **24** – furan **45**, HCN **24** – HCN **61**, HCN **24** – methylamine **91**, HCN **24** – methanol **71**, HCN **24** – methanethiol **53**, HCN **24** – water **75**, methanol **16** – HCN **61**, methanol **16** – methylamine **91**, methanol **16** – methanol **71**, methanol **16** – methanethiol **53**, methanol **16** – water **75**, pyrrole **21** – HCN **61**, pyrrole **21** – water **75**, methanethiol **3** – HCN **61**, methanethiol **3** – methylamine **91**, methanethiol **3** – methanethiol **53**, methanethiol **3** – water **75**, water **15** – dimethylamine **83**, water **15** – dimethyl ether **63**, water **15** – dimethyl sulfide **56**, water **15** – furan **45**, water **15** – HCN **61**, water **15** – methylamine **91**, water **15** – methanol **71**, water **15** – pyridine **82**, water **15** – methanethiol **53**, water **15** – trimethylamine **78** and water **15** – water **75**. Optimizations were performed with the counterpoise correction. The energy difference between B3LYP/cc-pVDZ and MP2/aug-cc-pVTZ is systematic and within about 1.0 kcal/mol, (Fig. S3), in agreement with previous studies.⁸⁰ The overall trend was preserved (Pearson correlation coefficient: 0.88) indicating a satisfactory agreement between the two methods.

As a validation, we tested that the relationship of $D_{[HF]}$ with strength for thioketone **57** ($\Delta E_{opt} = -7.36$) and acetone **67** ($\Delta E_{opt} = -10.11$) was consistent between results obtained with B3LYP and MP2. The values of $\varepsilon_{[HF]}$ calculated with MP2 (-3.7 and -2.9 kcal/mol) showed to be very close to those calculated with B3LYP (-3.49 kcal/mol and -2.56 kcal/mol).

Accuracy of anisotropy profiles. To assess the quality of directionality profiles calculated with B3LYP/cc-pVDZ we used MP2/aug-cc-pVTZ to compute domes for donor and acceptor sites of water (**15**, **75**), imidazole (**25**, **92**), methyl-acetamide (**18**, **90**), methyl-thioacetamide (not numbered as donor; compound **62** as acceptor), phenol (**26**, **55**), and TMAO (**97**). Results are summarized in Supplementary Table S2. For acceptors, the equal area projections obtained with the two levels of theory are nearly indistinguishable. The largest numerical deviation of $D_{[HF]}$ is observed for phenol (MP2: 0.37; B3LYP: 0.42). This is

Supplementary Figure S1 Donor groups ranked by their interaction energy (in kcal/mol, below each molecule) with the HCN acceptor probe 61. Donor target atoms are colored in blue

Supplementary Figure S2 Acceptor groups ranked by their interaction energy (in kcal/mol, below each molecule) with the HCN donor probe 24. Acceptor target atoms are colored in red

Supplementary Figure S3 Interaction energies of 33 dimers at the MP2/aug-cc-pVTZ and B3LYP/cc-pVDZ levels of theory. Each dimer was fully optimized with both levels of theory using the counterpoise correction. The red dashed line represents a regression model for this set of dimers.

likely due to better characterization of the delocalized electrons in the aromatic compound with MP2, while, visual inspection of the equal area projections revealed practically identical qualitative profiles (Fig. S4).

With respect to donor directionality, MP2/aug-cc-pVTZ results showed systematic profile broadening for imidazole, N-methylacetamide and N-methylthioacetamide, and to a lesser extent for water and phenol (Fig. S5). This was attributed to underestimation of the magnitude of dispersion interactions by B3LYP⁸⁷. Nevertheless, visual inspection of the domes shows very similar anisotropy maps. In summary, the confidence on $D_{[probe]}$ values calculated with B3LYP is higher for acceptor sites than for donors. This is probably due to the use of hydrogen fluoride (HF) to probe acceptors. HF has only 10 electrons and is highly polarized, minimizing the relative contribution of dispersion to the overall interaction energy. Overall, there is excellent agreement between B3LYP/cc-pVDZ and MP2/aug-cc-pVTZ for acceptor domes, as long as the probe is HF.

The B3LYP density function is adequate for the study of anisotropy in cooperative HBs (Supplementary Fig. S18), because three-body interaction energies are well described, even without dispersion corrections^{88,89}.

Supplementary Table S2 Directionality indexes calculated with MP2/aug-cc-pVTZ and B3LYP/cc-pVDZ. HF probes acceptor sites. HCN probes donor sites

target type	probe	target	MP2/aug-cc-pVTZ		B3LYP/cc-pVDZ	
			$\Delta E_{opt(probe)}$	$D_{[probe]}$	$\Delta E_{opt(probe)}$	$D_{[probe]}$
acceptor	HF	water 75	-8.0	0.34	-8.6	0.34
acceptor	HF	phenol 55	-7.3	0.37	-6.2	0.42
acceptor	HF	imidazole 92	-13.4	0.43	-11.6	0.45
acceptor	HF	N-methyl-acetamide 90	-11.4	0.21	-10.6	0.21
acceptor	HF	N-methyl-thioacetamide 62	-9.4	0.38	-8.0	0.36
acceptor	HF	TMAO 97	-20.0	0.25	-18.3	0.25
donor	HCN	water 15	-4.2	0.48	-3.2	0.47
donor	HCN	phenol 26	-6.1	0.47	-4.8	0.53
donor	HCN	imidazole 25	-5.9	0.32	-4.8	0.43
donor	HCN	N-methyl-acetamide 18	-5.0	0.35	-3.6	0.46
donor	HCN	N-methyl-thioacetamide	-5.8	0.33	-4.4	0.43

Internal geometry constraints. When calculating domes, we generated input geometries in an automated fashion and performed a single point energy evaluations for each geometry. During these calculations structures were not allowed to relax, specifically bond lengths and bond angles, in order to preserve the intended (r , ϕ , θ) values. This could create artificially high energies when angles are far from the optimum, and repulsions occur. To assess the effect of these restraints on calculated energies, we performed a relaxed energy scan along the ϕ angle of the HF-pyridine dimer (Fig. S6, black circles), and compared it with the energy profile obtained using rigid monomers at fixed distance (Fig. S6, red dashed line). The distance between the red dashed line and the black circles is the energy error in the domes if fixed r were used. When the distance between the donor

Supplementary Figure S4 Directionality profiles of acceptor sites calculated with B3LYP/cc-pVDZ (a) and MP2/aug-cc-pVTZ (b and c).

Supplementary Figure S5 Directionality profiles of donor sites calculated with B3LYP/cc-pVDZ (a) and MP2/aug-cc-pVTZ (b and c).

Supplementary Figure S6 Directionality profiles obtained using the rigid monomer approach compared to a relaxed energy scan (in which bond lengths and angles are allowed to relax).

Supplementary Figure S7 Comparison of donor strengths ΔE_{opt} ranked by their strength relative to water, and colored by element type. Values are calculated with HCN acceptor probe 61.

hydrogen and the acceptor atom is allowed to relax (Fig. S6, black solid line), negligible differences are found. Therefore the r distance scan was included in the calculation of domes.

Supplementary Figure S8 Comparison of acceptor strengths ΔE_{opt} ranked by their strength relative to water, and colored by element type. Values are calculated with HCN donor probe 24.

Supplementary Figure S9 Interaction energies between donor (*columns*) and acceptor (*rows*) groups, ranked by their interaction energy with the reference probe HCN (**61** and **24**, respectively). Annotations: *i* imaginary frequencies (optimized to transition state); *d* multi-furcated HB; *r* inversion of donor/acceptor roles; *a* dissociation of acidic proton; *n/hb* not a hydrogen bond.

Supplementary Table S3 Directionality and strength in cooperative HBs: acceptor sites.

Target	<i>In vacuo</i>		Cooperative	
	$\Delta E_{\text{opt}}[\text{HF}]$	$D_{\text{[HF]}}$	$\Delta E_{\text{wat}}[\text{HF}]$	$D_{\text{[HF]}}$
water 75	-8.7	0.34	-10.9	0.31
N-methylacetamide 90	-10.4	0.21	-10.9	0.21
imidazole 92	-11.6	0.45	-13.4	0.42

Supplementary Table S4 Directionality and strength in cooperative HBs: donor sites.

Target	<i>In vacuo</i>		Cooperative	
	$\Delta E_{\text{opt}}[\text{HCN}]$	$D_{\text{[HCN]}}$	$\Delta E_{\text{wat}}[\text{HCN}]$	$D_{\text{[HCN]}}$
water 15	-3.2	0.47	-3.9	0.48
N-methylacetamide 18	-3.6	0.46	-3.7	0.47
imidazole 25	-4.8	0.43	-5.2	0.43

Supplementary Figure S10 Directionality of donors. $D_{[HCN]}$ values range from 0.45 to 0.54, except for HCN 24 ($D_{[HCN]} = 0.36$).

Supplementary Figure S11 Directionality of amine series: methylamine 93, dimethylamine 83 and trimethylamine 78.

Supplementary Figure S12 Directionality of sp^2 sulfur: thioacetone **57** and N-methyl thioacetamide **62**.

Supplementary Figure S13 Directionality of oxygen bound to sulfur or phosphate: dimethyl sulfate **58**, methylsulfonamide **65**, dimethyl sulfone **77**, and trimethyl phosphate **89**.

Supplementary Figure S14 Directionality of various sp^2 oxygens: acetamide **85**, N,N-dimethyl acetamide **88**, benzamide **80**, methyl acetate **72**, acetic acid **74**, and nitrobenzene **59**.

Supplementary Figure S15 Directionality of various sp^3 oxygen and sulfur groups: diphenyl ether **51**, phenol **55**, dimethyl ether **63**, methanol **71**, methyl acetate **46**, dimethyl sulfide **56**, and methanethiol **53**.

Supplementary Figure S16 Directionality of nitrogen (pyridine **82**, *N*-methylacetimine **86**, guanidine **95**, HCN **61**) and fluorine (fluoromethyl **50**, HF **52**, sulfonyl fluoride **40**) acceptors.

Supplementary Figure S17 Directionality of anions methanethiolate **100** and acetate **101**.

Supplementary Figure S18 Directionality profiles with (b) and without (a) cooperative effects. An explicit water molecule was added to create a HB network involving the added water molecule, the target molecule under study, and the probe.