**Electronic Supplementary Information for**

***When are two hydrogen bonds better than one? Accurate first-principles models explain the balance of hydrogen bond donors and acceptors found in proteins***

Vyshnavi Vennelakanti[1,2], Helena W. Qi[1,2], Rimsha Mehmood[1,2], and Heather J. Kulik[1,*]
[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139
[2]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

**Contents**

**Table S1.** Step-by-step refinement of protein crystal structures from the data set. N, Q, S, T, and Y are the one-letter amino acid codes for Asparagine, Glutamine, Serine, Threonine, and Tyrosine, respectively. S-N, S-Q, T-N, T-Q, Y-N, and Y-Q are the residue pairs in the data set.

| Refinement criteria | S-N | S-Q | T-N | T-Q | Y-N | Y-Q | Total |
|---|---|---|---|---|---|---|---|
| resolution < 1.5 Å | 1,284 | 854 | 1,504 | 1,011 | 843 | 618 | 6,114 |
| HB distance cutoff applied | 889 | 545 | 1,016 | 591 | 578 | 416 | 4,035 |
| HB angle cutoff applied | 856 | 526 | 988 | 582 | 562 | 394 | 3,908 |

**Text S1.** Topology analysis of PDB files from protein data set.

To identify the HB distance criteria, residue pairs were selected at random from the initial protein data set. Hydrogen atoms were added to the PDB files using Avogadro v1.2.0[1] and the added hydrogen atoms were force-field optimized with MMFF94[2] while the heavy atoms were held fixed. Hydrogen atom positions were further refined with constrained geometry optimizations at the B3LYP-D3/6-31G* level of theory. The hybrid DFT optimization was selected for consistency with the level of theory with which the topology analysis could be readily carried out. Topology analysis was carried out to identify bond critical points (BCPs) and evaluate HB energies where BCPs were observed[3-5]. Negative energies were not used alone to determine the presence of HBs from the potential energy densities using Espinosa's equation (i.e., $E_{HB} = V_{BCP}/2$) because they can be significant overestimates[3]. They were used only for qualitative assessments.

Based on these considerations, we determined HBs to exist only for those structures for which HB energies computed from potential energy densities were equal to or stronger than -4 kcal/mol. N–H···O HBs satisfying this criterion were found between 2.50 and 3.20 Å in Ser/Thr/Tyr-Asn/Gln, and O–H···O HBs were found between 2.40 and 3.10 Å in Ser/Thr-Asn/Gln systems or between 2.40 and 3.20 Å in Tyr-Asn/Gln systems. To confirm that strong HBs were indeed absent in structures with HB distances outside of the above set HB distance criteria, we selected at random 10% of all the PDB structures that fell outside of the HB distance criteria as a representative sample. This corresponds to PDB structures with N···O distance < 2.50 Å or > 3.20 Å for Ser/Thr/Tyr-Asn/Gln systems and PDB structures with O···O distance < 2.40 Å or > 3.10 Å (3.20 Å) for Ser/Thr-Asn/Gln (Tyr-Asn/Gln) systems. We then carried out topology analysis for all the selected systems and computed HB energies from potential energy densities of the BCPs (when present). We found the resulting HB energies to be consistent with the set HB distance criteria, i.e., HBs of significant strength were absent from this set.

Following a similar protocol, HB angle criteria for N–H···O HBs and O–H···O HBs were determined. We performed topology analysis for all the structures with HB angles ≤ 110°. We found that four N–H···O HB systems showed strong HBs with HB angles < 110° (105.0°, 107.3°, 109.0°, and 109.9°) and very short HB distances (2.70 Å, 3.20 Å, 2.69 Å, and 2.68 Å, respectively). We did not observe strong O–H···O HBs with HB angles ≤ 110°. The complete set of results of these calculations is tabulated in an Excel file in Supporting Information.

**Table S2.** QTAIM analysis for some of the PDB files for the N–H···O and O–H···O HBs present between Ser/Thr/Tyr and Asn/Gln on B3LYP-D3/6-31G* geometry optimized PDB structures with heavy atoms constrained. The N–H···O and O–H···O HB distances for these PDB structures fall outside the HB distance criteria, and hence, HBs are either not observed or the observed HBs are very weak, as corroborated by the HB energy estimates. The HB distance and angle information are given in Å and °, respectively. The energy of HB ($E_{HB}$), which is estimated as half of the potential energy at the BCP obtained from QTAIM topology analysis, is given in kcal/mol.

| PDB ID | Chain Name | Residue 1 | Residue 2 | HB distance (Å) | HB angle (°) | $E_{HB}$ (kcal/mol) |
|---|---|---|---|---|---|---|
| N–H···O HBs | | | | | | |
| 3ir4 | A | S169 | N165 | 3.34 | 138.7 | -2.0 |
| 5vn4 | A | N213 | S210 | 3.26 | 104.6 | --- |
| 4esp | A | S37 | Q4 | 3.67 | 73.7 | -0.5 |
| 1llf | A | Q62 | S59 | 3.64 | 91.0 | -1.2 |
| 5akr | A | N305 | T252 | 3.33 | 101.0 | -1.2 |
| 5ta0 | A | T505 | N348 | 3.38 | 111.7 | -1.0 |
| 3e2d | B | Q364 | T283 | 3.62 | 104.7 | -0.5 |
| 2fgq | X | T226 | Q224 | 3.64 | 108.2 | -0.5 |
| 1lq9 | A | N62 | Y51 | 3.60 | 94.2 | -0.6 |
| 3i45 | A | Y24 | N12 | 3.47 | 134.1 | -1.2 |
| 5js4 | B | Q267 | Y265 | 3.30 | 108.9 | --- |
| 4brc | B | Q142 | Y138 | 3.62 | 66.0 | -0.6 |
| O–H···O HBs | | | | | | |
| 5f8s | B | N115 | S113 | 3.48 | 79.7 | --- |
| 3gzb | E | S48 | N45 | 3.56 | 157.8 | -1.1 |
| 3b7e | B | Q226 | S181 | 3.15 | 159.3 | -3.4 |
| 2x9g | D | Q236 | S233 | 3.51 | 155.4 | -1.5 |
| 2imi | B | T54 | N50 | 3.10 | 150.1 | -3.4 |
| 2r01 | A | T37 | N34 | 3.53 | 161.9 | -1.3 |
| 5ihv | A | Q98 | T95 | 3.47 | 142.7 | --- |
| 5a71 | A | Q46 | T43 | 3.40 | 162.5 | -2.0 |
| 3a72 | A | Y241 | N180 | 3.32 | 95.6 | -0.8 |
| 1u3w | A | N114 | Y110 | 3.36 | 155.2 | -2.0 |
| 5epu | E | Y52 | Q31 | 3.56 | 98.0 | --- |
| 4gwb | A | Q101 | Y63 | 3.60 | 113.6 | --- |

**Text S2.** Addition of hydrogen atoms to PDB files.

PDB structures in the protein data set do not contain hydrogen atoms. So, an in-house Python script was used to add hydrogen atoms to select atoms in the PDB files.

*N–H···O HBs.* Since the atoms involved in an N–H···O HB in Ser/Thr/Tyr-Asn/Gln systems are the side-chain hydroxyl oxygen of Ser/Thr/Tyr and side-chain amide nitrogen and hydrogen atoms of Asn/Gln, hydrogen atoms were added only to the side-chain amide nitrogen of Asn/Gln using an in-house Python script. The script reads each PDB file and writes the coordinates of side-chain hydroxyl oxygen of Ser/Thr/Tyr, side-chain amide carbon, oxygen, and nitrogen atoms of Asn/Gln to an XYZ file. Based on the following assumptions, two hydrogen atoms were added to the nitrogen atom in the XYZ file.

(i) Carbon, oxygen, nitrogen, and hydrogen atoms of the amide side-chain of Asn/Gln are in the same plane owing to the partial double-bond character of the C–N bond.

(ii) The ∠C–N–H angle is 120.6° in Ser/Thr-Asn/Gln systems and 120.4° in Tyr-Asn/Gln systems. These angles correspond to the mean value of ∠C–N–H angle of *syn* and *anti* N–H···O HBs observed in MP2/6-31G* optimized geometries of acetamide–methanol and acetamide–*p*-cresol N–H···O HBs, respectively.

(iii) The N–H distance is 1.02 Å in all the systems as observed in the MP2/6-31G* optimized geometries of acetamide–methanol and acetamide–*p*-cresol N–H···O HBs.

Each of the above assumptions leads to one mathematical equation, as elaborated below in the order of assumptions.

(i) We obtained the equation of plane formed by the side-chain amide carbon, oxygen, and nitrogen atoms using their coordinates. Because hydrogen atoms lie in the same plane, the x, y, and z coordinates of each hydrogen atom must satisfy the equation of plane formed by side-chain amide carbon, oxygen, and nitrogen atoms.

$$ah_x + bh_y + ch_z + d = 0$$

where a, b, c, and d are known and $h_x$, $h_y$, and $h_z$ are the x, y, and z coordinates of a hydrogen atom.

(ii) x, y, and z coordinates of the hydrogen atom must satisfy the dot-product equation of C–N and N–H bonds.

$$\cos(\angle C\text{–}N\text{–}H) = (\overrightarrow{C-N} \cdot \overrightarrow{N-H})/(||\overrightarrow{C-N}||\ ||\overrightarrow{N-H}||)$$

where $\overrightarrow{N-H} = (n_x - h_x, n_y - h_y, n_z - h_z)$.

(iii) x, y, and z coordinates of hydrogen atom must satisfy the equation of norm of $\overrightarrow{N-H}$.

$$||\overrightarrow{N-H}|| = \sqrt{(n_x - h_x)^2 + (n_y - h_y)^2 + (n_z - h_z)^2} = 1.02 \text{ Å}$$

The Python script solves the above three equations for the three unknowns. The presence of square terms in the equation of norm leads to two sets of coordinates, which correspond to the two hydrogen atoms. The script then writes the coordinates of hydrogen atoms to the XYZ file. The distances between side-chain hydroxyl oxygen of Ser/Thr/Tyr and side-chain amide hydrogen atoms of Asn/Gln were measured. The hydrogen atom closest to the side-chain hydroxyl oxygen would result in favorable hydrogen bonding interaction. The position of this hydrogen atom with respect to the amide oxygen determines whether the N–H···O HB is *syn* or *anti*. The script classifies the HBs as *syn* or *anti*, and also computes the N–H···O HB angle.

*O–H···O HBs.* Addition of a hydrogen atom to side-chain hydroxyl oxygen of Ser/Thr/Tyr follows a protocol very similar to the one followed while adding hydrogen atoms to N–H···O HB systems. The hydrogen atom is added using an in-house Python script only to the side-chain hydroxyl oxygen since the atoms involved in O–H···O HB are side-chain amide oxygen of Asn/Gln and side-chain hydroxyl group of Ser/Thr/Tyr.

The script reads each PDB file and writes the coordinates of side-chain hydroxyl oxygen, C$\alpha$, and C$\beta$ of Ser/Thr, and side-chain amide oxygen of Asn/Gln to an XYZ file. For Tyr systems, the coordinates of side-chain hydroxyl oxygen, C$\varepsilon$, and C$\zeta$ of Tyr, and side-chain amide oxygen of Asn/Gln are written to an XYZ file. Based on the following assumptions, a hydrogen atom was added to the hydroxyl oxygen atom in the XYZ file.

(i) O–H distance is 0.98 Å in Ser/Thr-Asn/Gln systems and 0.99 Å in Tyr-Asn/Gln systems as observed in the MP2/6-31G* optimized geometry of acetamide-methanol O–H···O HB.

(ii) The ∠Cβ–O–H angle is 106.4° in Ser/Thr-Asn/Gln systems and ∠Cζ–O–H angle is 108.7° in Tyr-Asn/Gln systems as observed in MP2/6-31G* optimized geometry of acetamide-methanol O–H···O HB.

(iii) C$\alpha$, C$\beta$, hydroxyl oxygen and hydrogen atoms of Ser/Thr lie in the same plane. C$\varepsilon$, Cζ, hydroxyl oxygen and hydrogen atoms of Tyr lie in the same plane.

Each of the above assumptions leads to one mathematical equation. Although, the third assumption is not true, assuming this results in three equations in three unknowns as elaborated below.

(i) x, y, and z coordinates of hydrogen atom must satisfy the equation of norm of $\overrightarrow{O-H}$.

$$||\overrightarrow{O-H}|| = \sqrt{(o_x - h_x)^2 + (o_y - h_y)^2 + (o_z - h_z)^2} = 0.98 \text{ Å (Ser/Thr systems)}$$

$$||\overrightarrow{O-H}|| = \sqrt{(o_x - h_x)^2 + (o_y - h_y)^2 + (o_z - h_z)^2} = 0.99 \text{ Å (Tyr systems)}$$

(ii) x, y, and z coordinates of hydrogen atom must satisfy the dot-product equation of C$\beta$–O and O–H bonds for Ser/Thr systems. Similarly, x, y, and z coordinates of hydrogen atom must satisfy the dot-product equation of Cζ–O and O–H bonds for Tyr systems.

$$\cos(\angle C\beta\text{–O–H}) = (\overrightarrow{C\beta - O} \cdot \overrightarrow{O-H})/(||\overrightarrow{C\beta - O}||\ ||\overrightarrow{O-H}||)$$

$$\cos(\angle C\zeta\text{–O–H}) = (\overrightarrow{C\zeta - O} \cdot \overrightarrow{O-H})/(||\overrightarrow{C\zeta - O}||\ ||\overrightarrow{O-H}||)$$

where $\overrightarrow{O-H} = (o_x - h_x, o_y - h_y, o_z - h_z)$.

(iii) We obtain the equation of plane formed by C$\alpha$, C$\beta$, and hydroxyl oxygen of Ser/Thr, and that formed by C$\varepsilon$, Cζ, and hydroxyl oxygen of Tyr. From the third assumption, the x, y, and z coordinates of hydrogen atom must satisfy this equation of plane.

$$ah_x + bh_y + ch_z + d = 0$$

where a, b, c, and d are known and $h_x$, $h_y$, and $h_z$ are the x, y, and z coordinates of the hydrogen atom.

The Python script solves the above three equations in three unknowns. The presence of square terms in the equation of norm leads to two sets of coordinates, of which the set of coordinates closest to the side-chain amide oxygen of Asn/Gln is selected using the script. With these coordinates of hydrogen atom as the starting point, the O–H bond is rotated around C$\beta$–O bond (Ser/Thr) or Cζ–O bond (Tyr) from 0° to 360° in steps of 0.1°. At each step of rotation, the distance between side-chain amide oxygen of Asn/Gln and hydrogen atom is measured. The coordinates of hydrogen atom in the configuration with shortest distance between hydrogen and amide oxygen are written to the XYZ file. The script then computes O–H···O HB angle using these hydrogen atom coordinates.

*Ambifunctional HBs.* For ambifunctional HBs, two hydrogen atoms were added to the sidechain amide nitrogen of Asn/Gln and one hydrogen atom was added to the sidechain hydroxyl oxygen of Ser/Thr/Tyr. The addition of amide hydrogen atoms was done as per the protocol mentioned above for adding hydrogen atoms to residue-pairs forming N–H···O HBs, and the addition of sidechain hydroxyl hydrogen was done following the protocol for adding hydrogen atoms to residue-pairs forming O–H···O HBs.

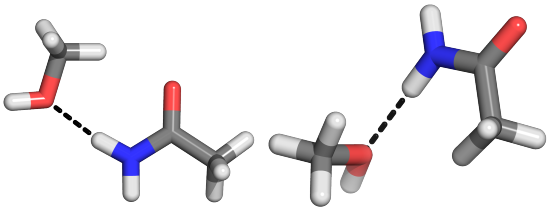


**Figure S1.** MP2/6-31G* geometry-optimized structures of acetamide–methanol N–H···O HBs. (a) *syn* N–H···O HB (b) *anti* N–H···O HB. Oxygen, nitrogen, carbon, and hydrogen atoms are shown in red, blue, gray, and white colors, respectively. The dashed lines indicate the N–H···O HB interactions.

**Table S3.** Interaction energies (in kcal/mol) with larger basis sets and higher levels of theory using the MP2/6-31G*-optimized geometries of all four acetamide–methanol and acetamide–*p*-cresol HB conformations: N–H···O *syn* HB, N–H···O *anti* HB, O–H···O HB, and ambifunctional (ambi.) HB. The interaction energy is obtained as the difference in energy of the dimer from that of the isolated molecules, all evaluated on MP2/6-31G*-optimized geometries. The two-point extrapolation formula based on the aug-cc-pVDZ and aug-cc-pVTZ energies is used to extrapolate to the limit following Refs. 6-8 for DLPNO-CCSD(T) and CCSD(T). Here, Normal refers to default thresholds of TCutPairs = $10^{-4}$, TCutPNO = 3.33 x $10^{-7}$, TCutMKN = $10^{-3}$, and Tight refers to default thresholds of TCutPairs = $10^{-5}$, TCutPNO = 1.00 x $10^{-7}$, TCutMKN = $10^{-3}$. The interaction energy of *syn* N–H···O HB obtained from CCSD(T)/CBS calculations was 0.5 kcal/mol stronger than that of *anti* N–H···O HB of acetamide–methanol, while the interaction energy of *anti* N–H···O HB obtained from DLPNO-CCSD(T)/CBS calculations is 0.1 kcal/mol stronger than that of *syn* N–H···O HB of acetamide–*p*-cresol. DLPNO-CCSD(T)/CBS calculations also reveal that *syn* and *anti* N–H···O HBs of acetamide–methanol are 0.9 kcal/mol and 0.3 kcal/mol stronger than those of acetamide–*p*-cresol, respectively.

| Method/Basis | Interaction Energy (kcal/mol) | | | |
|---|---|---|---|---|
| | N–H···O *syn* HB | N–H···O *anti* HB | O–H···O HB | Ambi. HB |
| acetamide–methanol | | | | |
| DLPNO-CCSD(T) / aug-cc-pVDZ (Normal) | -7.2 | -6.8 | -8.2 | -10.7 |
| DLPNO-CCSD(T) / aug-cc-pVDZ (Tight) | -7.5 | -6.9 | -8.4 | -10.9 |
| CCSD(T) / aug-cc-pVDZ | -7.6 | -7.1 | -8.5 | -11.2 |
| DLPNO-CCSD(T) / aug-cc-pVTZ (Normal) | -6.9 | -6.5 | -7.8 | -10.5 |
| DLPNO-CCSD(T) / aug-cc-pVTZ (Tight) | -7.1 | -6.6 | -8.0 | -10.7 |
| CCSD(T) / aug-cc-pVTZ | -7.2 | -6.7 | -8.1 | -10.9 |
| DLPNO-CCSD(T)/CBS (Tight) | -7.0 | -6.5 | -7.9 | -10.6 |
| CCSD(T)/CBS | -7.1 | -6.6 | -8.0 | -10.9 |
| acetamide–*p*-cresol | | | | |
| DLPNO-CCSD(T) / aug-cc-pVDZ (Normal) | -6.9 | -7.3 | -12.1 | -12.7 |
| DLPNO-CCSD(T) / aug-cc-pVDZ (Tight) | -7.2 | -7.4 | -12.4 | -13.1 |
| CCSD(T) / aug-cc-pVDZ | -7.4 | -7.6 | -12.7 | -13.5 |
| DLPNO-CCSD(T) / aug-cc-pVTZ (Normal) | -6.2 | -6.4 | -11.2 | -12.1 |
| DLPNO-CCSD(T) / aug-cc-pVTZ (Tight) | -6.4 | -6.5 | -11.4 | -12.5 |
| DLPNO-CCSD(T)/CBS (Tight) | -6.1 | -6.2 | -11.0 | -12.2 |

**Table S4.** Comparison of interaction energies (in kcal/mol) and optimized geometries (relevant distances in Å and angles in °) of all four acetamide–methanol HB configurations obtained using the listed methods with and without incorporating semi-empirical dispersion (D3[9]) for the hybrid DFT (B3LYP[10-12]) results. The counterpoise corrected results were obtained on the B3LYP-D3/6-31G* geometries. HB distances and angles from B3LYP and B3LYP-D3 calculations were found to differ by at most 0.05 Å and 6.4º, respectively, and the interaction energies from B3LYP-D3 calculations were stronger than those from B3LYP calculations by at most 2.3 kcal/mol. Comparison of optimized geometries obtained from B3LYP-D3 and MP2 calculations revealed that the HB distances and angles differed by at most 0.06 Å and 1.6º, respectively, and the B3LYP-D3 interaction energies were found to be stronger than those obtained from MP2 calculations by at most 1.9 kcal/mol. As a result, MP2-optimized geometries are used throughout in this work.

| Method/Basis | Int. E (kcal/mol) | O–H···O=C (Å) | H–O···H–N (Å) | O···O (Å) | N···O (Å) | ∠O–H···O (°) | ∠N–H···O (°) |
|---|---|---|---|---|---|---|---|
| N–H···O *syn* HB | | | | | | | |
| B3LYP-D3/6-31G* | -10.1 | -- | 1.91 | -- | 2.92 | -- | 169.8 |
| Counterpoise-corrected B3LYP-D3/6-31G* | -7.8 | -- | -- | -- | -- | -- | -- |
| B3LYP/6-31G* | -8.1 | -- | 1.95 | -- | 2.96 | -- | 170.6 |
| MP2/6-31G* | -9.1 | -- | 1.96 | -- | 2.97 | -- | 171.4 |
| N–H···O *anti* HB | | | | | | | |
| B3LYP-D3/6-31G* | -8.9 | -- | 1.94 | -- | 2.92 | -- | 161.1 |
| Counterpoise-corrected B3LYP-D3/6-31G* | -6.8 | -- | -- | -- | -- | -- | -- |
| B3LYP/6-31G* | -7.1 | -- | 1.97 | -- | 2.97 | -- | 167.5 |
| MP2/6-31G* | -8.0 | -- | 1.99 | -- | 2.97 | -- | 162.3 |
| O–H···O HB | | | | | | | |
| B3LYP-D3/6-31G* | -12.2 | 1.85 | -- | 2.80 | -- | 161.8 | -- |
| Counterpoise-corrected B3LYP-D3/6-31G* | -8.7 | -- | -- | -- | -- | -- | -- |
| B3LYP/6-31G* | -9.9 | 1.86 | -- | 2.82 | -- | 164.9 | -- |
| MP2/6-31G* | -10.5 | 1.90 | -- | 2.85 | -- | 162.7 | -- |
| Ambifunctional HB | | | | | | | |
| B3LYP-D3/6-31G* | -15.9 | 1.83 | 1.96 | 2.75 | 2.83 | 153.3 | 140.8 |
| Counterpoise-corrected B3LYP-D3/6-31G* | -12.3 | -- | -- | -- | -- | -- | -- |
| B3LYP/6-31G* | -13.6 | 1.86 | 1.98 | 2.78 | 2.85 | 152.8 | 141.4 |
| MP2/6-31G* | -14.0 | 1.90 | 2.01 | 2.81 | 2.88 | 152.6 | 141.8 |

**Table S5.** Comparison of interaction energies (in kcal/mol) and optimized geometries (relevant distances in Å and angles in °) of all four acetamide–*p*-cresol HB configurations obtained using the listed methods with and without incorporating semi-empirical dispersion (D3[9]) for the hybrid DFT (B3LYP[10-12]) results. The counterpoise corrected results were obtained on the B3LYP-D3/6-31G* geometries. B3LYP and B3LYP-D3 calculations were found to differ by at most 0.07 Å and 7.4°, respectively, and the interaction energies from B3LYP-D3 calculations were stronger than those from B3LYP calculations by at most 3.7 kcal/mol. Comparison of optimized geometries obtained from B3LYP-D3 and MP2 calculations revealed that the HB distances and angles differed by at most 0.05 Å and 1.0°, respectively, and the B3LYP-D3 interaction energies were found to be stronger than those obtained from MP2 calculations by at most 2.0 kcal/mol. For the ambifunctional HB in both model systems, the N···O HB distance and N–H···O HB angle are more distorted than the O···O HB distance and O–H···O HB angle with respect to their equilibrium values in a single HB configuration. This effect is most pronounced with *p*-cresol, which has both a shorter O···O HB distance and longer N···O HB distance (both by ca. 0.05 Å) and a more linear O–H···O HB angle at the cost of a less favorable N–H···O HB angle (both by 4°). These observations are in line with a shallower PEC of the N–H···O HB than that of O–H···O HB in *p*-cresol.

| Method/Basis | Int. E (kcal/mol) | O–H···O=C (Å) | H–O···H–N (Å) | O···O (Å) | N···O (Å) | ∠O–H···O (°) | ∠N–H···O (°) |
|---|---|---|---|---|---|---|---|
| N–H···O *syn* HB | | | | | | | |
| B3LYP-D3/6-31G* | -8.8 | -- | 2.01 | -- | 3.03 | -- | 177.0 |
| Counterpoise Corrected B3LYP-D3/6-31G* | -6.5 | -- | -- | -- | -- | -- | -- |
| B3LYP/6-31G* | -6.5 | -- | 2.06 | -- | 3.07 | -- | 177.9 |
| MP2/6-31G* | -8.3 | -- | 2.06 | -- | 3.07 | -- | 177.7 |
| N–H···O *anti* HB | | | | | | | |
| B3LYP-D3/6-31G* | -8.5 | -- | 2.03 | -- | 3.01 | -- | 163.9 |
| Counterpoise Corrected B3LYP-D3/6-31G* | -6.3 | -- | -- | -- | -- | -- | -- |
| B3LYP/6-31G* | -5.1 | -- | 2.07 | -- | 3.08 | -- | 171.3 |
| MP2/6-31G* | -7.7 | -- | 2.07 | -- | 3.06 | -- | 164.8 |
| O–H···O HB | | | | | | | |
| B3LYP-D3/6-31G* | -15.5 | 1.78 | -- | 2.75 | -- | 165.9 | -- |
| Counterpoise Corrected B3LYP-D3/6-31G* | -12.0 | -- | -- | -- | -- | -- | -- |
| B3LYP/6-31G* | -11.8 | 1.80 | -- | 2.77 | -- | 167.7 | -- |
| MP2/6-31G* | -13.5 | 1.83 | -- | 2.80 | -- | 166.0 | -- |
| Ambifunctional HB | | | | | | | |
| B3LYP-D3/6-31G* | -17.2 | 1.77 | 2.07 | 2.71 | 2.90 | 157.1 | 136.7 |
| Counterpoise Corrected B3LYP-D3/6-31G* | -13.6 | -- | -- | -- | -- | -- | -- |
| B3LYP/6-31G* | -14.0 | 1.79 | 2.12 | 2.73 | 2.95 | 157.3 | 136.9 |
| MP2/6-31G* | -15.5 | 1.82 | 2.11 | 2.76 | 2.94 | 156.6 | 137.7 |

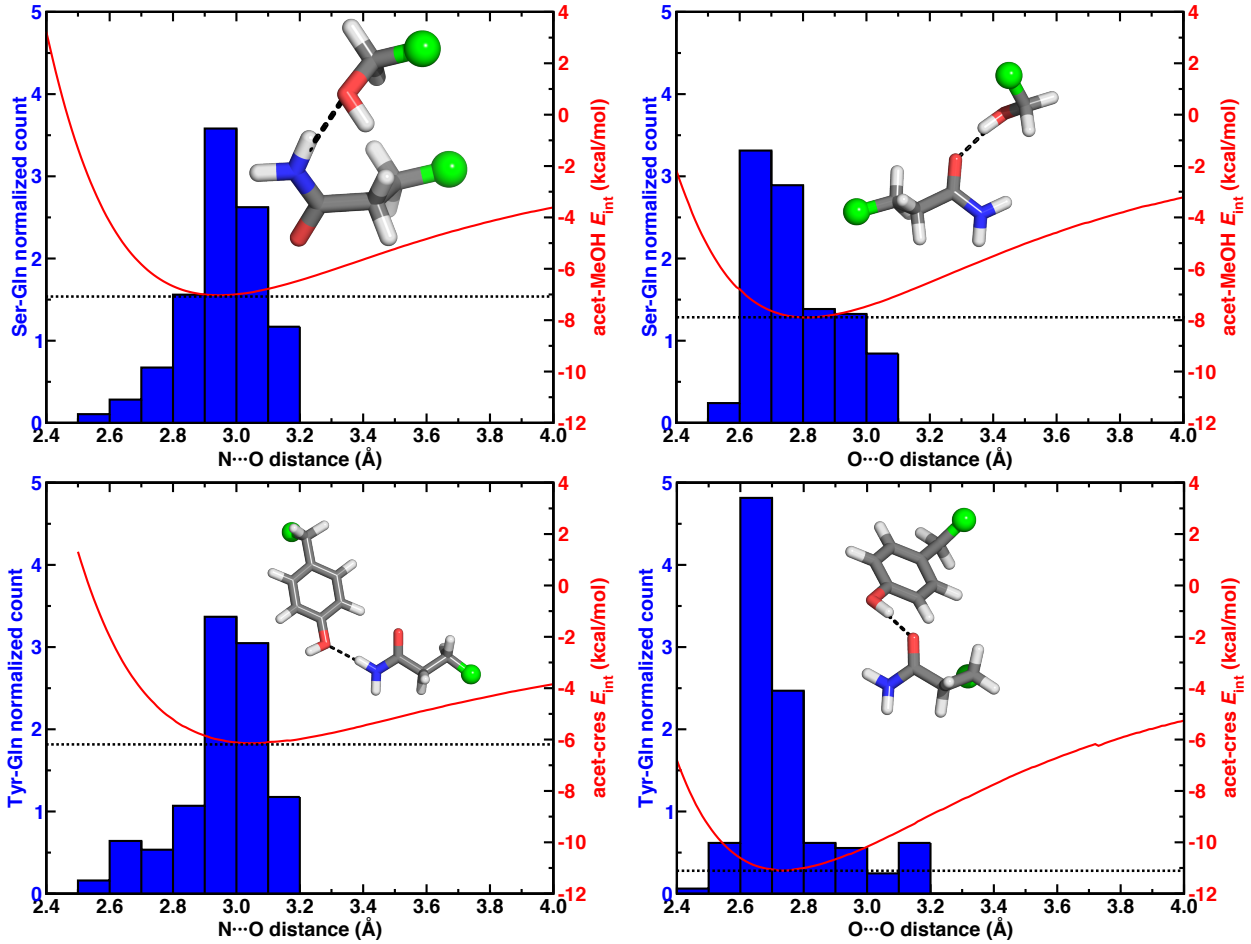


**Figure S2.** Normalized histograms (blue, left axes) of heavy-atom HB distances (in Å, bin width of 0.1 Å) for Ser-Gln (top) and Tyr-Gln (bottom) X-ray crystal structures with the 1D PECs (red, right axes) for acetamide–methanol (top) and acetamide–*p*-cresol (bottom) overlaid. The left panes show the N···O HB distance histograms and PECs of *syn* N–H···O HBs, and the right panes show the O···O HB distance histograms and PECs of O–H···O HBs. The interaction energies ($E_{int}$) shown are obtained from DLPNO-CCSD(T)/CBS calculations. The structure insets are representative PDB structures for the relevant HB where the HB distance is indicated by black dotted lines. Cα of the residues in the insets is represented as a green sphere indicating that the residues are truncated to show only the side-chains, and the remaining atoms in the side-chains are shown as sticks with carbon in gray, hydrogen in white, nitrogen in blue, and oxygen in red. Cα of the residues in the insets is represented as a green sphere indicating that the residues are truncated to show only the side-chains, and the remaining atoms in the side-chains are shown as sticks with carbon in gray, hydrogen in white, nitrogen in blue, and oxygen in red.

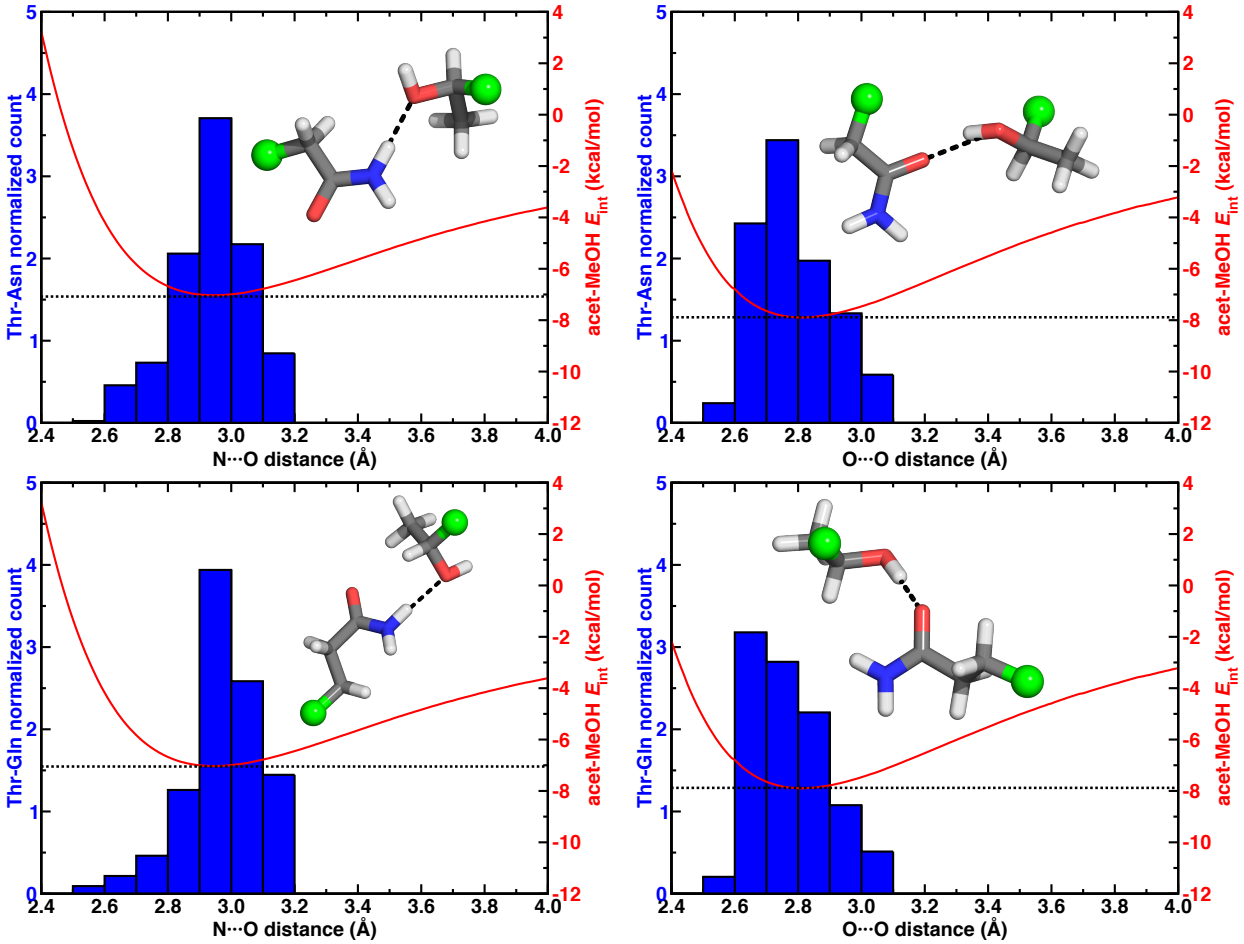


**Figure S3.** Normalized histograms (blue, left axes) of heavy-atom HB distances (in Å, bin width of 0.1 Å) for Thr-Asn (top) and Thr-Gln (bottom) X-ray crystal structures with the 1D PECs (red, right axes) for acetamide–methanol (top and bottom) overlaid. The left panes show the N···O HB distance histograms and PECs of *syn* N–H···O HBs, and the right panes show the O···O HB distance histograms and PECs of O–H···O HBs. The interaction energies ($E_{int}$) shown are obtained from DLPNO-CCSD(T)/CBS calculations. The structure insets are representative PDB structures for the relevant HB where the HB distance is indicated by black dotted lines. Cα of the residues in the insets is represented as a green sphere indicating that the residues are truncated to show only the side-chains, and the remaining atoms in the side-chains are shown as sticks with carbon in gray, hydrogen in white, nitrogen in blue, and oxygen in red.

**Text S3.** Presence of a higher number of *anti* N–H···O HBs than *syn* N–H···O HBs in the protein data set.

Given the ability of our models to recapitulate key differences between Ser/Thr and Tyr residues, we return to the question of the extent to which the models could capture other kcal/mol-scale trends. Specifically, the DLPNO-CCSD(T)/CBS model energetics for methanol (i.e., representing Ser/Thr) had exhibited a weak preference for *syn* over *anti* N–H···O HBs, whereas they were degenerate for the Tyr-model *p*-cresol (main text Figure 2 and ESI Table S3). Analysis of the number of *syn* and *anti* N–H···O HB structures from the protein data set indicates a weak preference towards the formation of *anti* N–H···O HBs over *syn* N–H···O HBs for all amino acid pairs (ESI Table S6). These differences could arise due to our neglect of the

protein environment, e.g., from the constraining presence of the protein backbone or additional backbone–backbone HBs.

We thus compared our model acetamide–methanol *syn* and *anti* N–H···O HB interaction energies to those of full Ser and Asn residues (see Sec. 5). The full Ser-Asn model *anti* N–H···O HB is stronger than that of the *syn* N–H···O HB due to the presence of two additional backbone HBs in the *anti* conformation (by ca. 4.6 kcal/mol, ESI Table S7 and Figure S4). Over the full protein data set, we observed similar additional HB interactions in many of the *anti* N–H···O HB conformations (ESI Figure S5). Hence, while our models accurately capture key structural features, the higher abundance of *anti* N–H···O HBs in X-ray crystal structures is likely due to their higher compatibility with simultaneous backbone–backbone stabilization.

From the DLPNO-CCSD(T)/CBS results and the MP2/6-31G* optimized geometries, we inferred that additional HB interactions are present in *anti* N–H···O HBs (ESI Table S7 and Figure S4). Given that these interactions are observed between the backbones of the hydrogen bonding residues, it is very likely that *anti* N–H···O HBs are observed between residues on adjacent beta sheets or loops (ESI Figure S5). Since the backbones of residues were not observed to hydrogen bond in *syn* N–H···O HB, it is likely that the *syn* N–H···O HBs are observed between residues with one residue on an alpha helix while the other is on a beta sheet or a loop (ESI Figure S5). These features may not be observed in all the *syn* and *anti* N–H···O HBs, but they might be part of why we see more *anti* N–H···O HBs than *syn* N–H···O HBs despite similar energetics in model systems.

**Table S6.** Number of *syn* and *anti* N–H···O HBs in Ser-Asn, Ser-Gln, Thr-Asn, Thr-Gln, Tyr-Asn, and Tyr-Gln residue pairs in the protein data set, mean N–H···O angle, and standard deviation of N–H···O angle of N–H···O HBs.

| Residue pair | Number of *syn* N–H···O HBs | Number of *anti* N–H···O HBs | Mean N–H···O HB angle | Standard deviation |
|---|---|---|---|---|
| Ser-Asn | 153 | 260 | 153.3° | 16.7° |
| Ser-Gln | 73 | 209 | 153.4° | 14.6° |
| Thr-Asn | 157 | 280 | 156.6° | 16.2° |
| Thr-Gln | 85 | 240 | 155.0° | 16.4° |
| Tyr-Asn | 87 | 187 | 152.4° | 15.7° |
| Tyr-Gln | 54 | 133 | 154.5° | 15.7° |

**Table S7.** Comparison of interaction energies (Int. E, in kcal/mol) of *syn* and *anti* N–H···O HBs of Ser and Asn with N-terminal acetyl (ACE) and C-terminal N-methyl (NME) capping for both the residues. The interaction energy is obtained as the difference in energy of the dimer from that of the isolated molecules, all evaluated on MP2/6-31G*-optimized geometries. The two-point extrapolation formula based on the aug-cc-pVDZ and aug-cc-pVTZ energies is used to extrapolate to the complete basis set limit following Refs. [6-8] for DLPNO-CCSD(T) and canonical CCSD(T). Here, Tight refers to default thresholds of TCutPairs = $10^{-5}$, TCutPNO = $1.00 \times 10^{-7}$, and TCutMKN = $10^{-3}$.

| Method/Basis | *syn* N–H···O HB Int. E (kcal/mol) | *anti* N–H···O HB Int. E (kcal/mol) | *syn-anti* (kcal/mol) |
|---|---|---|---|
| DLPNO-CCSD(T) / aug-cc-pVDZ (Tight) | -15.7 | -21.7 | -6.0 |
| DLPNO-CCSD(T) / aug-cc-pVTZ (Tight) | -15.0 | -20.1 | -5.1 |
| DLPNO-CCSD(T)/CBS (Tight) | -14.8 | -19.4 | -4.6 |

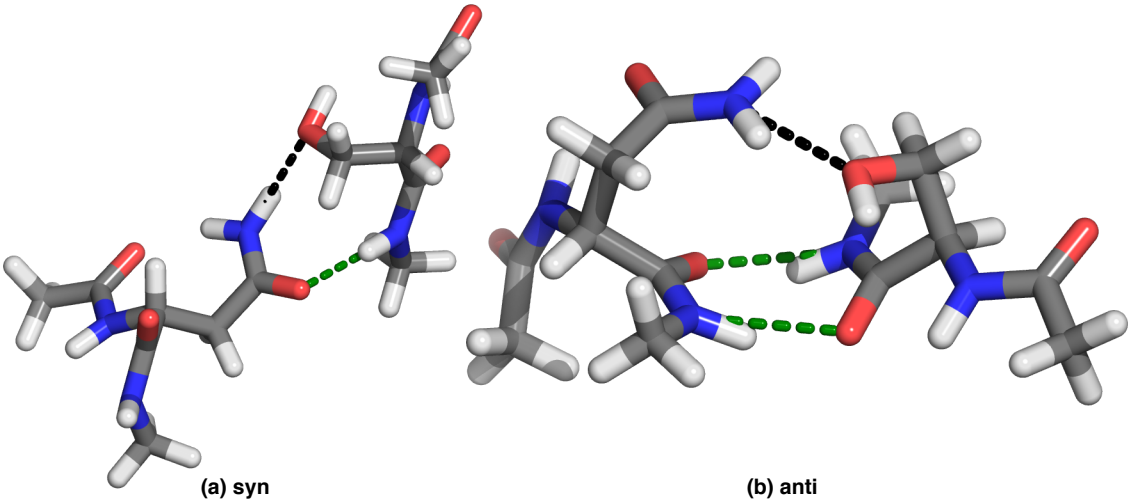


**Figure S4.** MP2/6-31G* geometry-optimized structures of Ser-Asn N–H···O HBs with N-terminal acetyl (ACE) and C-terminal N-methyl (NME) capping for backbones of both residues: (a) *syn* N–H··O HB (b) *anti* N–H···O HB. Oxygen, nitrogen, carbon, and hydrogen atoms are shown in red, blue, gray, and white colors, respectively. The black dashed lines indicate the N–H···O HB of Ser and Asn sidechains. The green dashed lines indicate the additional HB interactions between the residues. In the case of *syn* N–H···O HB, an additional HB is present between sidechain amide carbonyl oxygen of Asn and backbone N–H hydrogen of Ser. For *anti* N–H···O HB, additional HBs are observed between the backbone carbonyl oxygen and N–H hydrogen of both the residues.

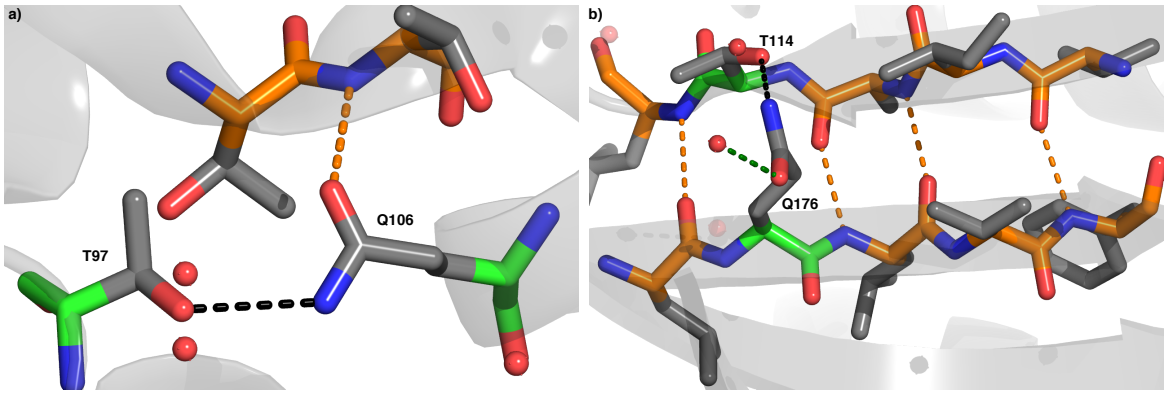


**Figure S5.** Representative proteins showing the HB interactions observed in the surrounding protein environment of N–H···O HBs between Thr and Gln. The protein cartoon is shown in translucent gray. The hydrogen bonding residues Thr and Gln are shown in green sticks while the nearby residues involved in HB interactions are shown in orange sticks. Oxygen and nitrogen atoms are shown in red and blue, respectively. Sidechain carbon atoms are shown in gray. (a) *syn* N–H···O HB between Thr located on a loop and Gln located on an alpha helix in the protein (PDB ID: 4URF) is shown in black dashed lines. HB between sidechain amide carbonyl oxygen of Gln and a nearby backbone N–H hydrogen is shown in orange dashed lines. (b) *anti* N–H···O HB between Thr and Gln in the protein (PDB ID: 4MHP) is shown in black dashed lines. Thr and Gln are located on adjacent beta sheets. The HBs between the backbone carbonyl oxygen and N–H hydrogen of the adjacent beta sheets are shown in orange dashed lines. The HB of Gln sidechain amide carbonyl oxygen with a nearby solvent molecule (red sphere) is shown in green

dashed lines. Thr and Gln residues are labeled with the one-letter amino acid code followed by the residue number.

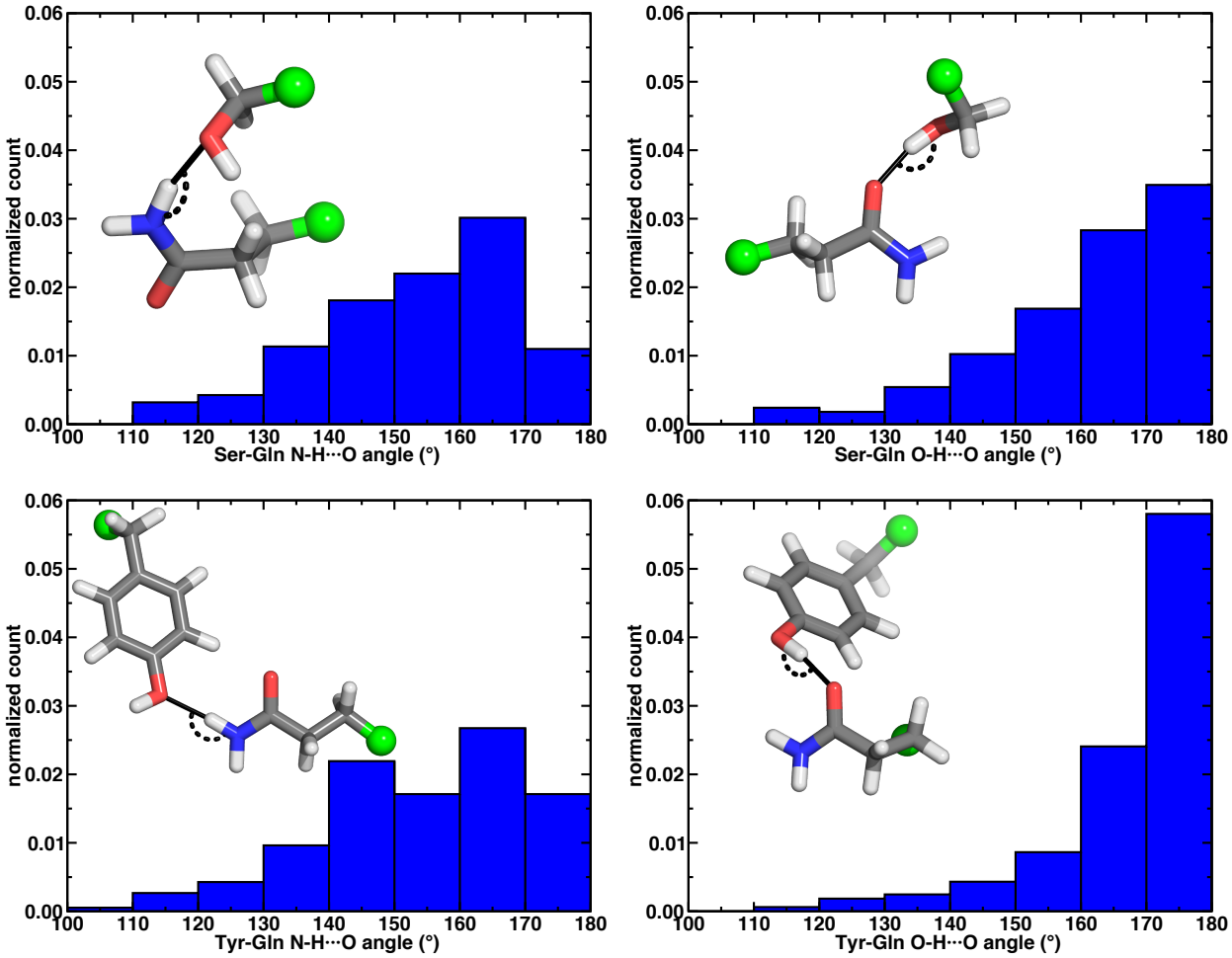


**Figure S6.** Normalized histograms of N–H···O HB (left) and O–H···O HB (right) angles (in °) for Ser-Gln (top) and Tyr-Gln (bottom) residue pairs from X-ray crystal structures. All histograms have 10° bin widths. The insets depict the HB angle on representative PDB structures with the corresponding Cα of the residues represented as a green sphere, and the remaining atoms are shown as sticks with carbon in gray, hydrogen in white, nitrogen in blue, and oxygen in red.

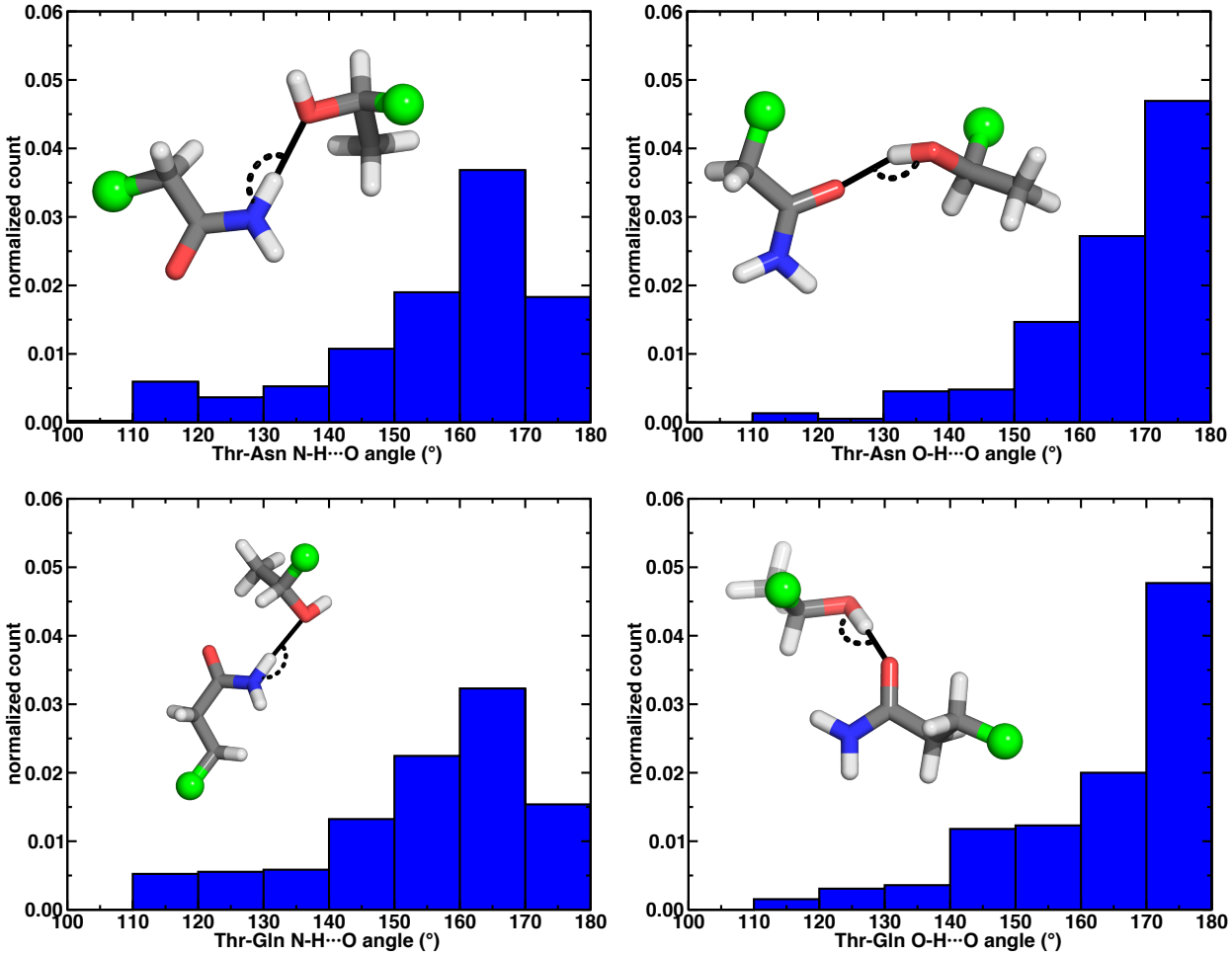


**Figure S7.** Normalized histograms of N–H···O HB (left) and O–H···O HB (right) angles (in °) for Thr-Asn (top) and Thr-Gln (bottom) residue pairs from X-ray crystal structures. All histograms have 10° bin widths. The insets depict the HB angle on representative PDB structures with the corresponding Cα of the residues represented as a green sphere, and the remaining atoms are shown as sticks with carbon in gray, hydrogen in white, nitrogen in blue, and oxygen in red.

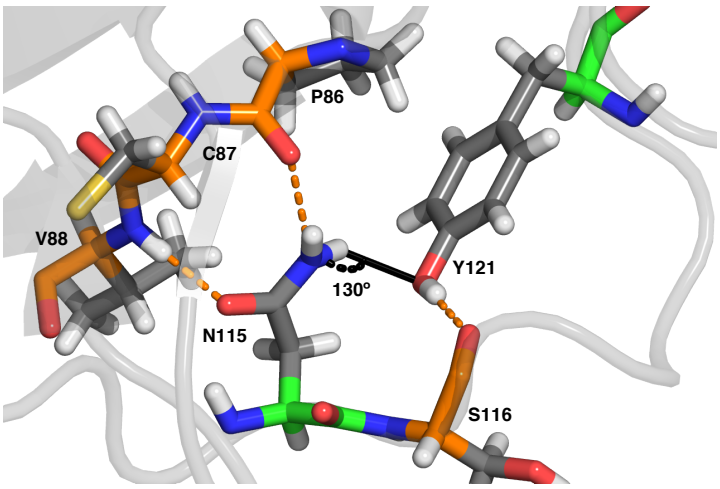


**Figure S8.** A representative protein (PDB ID: 2XJP) showing the additional stabilizing HB interactions observed in the surrounding protein environment of an N–H···O HB between Tyr and Asn (HB angle shown in black dashed lines). The protein cartoon is shown in translucent gray. The hydrogen bonding residues Tyr and Asn are shown in green sticks while the nearby residues involved in additional HB interactions with Asn and Tyr are shown in orange sticks. Oxygen, nitrogen, and hydrogen atoms are shown in red, blue, and white, respectively. Sidechain carbon atoms are shown in gray. Tyr, Asn, Pro, Cys, Val, and Ser residues are labeled with the one-letter amino acid code followed by the residue number. HBs between the Asn sidechain amide carbonyl oxygen and a nearby Val backbone N–H hydrogen, the Asn sidechain amide hydrogen and a nearby Pro backbone carbonyl oxygen, the Tyr sidechain hydroxyl hydrogen and a nearby Ser backbone carbonyl oxygen are shown in orange dashed lines.

**Table S8.** Comparison of optimized geometries (relevant distances in Å and angles in °) of all four acetamide–methanol and acetamide-*p*-cresol HB configurations obtained using MP2/6-31G* geometry optimization using implicit solvent correction with ε =10, which was selected to approximately mimic the protein environment. The difference in HB distances and angles of optimized geometries that incorporate implicit solvent correction and those that are gas-phase is also given in this table. The differences are computed as solvent corrected geometric parameter – gas-phase geometric parameter, with appropriate units.

| Type of HB | O–H···O=C (Å) | H–O···H–N (Å) | O···O (Å) | N···O (Å) | ∠O–H···O (°) | ∠N–H···O (°) |
|---|---|---|---|---|---|---|
| acetamide-methanol | | | | | | |
| N–H···O *syn* HB | -- | 1.94 | -- | 2.96 | -- | 173.0 |
| N–H···O *anti* HB | -- | 1.94 | -- | 2.95 | -- | 172.1 |
| O–H···O HB | 1.84 | -- | 2.82 | -- | 172.1 | -- |
| Ambifunctional HB | 1.86 | 2.12 | 2.80 | 2.96 | 157.8 | 138.6 |
| acetamide-*p*-cresol | | | | | | |
| N–H···O *syn* HB | -- | 2.02 | -- | 3.04 | -- | 179.5 |
| N–H···O *anti* HB | -- | 2.04 | -- | 3.04 | -- | 169.5 |
| O–H···O HB | 1.76 | -- | 2.75 | -- | 174.9 | -- |
| Ambifunctional HB | 1.77 | 2.28 | 2.74 | 3.07 | 163.7 | 133.4 |
| Difference in geometries of acetamide-methanol | | | | | | |
| N–H···O *syn* HB | -- | -0.02 | -- | -0.01 | -- | 1.6 |
| N–H···O *anti* HB | -- | -0.05 | -- | -0.02 | -- | 9.8 |
| O–H···O HB | -0.06 | -- | -0.03 | -- | 9.4 | -- |
| Ambifunctional HB | -0.04 | 0.11 | -0.01 | 0.08 | 5.2 | -3.2 |
| Difference in geometries of acetamide-*p*-cresol | | | | | | |
| N–H···O *syn* HB | -- | -0.04 | -- | 0.03 | -- | 1.8 |
| N–H···O *anti* HB | -- | -0.03 | -- | -0.02 | -- | 4.7 |
| O–H···O HB | -0.07 | -- | -0.05 | -- | 8.9 | -- |
| Ambifunctional HB | -0.05 | 0.17 | -0.02 | 0.13 | 7.1 | -4.3 |

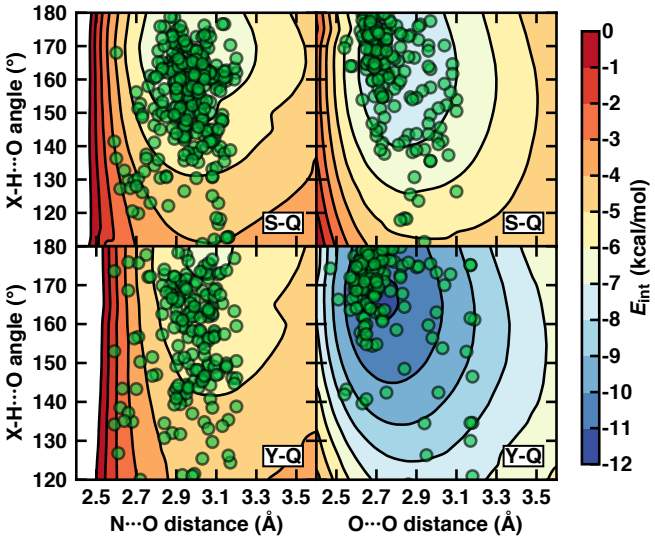


**Figure S9.** 2D PESs depicting interaction energies ($E_{int}$ in kcal/mol, colorbar at right) of N–H···O HBs (left) and O–H···O HBs (right) in acetamide–methanol (top) and acetamide–*p*-cresol (bottom). The heavy-atom (i.e., N···O and O···O) distances (in Å) and X–H···O angles (in °) are

shown as labeled on the axes, where X–H···O corresponds to N–H···O (left) or O–H···O (right). The same color scale is used for all inset PESs with 1 kcal/mol contour lines. The X-ray crystal structure distances and angles (translucent green circles) from the data set are overlaid onto the PESs for the corresponding Ser-Gln (labeled as S-Q) and Tyr-Gln (labeled as Y-Q) residue pairs.

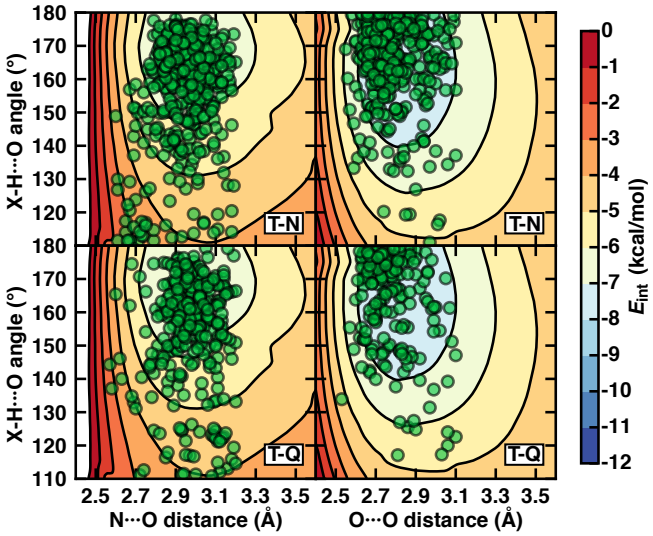


**Figure S10.** 2D PESs depicting interaction energies ($E_{int}$ in kcal/mol, colorbar at right) of N–H···O HBs (left) and O–H···O HBs (right) in acetamide-methanol (top and bottom). The heavy-atom (i.e., N···O and O···O) distances (in Å) and X–H···O angles (in °) are shown as labeled on the axes, where X–H···O corresponds to N–H···O (left) or O–H···O (right). The same color scale is used for all inset PESs with 1 kcal/mol contour lines. The X-ray crystal structure distances and angles (translucent green circles) from the data set are overlaid onto the PESs for the corresponding Thr-Asn (labeled as T-N) and Thr-Gln (labeled as T-Q) residue pairs.

**Table S9.** Distribution of PDB structures inside and outside the first two contours of the 2D potential energy surfaces of N–H···O and O–H···O HBs for Ser-Asn, Ser-Gln, Thr-Asn, Thr-Gln, Tyr-Asn and Tyr-Gln residue pairs in the protein data set, i.e., the number of PDB structures within 2 kcal/mol of the strongest interaction energy (columns 2 and 4) and the number of PDB structures outside this energy range (columns 3 and 5). Columns 6 and 7 indicate the number of PDB structures with less favorable HB angles (i.e., between 110° and 130°) that correspond to less favorable model interaction energies due to a short N···O distance but relatively long HBD to HBA (i.e., H···O) distance. The H···O distance in these N–H···O HBs ranges from 1.81 Å to 2.75 Å, and in O–H···O HBs, it ranges from 2.03 Å to 2.61 Å.

| Residue pair | No. of N–H···O HBs inside | No. of N–H···O HBs outside | No. of O–H···O HBs inside | No. of O–H···O HBs outside | No. of N–H···O HBs with 110-130° angles | No. of O–H···O HBs with 110-130° angles |
|---|---|---|---|---|---|---|
| Ser-Asn | 345 | 68 | 289 | 24 | 50 | 19 |
| Ser-Gln | 246 | 36 | 159 | 7 | 21 | 7 |
| Thr-Asn | 386 | 51 | 367 | 8 | 43 | 7 |
| Thr-Gln | 278 | 47 | 187 | 8 | 35 | 9 |
| Tyr-Asn | 190 | 84 | 186 | 34 | 32 | 6 |
| Tyr-Gln | 129 | 58 | 138 | 24 | 14 | 4 |

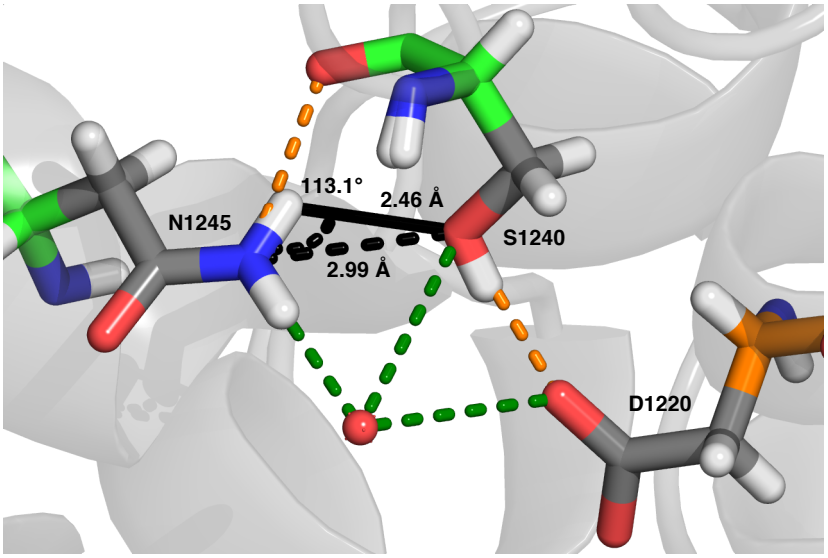


**Figure S11.** A representative protein (PDB ID: 4F1V) showing the additional stabilizing HB interactions observed in the surrounding protein environment of the N–H···O HB between Ser and Asn. The protein cartoon is shown in translucent gray. The hydrogen bonding residues Ser and Asn are shown in green sticks while the nearby Asp residue involved in HB interaction with Ser is shown in orange sticks. Oxygen, nitrogen, and hydrogen atoms are shown in red, blue, and white, respectively. Sidechain carbon atoms are shown in gray. The N···O HB distance and N–H···O HB angle are shown in black dashed lines. The N···O and H···O HB distances are 2.99 Å and 2.46 Å, respectively, and the N–H···O HB angle is 113.1°. Ser, Asn, and Asp residues are labeled with the one-letter amino acid code followed by their residue numbers. Orange dashed lines are shown for the HB between Asn sidechain amide hydrogen and Ser backbone carbonyl oxygen, HB between Ser sidechain hydroxyl hydrogen, and a nearby Asp sidechain carboxylate oxygen. Green dashed lines are shown for the HB of an Asn sidechain amide hydrogen atom with a solvent molecule, a HB between a Ser sidechain hydroxyl oxygen and the solvent molecule, and a HB between an Asp sidechain carboxylate oxygen with the solvent molecule.

**Table S10.** Mean O–H···O angle (in °) and standard deviation (in °) of O–H···O angle of O–H···O HBs in Ser-Asn, Ser-Gln, Thr-Asn, Thr-Gln, Tyr-Asn, and Tyr-Gln residue pairs in the protein data set.

| Residue pair | Mean O–H···O HB angle | Standard deviation |
|---|---|---|
| Ser-Asn | 163.7° | 15.4° |
| Ser-Gln | 161.5° | 14.8° |
| Thr-Asn | 165.6° | 13.0° |
| Thr-Gln | 163.5° | 14.8° |
| Tyr-Asn | 168.3° | 11.2° |
| Tyr-Gln | 168.0° | 12.4° |

**Table S11.** Comparison of different components of symmetry-adapted perturbation theory (SAPT) energies (in kcal/mol) obtained using MP2/6-31G*-optimized geometries for all four acetamide–methanol and acetamide–*p*-cresol HB configurations and also the HB configurations at the top of the energy barriers in the reaction coordinate (RC) plots evaluated at SAPT2+3/aug-cc-pVTZ level of theory, as implemented in Psi4.

| HB configuration | Electrostatics (kcal/mol) | Exchange (kcal/mol) | Induction (kcal/mol) | Dispersion (kcal/mol) | Total SAPT2+3 (Col. 1+2+3+4, in kcal/mol) |
|---|---|---|---|---|---|
| acetamide–methanol | | | | | |
| *syn* N–H···O | -10.2 | 11.3 | -3.7 | -4.9 | -7.5 |
| *anti* N–H···O | -8.3 | 8.5 | -2.5 | -4.2 | -6.5 |
| O–H···O | -12.6 | 14.3 | -4.5 | -5.6 | -8.4 |
| ambifunctional | -19.2 | 22.1 | -7.0 | -7.3 | -11.4 |
| HB at RC energy barrier | -7.6 | 7.2 | -2.6 | -3.4 | -6.4 |
| acetamide–*p*-cresol | | | | | |
| *syn* N–H···O | -8.4 | 9.8 | -3.1 | -5.0 | -6.7 |
| *anti* N–H···O | -7.1 | 8.6 | -2.0 | -5.7 | -6.2 |
| O–H···O | -16.0 | 18.7 | -6.4 | -8.4 | -12.1 |
| ambifunctional | -20.2 | 23.5 | -8.2 | -8.7 | -13.6 |
| HB at RC energy barrier | -9.9 | 9.8 | -4.0 | -4.9 | -9.0 |

**Table S12.** Number of N–H···O HBs, O–H···O HBs, and ambifunctional HBs in Ser-Asn, Ser-Gln, Thr-Asn, Thr-Gln, Tyr-Asn, and Tyr-Gln residue pairs in the protein data set.

| Residue pair | No. of N–H···O HBs | No. of O–H···O HBs | No. of ambifunctional HBs |
|---|---|---|---|
| Ser-Asn | 413 | 313 | 130 |
| Ser-Gln | 282 | 166 | 78 |
| Thr-Asn | 437 | 375 | 176 |
| Thr-Gln | 325 | 195 | 62 |
| Tyr-Asn | 274 | 220 | 68 |
| Tyr-Gln | 187 | 162 | 45 |

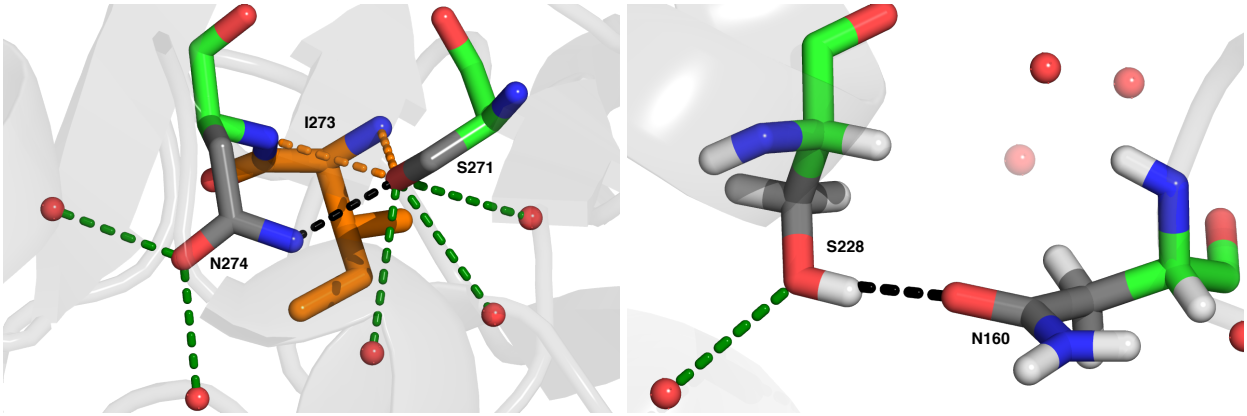


**Figure S12.** Representative proteins showing the HB interactions observed in the surrounding protein environment of N–H···O (left, PDB ID: 3VLA) and O–H···O (right, PDB ID: 3ZOJ) HBs between Ser and Asn. The protein cartoon is shown in translucent gray. The hydrogen bonding residues Ser and Asn are shown in green sticks while the nearby residues involved in HB interactions are shown in orange sticks. Oxygen atoms are shown in red, nitrogen in blue, sidechain carbon atoms in gray, and hydrogen atoms are shown in white. Ser, Asn, and Ile residues are labeled with the one-letter amino acid code followed by their residue numbers. In addition to the N–H···O HB between Ser and Asn (left, shown with black dashed lines), the sidechain hydroxyl of Ser interacts with two nearby backbone N–H (shown in orange dashed lines) and three nearby solvent molecules (shown in green dashed lines), while the sidechain amide oxygen of Asn interacts with two nearby solvent molecules (shown in green dashed lines). In addition to the O–H···O HB between Ser and Asn (right, shown with black dashed lines), there is only one other interaction of sidechain hydroxyl oxygen of Ser with a nearby solvent molecule (shown in green dashed lines).

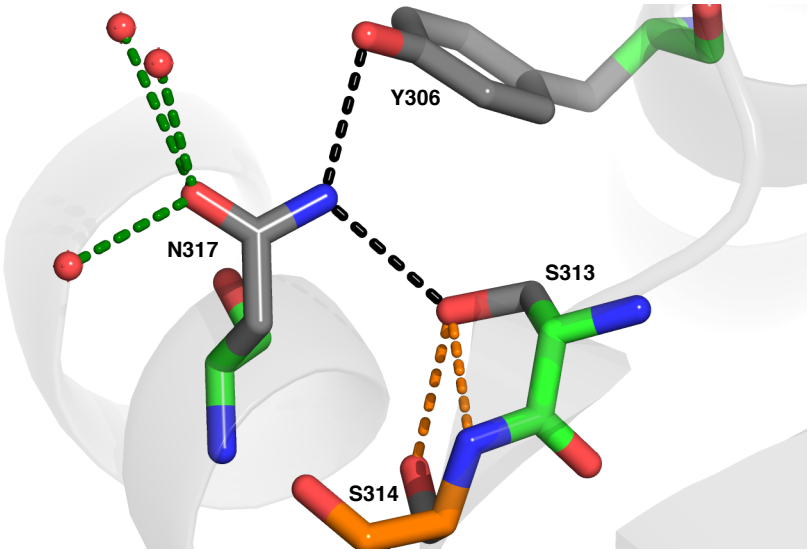


**Figure S13.** A representative protein (PDB ID: 1XMK) showing the presence of two simultaneous N–H···O HBs of Asn with Ser (*anti* N–H···O HB) and Tyr (*syn* N–H···O HB) residues (shown in black dashed lines). The protein cartoon is shown in translucent gray. The hydrogen bonding residues Ser, Tyr and Asn are shown in green sticks while the nearby Ser (S314) residue involved in a HB interaction with Ser (S313) is shown in orange sticks. Oxygen

atoms are shown in red, nitrogen atoms in blue, and sidechain carbon atoms are shown in gray. Ser, Asn, and Tyr residues are labeled with the one-letter amino acid code followed by their residue numbers. Green dashed lines represent the HB interactions of the Asn sidechain amide oxygen with three nearby solvent molecules. Orange dashed lines represent the HB interactions of sidechain hydroxyl group of Ser with nearby backbone nitrogen and sidechain hydroxyl oxygen atoms.

**Table S13.** O···O and N···O HB distances observed in ambifunctional HBs of acetamide–methanol and acetamide–*p*-cresol and the corresponding O–H···O and N–H···O HB angles, compared alongside the O–H···O and N–H···O HB interaction energies in the respective single O–H···O and N–H···O HBs, obtained from the optimized geometries on their 1D PECs.

| O···O dist. (Å) | N···O dist. (Å) | O–H···O angle (°) | N–H···O angle (°) | $E_{int}$ O–H···O (kcal/mol) | $E_{int}$ N–H···O (kcal/mol) |
|---|---|---|---|---|---|
| **Acetamide–methanol** | | | | | |
| 2.81 | 2.88 | 162.2 | 171.1 | -7.9 | -7.0 |
| **Acetamide–*p*-cresol** | | | | | |
| 2.76 | 2.94 | 166.3 | 176.5 | -11.1 | -6.0 |

**Table S14.** Comparison of zero point vibrational energies (ZPE in units of kcal/mol), ZPE + total thermal correction (E(ZPE) + E(trans) + E(rot) + E(vib)) to the electronic energy ($E_{el}$), entropic contribution (TS in units of kcal/mol, where T = 298.15 K), and the energy required to transform $E_{el}$ into Gibbs free energy (G-$E_{el}$), i.e., E(ZPE) + E(trans) + E(rot) + E(vib) + $k_BT$, where $k_BT$ is the thermal enthalpy correction with a value of 0.6 kcal/mol, for all four acetamide–methanol and acetamide–*p*-cresol HB configurations obtained using MP2/6-31G*-optimized geometries evaluated at the MP2/6-31G* level of theory. We also report the net effect of G-$E_{el}$ relative to *syn* N–H···O energies that Gibbs free energy corrections are expected to have on electronic energies.

| HB configuration | ZPE (kcal/mol) | ZPE + total thermal correction (kcal/mol) | Entropic contribution (TS) (kcal/mol) | ZPE+total thermal + enthalpy – entropic corrections (kcal/mol) |
|---|---|---|---|---|
| acetamide–methanol | | | | |
| *syn* N–H···O | 81.4 | 87.0 | 27.0 | 60.6 (0.0) |
| *anti* N–H···O | 81.4 | 87.1 | 27.4 | 60.3 (-0.3) |
| O–H···O | 81.6 | 87.1 | 26.9 | 60.8 (0.2) |
| ambifunctional | 82.2 | 88.0 | 27.3 | 61.3 (0.7) |
| acetamide–*p*-cresol | | | | |
| *syn* N–H···O | 131.2 | 139.9 | 34.8 | 105.7 (0.0) |
| *anti* N–H···O | 131.5 | 140.6 | 35.1 | 106.1 (0.4) |
| O–H···O | 132.0 | 140.9 | 34.5 | 107.0 (1.3) |
| ambifunctional | 132.3 | 141.0 | 34.4 | 107.2 (1.5) |

**Table S15.** Total energetic penalty ($\Delta E_{total}$) in ambifunctional HBs due to shorter HB distances and smaller HB angles. The energetic cost due to shorter HB distances in ambifunctional HBs was evaluated using energies of single HBs with HB distances observed in ambifunctional HBs but near-linear HB angles observed in freely optimized geometries of single HBs, obtained from the 2D energy surfaces ($E_{dist}$, column 3). We then evaluated the difference in energies between $E_{dist}$ and the energies of freely optimized single HBs ($E_{single}$, column 2) in both the model systems. We used a similar approach to evaluate the penalty related to smaller HB angles, i.e., we obtained energies of single HBs with HB angles observed in ambifunctional HBs but HB distances observed in freely optimized single HBs from the 2D contours ($E_{angle}$, column 5). We then evaluated the difference in energies between $E_{angle}$ and the energies of single HBs in both the model systems.

| HB | $E_{single}$ (kcal/mol) | $E_{dist}$ (kcal/mol) | Dist. (Å), angle (°) of $E_{dist}$ | $E_{angle}$ (kcal/mol) | Dist. (Å), angle (°) of $E_{angle}$ | $\Delta E_{dist}$ (kcal/mol) | $\Delta E_{angle}$ (kcal/mol) | $\Delta E_{total}$ (kcal/mol) |
|---|---|---|---|---|---|---|---|---|
| **Acetamide–methanol** | | | | | | | | |
| N–H···O | -6.9 | -6.9 | 2.90, 170 | -5.6 | 3.00, 140 | 0.0 | 1.3 | 1.5 |
| O–H···O | -7.9 | -7.9 | 2.80, 165 | -7.7 | 2.85, 155 | 0.0 | 0.2 | |
| **Acetamide–*p*-cresol** | | | | | | | | |
| N–H···O | -6.1 | -6.0 | 2.95, 180 | -4.9 | 3.05, 140 | 0.1 | 1.2 | 1.5 |
| O–H···O | -11.0 | -11.1 | 2.75, 165 | -10.7 | 2.80, 155 | -0.1 | 0.3 | |

**Table S16.** Comparison of generalized amber force field (GAFF)[13] and DLPNO-CCSD(T)/CBS (using the default Tight PNO thresholds of TCutPairs = $10^{-5}$, TCutPNO = 1.00 x $10^{-7}$, TCutMKN = $10^{-3}$) energetics for select intermediates along the reaction coordinate: the *syn* N–H···O, O–H···O, ambifunctional (ambi) intermediates, the barrier between the ambifunctional and O–H···O configuration, and the difference between the ambifunctional and O–H···O configurations as indicated as well as the relative stabilization of *p*-cresol (p) versus methanol (m) as indicated in the legends in the table. All energies are listed in kcal/mol.

| HB configuration | GAFF (kcal/mol) | | | DLPNO-CCSD(T)/CBS (kcal/mol) | | |
|---|---|---|---|---|---|---|
| | *p*-cresol (p) | methanol (m) | diff m-p | *p*-cresol (p) | methanol (m) | diff m-p |
| *syn* N–H···O (1) | -4.2 | -7.1 | -2.9 | -6.1 | -7.1 | -1.0 |
| O–H···O (2) | -9.1 | -8.5 | 0.6 | -11.0 | -8.0 | 3.0 |
| ambi (3) | -10.3 | -10.2 | 0.1 | -12.2 | -10.9 | 1.3 |
| barrier | -8.7 | -7.8 | 0.9 | -8.6 | -6.2 | 2.4 |
| (3) – (2) | -1.2 | -1.7 | -0.5 | -1.2 | -2.9 | -1.7 |

**Text S4.** Details of reaction coordinate construction and transformation.

The (H)O···C=N angle was selected by trial and error to construct a reaction coordinate (ESI Figure S14). This angle was then varied in increments of 0.1° such that methanol or *p*-cresol rotates around acetamide resulting in singly- (single HBs) and doubly-hydrogen-bonded (ambifunctional HB) conformations (Figure S14). Constrained optimizations were carried out on the resulting structures at the MP2/6-31G* level of theory by constraining the (H)O···C-N angle that accounts for simultaneous rotation and translation of methanol or *p*-cresol along with the O···O=C-N dihedral angle to prevent unphysical orientation of the molecules. The geometry optimizations were repeated on most of the structures until converged results were obtained.

Single points were then computed on the converged geometries at the DLPNO-CCSD(T)/CBS level of theory using Tight PNO thresholds. Here, Tight refers to default thresholds of TCutPairs = $10^{-5}$, TCutPNO = 1.00 x $10^{-7}$, TCutMKN = $10^{-3}$.

We then measured the N–H···O angles in all the optimized conformations sampled along the reaction coordinate for both the model systems. The N–H···O angles increase from one conformation to another up to 180°, after which the smaller of [N–H···O angle, 360° - N–H···O angle] is reported. We then obtained potential energy surfaces of the interaction energies of our model systems as a function of the N–H···O angle (ESI Figure S15).

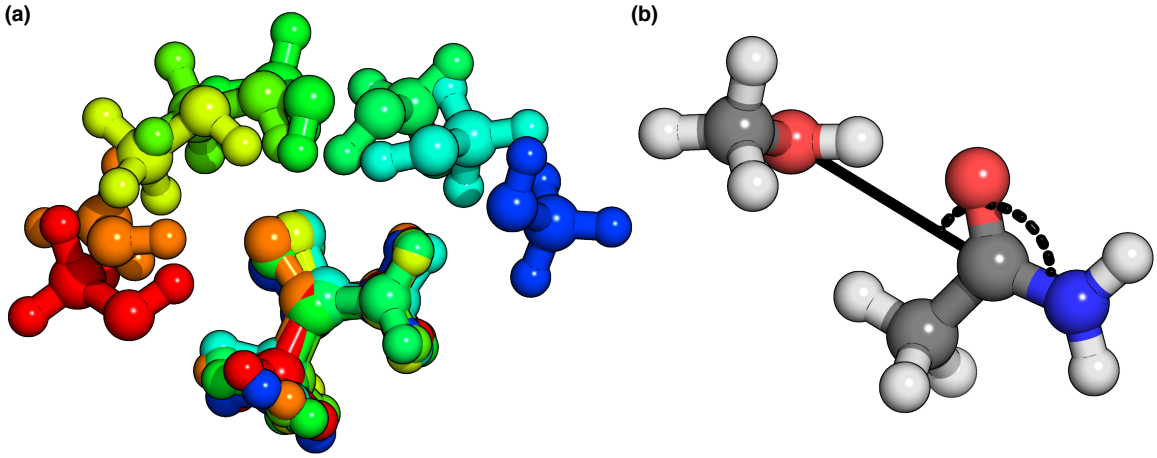


**Figure S14.** (a) Trajectory of the acetamide–methanol system with respect to reaction coordinate, (H)O···C=N intermolecular angle colored by progress along the reaction coordinate (from red to blue). (b) (H)O···C-N reaction coordinate angle on a single representative structure with atoms colored as C in gray, O in red, N in blue, and H in white.

**Table S17.** Frequencies of relevant modes for acetamide-methanol structures along the reaction coordinate. The ambifunctional HB, O-H···O HB, peak of the approximate barrier, and N-H···O HB structures are annotated accordingly. These points correspond to Hessians evaluated at the MP2/6-31G* level of theory on MP2/6-31G*-optimized geometries. The N-H···O angle (in °) is given for these structures (column 1) along with the relevant rotational and vibrational modes (columns 2 and 3, respectively). The negative frequencies reported here correspond to imaginary modes. The rotational frequency reported in column 2 corresponds to a weak rotation where the hydroxyl group (O-H) of methanol moves from N-H of acetamide towards C=O of acetamide or vice-versa. The vibrational frequency reported in column 3 corresponds to the strong O-H stretch of methanol that brings the hydroxyl group of methanol closer to the carbonyl oxygen of acetamide to enable O-H···O HB formation. Up to three imaginary modes were observed in some of the structures which are reported in column 4. These correspond to either weak $-CH_3$ rotations or bending modes of the amide group of acetamide, and do not help to facilitate the transition from one minimum to the other.

| N-H···O angle (°) | Rotation frequency ($cm^{-1}$) | O-H stretching frequency ($cm^{-1}$) | Imaginary modes ($cm^{-1}$) |
|---|---|---|---|
| 167.8 (N-H···O HB) | -66.62 | | |
| 161.0 | -61.80 | 3674.78 | |
| 155.6 | | 3639.55 | -28.19 |
| 151.9 | | 3615.51 | |
| 147.7 | | 3585.49 | |
| 145.2 | | 3570.59 | |
| 141.2 (ambifunctional HB) | | 3564.28 | |
| 135.1 | | 3591.39 | -29.80 |
| 130.0 | -43.50 | 3628.53 | -63.09, -16.84 |
| 127.0 | -64.59 | 3649.71 | -44.94, -8.76 |
| 119.2 | -54.76 | 3693.57 | -17.62 |
| 110.0 | -17.13 | 3725.25 | |
| 92.9 (barrier) | -37.62 | 3765.19 | -22.53, -8.04 |
| 90.0 | -39.47 | 3764.60 | -56.72, -25.36 |
| 80.0 | -10.65 | 3724.80 | -20.21 |
| 70.0 | -52.33 | 3696.67 | -24.45, -13.47 |
| 59.8 (OHO HB) | | 3645.79 | -12.51 |

**Table S18.** Frequencies of relevant modes for acetamide-*p*-cresol structures along the reaction coordinate. The ambifunctional HB, O-H···O HB, peak of the approximate barrier, and N-H···O HB structures are annotated accordingly. These points correspond to Hessians evaluated at the MP2/6-31G* level of theory on MP2/6-31G*-optimized geometries. The N-H···O angle (in °) is given for these structures (column 1) along with the relevant rotational and vibrational modes (columns 2 and 3, respectively). The negative frequencies reported here correspond to imaginary modes. The rotational frequency reported in column 2 corresponds to a weak rotation where the hydroxyl group (O-H) of *p*-cresol moves from N-H of acetamide towards C=O of acetamide or vice-versa. The vibrational frequency reported in column 3 corresponds to the strong O-H stretch of *p*-cresol that brings the hydroxyl group of *p*-cresol closer to the carbonyl oxygen of acetamide to enable O-H···O HB formation. Up to three imaginary modes were observed in some of the structures which are reported in column 4. These correspond to either weak $-CH_3$ rotations or bending modes of the amide group of acetamide, and do not help to facilitate the transition from one minimum to the other.

| N-H···O angle (°) | Rotation frequency ($cm^{-1}$) | O-H stretching frequency ($cm^{-1}$) | Imaginary modes ($cm^{-1}$) |
|---|---|---|---|
| 176.1 (N-H···O HB) | 58.68 | | -40.88, -25.19 |
| 175.7 | -54.45 | 3698.58 | -29.22 |
| 170.3 | -66.02 | 3679.24 | -12.48 |
| 160.1 | -64.22 | 3626.08 | -70.56, -36.72 |
| 150.0 | 69.30 | 3508.48 | -37.51 |
| 144.0 | 138.46 | 3447.85 | -19.12 |
| 137.5 (ambifunctional HB) | 176.41 | 3417.95 | |
| 133.9 | 179.56 | 3431.49 | -7.85 |
| 129.9 | 178.19 | 3469.63 | -16.35 |
| 120.6 | | 3534.41 | |
| 110.0 | -13.81 | 3632.57 | -54.47 |
| 104.0 | | 3652.48 | -37.09, -20.71, -12.31 |
| 90.5 (barrier) | -30.77 | 3673.34 | -14.56 |
| 79.0 | | 3245.02 | -23.73, -20.92 |
| 65.3 | | 3525.75 | -34.11 |
| 57.2 (OHO HB) | | 3497.89 | |

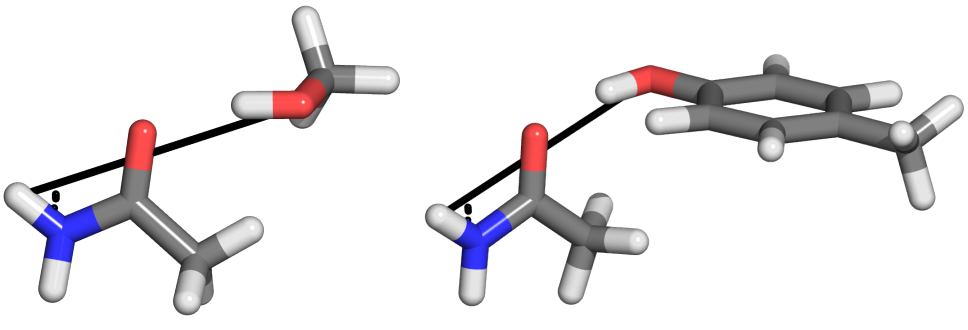


**Figure S15.** The translated reaction coordinate, N–H···O angle, in (a) the acetamide–methanol model system and (b) the acetamide–*p*-cresol model system. Oxygen, nitrogen, carbon, and hydrogen atoms are shown in red, blue, gray, and white, respectively.

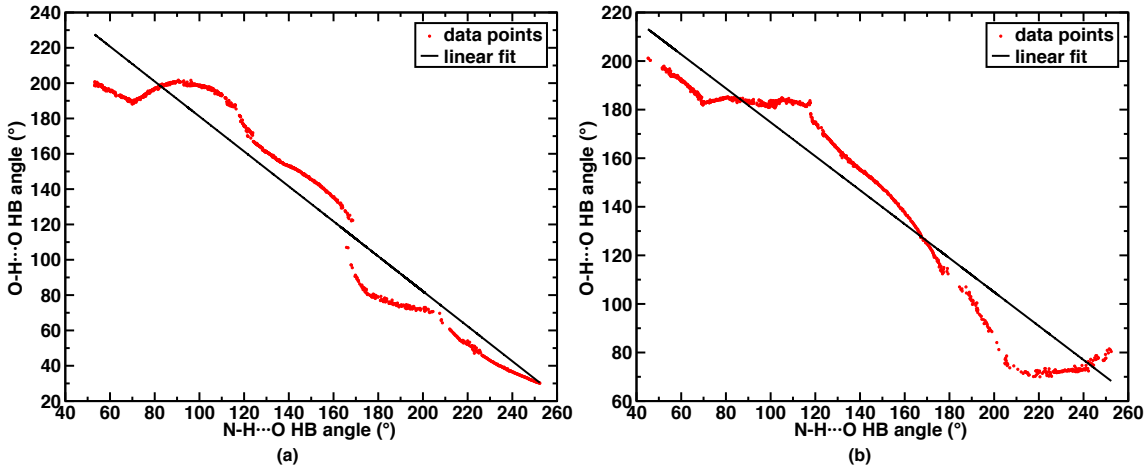


**Figure S16.** O–H···O angle vs N–H···O angle in (a) acetamide–methanol and (b) acetamide–*p*-cresol model systems. The black line in each plot indicates the linear fit through the plot while the red circles are the data points. The slope of the linear fit for acetamide–methanol is -0.99 and for acetamide–*p*-cresol, it is -0.70. Discontinuous data points were pruned in the plots.

**Table S19.** DLPNO-CCSD(T)/CBS energies of O–H···O HB and the transition state corresponding to its transition to ambifunctional HB for acetamide–methanol and acetamide–*p*-cresol model systems. This transition state is qualitative in nature and represents the maximum energy point along this minimum energy pathway obtained through constrained optimizations.

| Model system | O–H···O HB energy (kcal/mol) | Transition state energy (kcal/mol) |
|---|---|---|
| Acetamide–methanol | -7.9 | -6.2 |
| Acetamide–*p*-cresol | -11.0 | -8.7 |

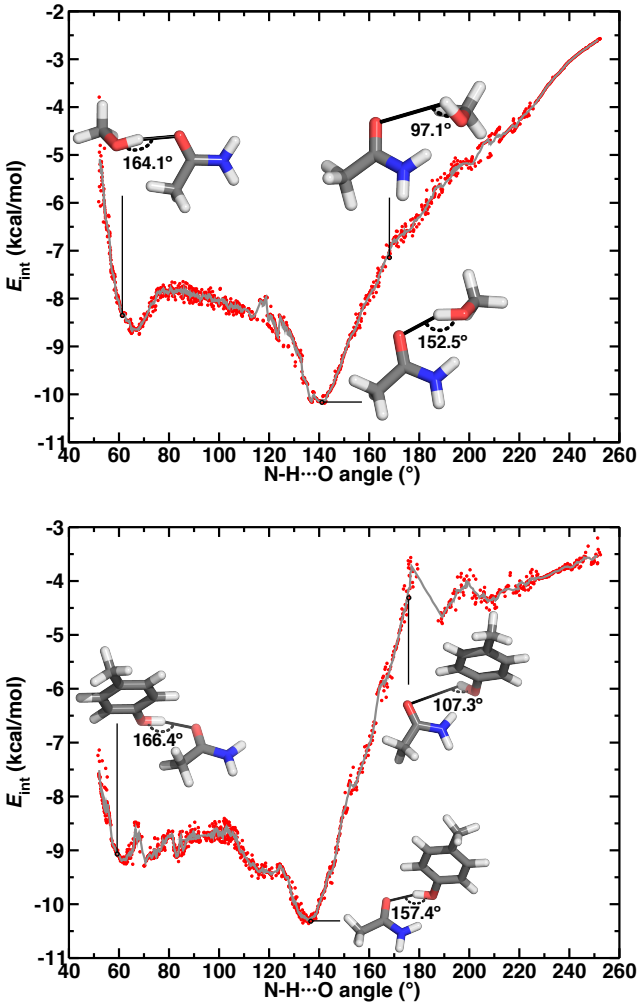


**Figure S17.** Interaction energies ($E_{int}$, in kcal/mol) obtained from the Generalized Amber Force Field (GAFF) of HB conformations shown (red dots) as a function of N–H···O HB angle (in °) and a corresponding 10-point running average (gray line) for (top) acetamide–methanol and (bottom) acetamide–*p*-cresol. Representative structures with measured O–H···O HB angles are shown for the O–H···O HB (top left inset), *syn* N–H···O HB (top right inset), and ambifunctional HB (bottom inset) with the relevant O–H···O HB angle annotated in black. The value of the N–H···O HB angle can be read from the *x*-axis. Discontinuous data points were pruned in the plots. The *p*-cresol points have a greater number of discontinuities due to changes in the unconstrained degrees of freedom.

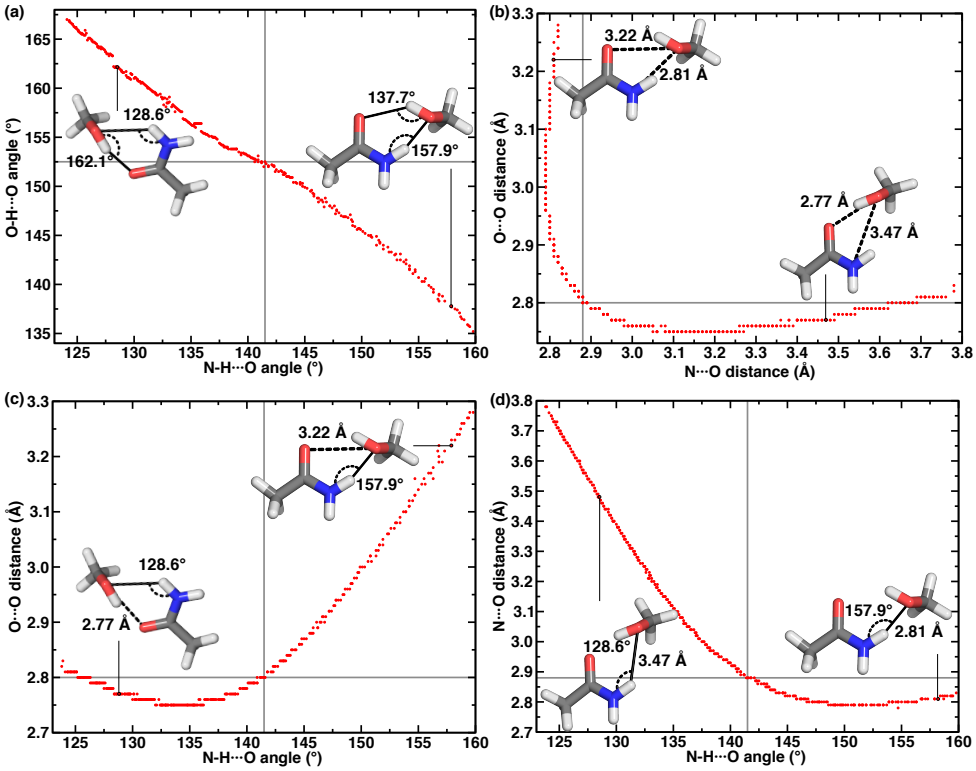


**Figure S18.** Plots of (a) O–H···O angle (in °) vs N–H···O angle (in °), (b) O···O distance (in Å) vs N···O distance (in Å), (c) O···O distance (in Å) vs N–H···O angle (in °) and (d) N···O distance (in Å) vs N–H···O angle (in °) observed in structures in the ambifunctional HB basin of acetamide–methanol model system. (a) The vertical and horizontal gray lines indicate the N–H···O and O–H···O angles observed in the most stable ambifunctional HB, respectively. The data points to the left of the vertical gray line indicate structures from the transition state (connecting O–H···O HB and ambifunctional HB) moving towards the ambifunctional HB configuration while those to the right indicate the structural rearrangements from N–H···O HB to the ambifunctional HB. Two insets are shown, where one exhibits a strong O–H···O HB but a weak N–H···O HB (left) while the other shows the reverse trend (right). (b) The vertical and horizontal gray lines indicate the N···O and O···O distances observed in the most stable ambifunctional HB, respectively. The data points to the left of the vertical gray line indicate the structural rearrangements from N–H···O HB to the ambifunctional HB while those to the right are structures from the transition state (connecting O–H···O HB and ambifunctional HB) moving towards the ambifunctional HB configuration (i.e., the reverse of the orientation on the left plot). Two insets are shown, where one exhibits a strong N–H···O HB but a weak O–H···O HB (left) while the other shows the reverse trend (right). (c) The vertical and horizontal gray lines indicate the N–H···O angle and O···O distance observed in the most stable ambifunctional HB, respectively. (d) The vertical and horizontal gray lines indicate the N–H···O angle and N···O distance observed in the most stable ambifunctional HB, respectively. For both (c) and (d), the data points to the left of the vertical gray line indicate structures from the transition state (connecting O–H···O HB and ambifunctional HB) moving towards the ambifunctional HB configuration while those to the right indicate the structural rearrangements from N–H···O HB to the ambifunctional HB. Two insets are shown, where one exhibits a strong O–H···O HB but a weak N–H···O HB (left) while the other shows the reverse trend (right). Discontinuous data points were pruned in the plots.

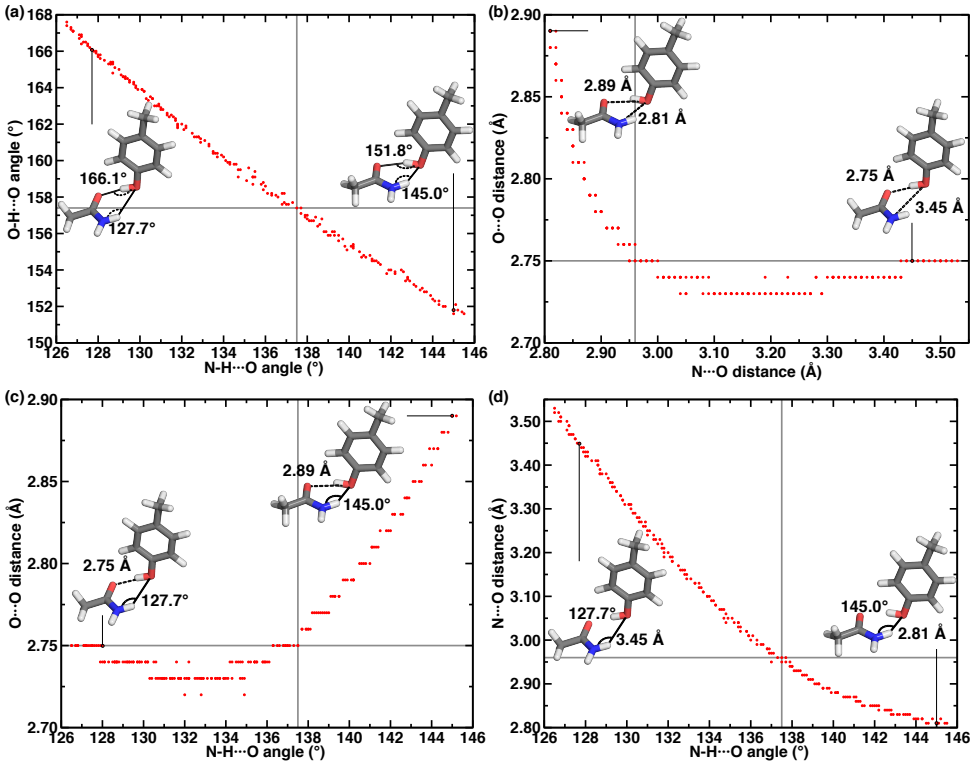


**Figure S19.** Plots of (a) O–H···O angle (in °) vs N–H···O angle (in °), (b) O···O distance (in Å) vs N···O distance (in Å), (c) O···O distance (in Å) vs N–H···O angle (in °) and (d) N···O distance (in Å) vs N–H···O angle (in °) observed in structures in the ambifunctional HB basin of acetamide–*p*-cresol model system. (a) The vertical and horizontal gray lines indicate the N–H···O and O–H···O angles observed in the most stable ambifunctional HB, respectively. The data points to the left of the vertical gray line indicate structures from the transition state (connecting O–H···O HB and ambifunctional HB) moving towards the ambifunctional HB configuration while those to the right indicate the structural rearrangements from N–H···O HB to the ambifunctional HB. Two insets are shown, where one exhibits a strong O–H···O HB but a weak N–H···O HB (left) while the other shows the reverse trend (right). (b) The vertical and horizontal gray lines indicate the N···O and O···O distances observed in the most stable ambifunctional HB, respectively. The data points to the left of the vertical gray line indicate the structural rearrangements from N–H···O HB to the ambifunctional HB while those to the right are structures from the transition state (connecting O–H···O HB and ambifunctional HB) moving towards the ambifunctional HB configuration (i.e., the reverse of the left plot). Two insets are shown, where one exhibits a strong N–H···O HB but a weak O–H···O HB (left) while the other shows the reverse trend (right). (c) The vertical and horizontal gray lines indicate the N–H···O angle and O···O distance observed in the most stable ambifunctional HB, respectively. (d) The vertical and horizontal gray lines indicate the N–H···O angle and N···O distance observed in the most stable ambifunctional HB, respectively. For both (c) and (d), the data points to the left of the vertical gray line indicate structures from the transition state (connecting O–H···O HB and ambifunctional HB) moving towards the ambifunctional HB configuration while those to the right indicate the structural rearrangements from N–H···O HB to the ambifunctional HB. Two insets are shown, where one exhibits a strong O–H···O HB but a weak N–H···O HB (left) while the other shows the reverse trend (right). Discontinuous data points were pruned in the plots.

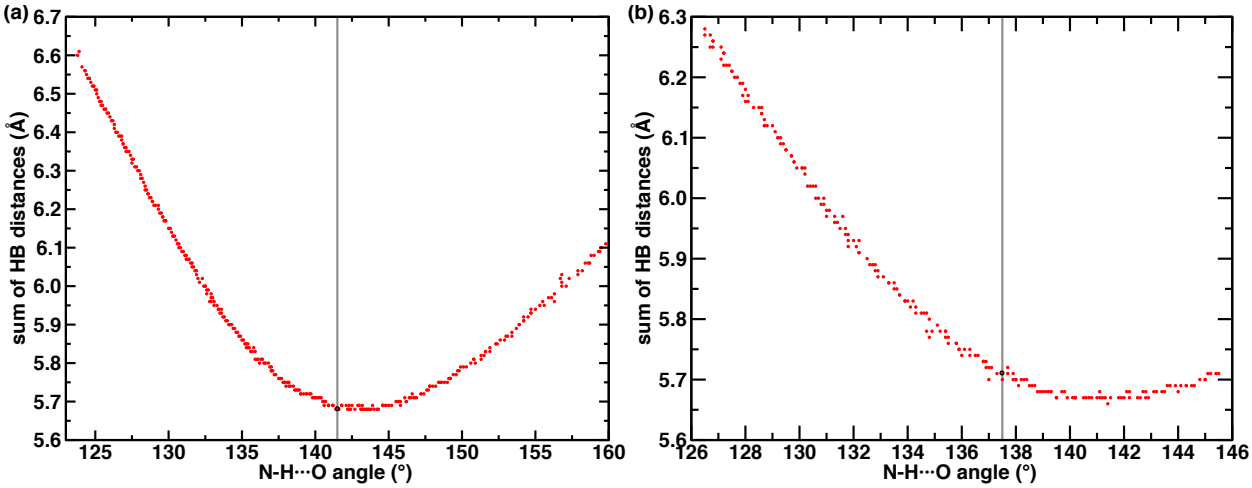


**Figure S20.** Plots showing the sum of HB distances as a function of N–H···O HB angle for the structures in the ambifunctional HB basin for (a) acetamide–methanol and (b) acetamide–*p*-cresol model systems. The data point corresponding to the most stable ambifunctional HB is circled in black. The gray vertical line indicates the N–H···O HB angle in the most stable ambifunctional HB. Discontinuous data points were pruned in the plots.

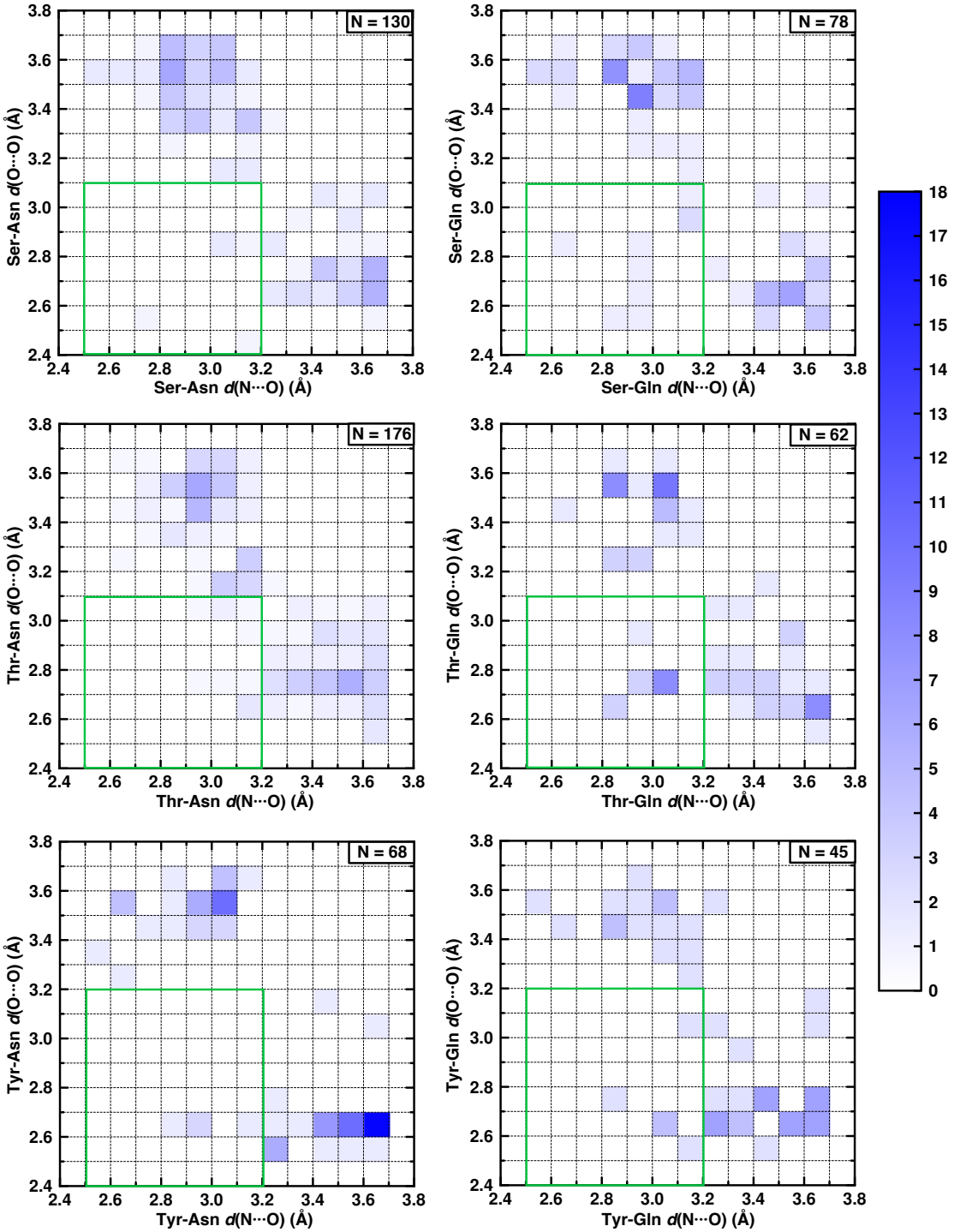


**Figure S21.** Normalized 2D histograms of O···O HB distance ($d$(O···O) in Å) vs. N···O HB distance ($d$(N···O) in Å) of residue pairs in high-resolution crystal structures from our data set that we classify as forming ambifunctional HBs, and the normalized frequency is colored according to the colorbar. The green rectangular box indicates the O···O and N···O HB distance ranges over which the strongest ambifunctional HBs are observed. The residue pairs shown are Ser-Asn (top left), Ser-Gln (top right), Thr-Asn (middle left), Thr-Gln (middle right), Tyr-Asn (bottom left), and Tyr-Gln (bottom right). The bin width along each axis is 0.1 Å.

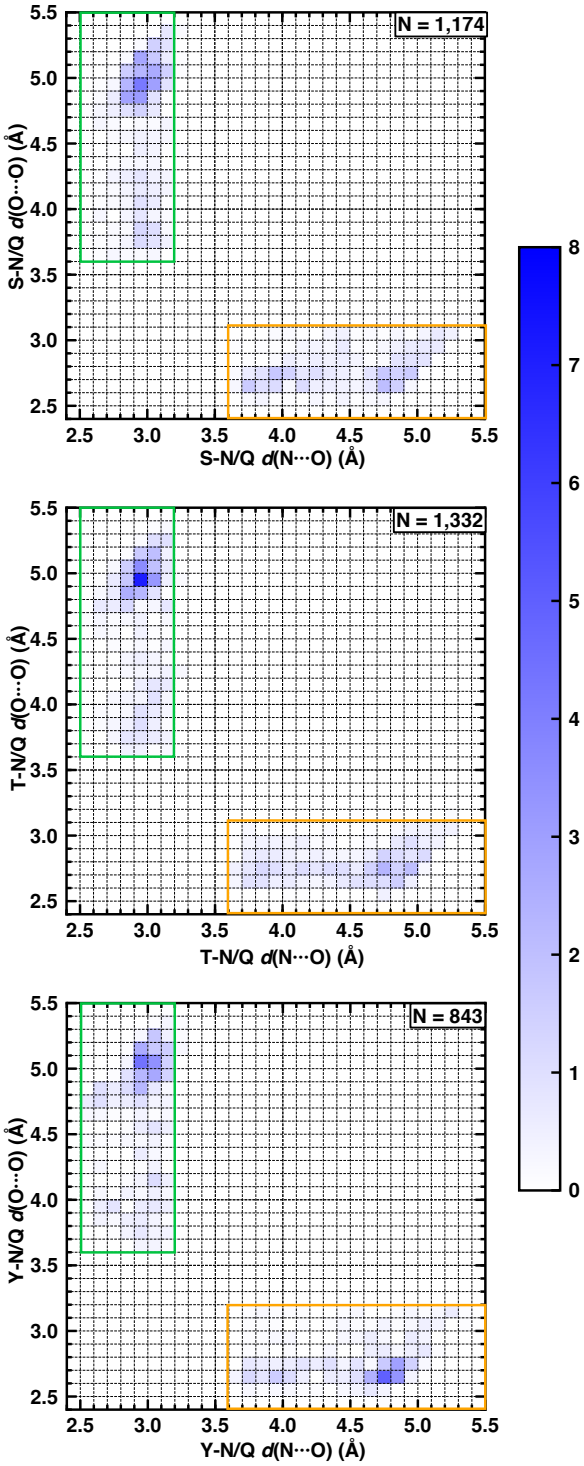


**Figure S22.** Normalized 2D histograms (frequency in colorbar at right) of O···O HB distance ($d$(O···O) in Å) vs N···O HB distance ($d$(N···O) in Å) of all residue pairs in with single HBs in the high-resolution crystal structure data set. The green square indicates the singly hydrogen bonded N–H···O HBs while the orange square indicates the singly hydrogen bonded O–H···O HBs. The residue pairs shown are Ser-Asn/Gln (top), Thr-Asn/Gln (middle), and Tyr-Asn/Gln (bottom). The residue counts of both singly hydrogen bonded N–H···O and O–H···O HBs (N) are shown in the top right corners of the plots. The bin width along each axis is 0.1 Å.

**Table S20.** Number of extremely strong ambifunctional HBs (column 1), moderately strong ambifunctional HBs (column 2), and total number of ambifunctional HBs (column 3 = column 1 + column 2) in Ser-Asn, Ser-Gln, Thr-Asn, Thr-Gln, Tyr-Asn, and Tyr-Gln residue pairs in the protein data set. The distance criteria for extremely strong ambifunctional HBs is: N···O HB distance ranges between 2.5 Å and 3.2 Å for Ser/Thr/Tyr-Asn/Gln systems while O···O HB distance ranges from 2.4–3.1 Å (3.2 Å) for Ser/Thr-Asn/Gln (Tyr-Asn/Gln) systems, respectively. Moderately strong HBs comprise of one of the HB distances within the above-mentioned range while the other HB distance is outside this range.

| Residue pair | Extremely strong | Moderately strong | Total |
|---|---|---|---|
| Ser-Asn | 5 | 125 | 130 |
| Ser-Gln | 9 | 69 | 78 |
| Thr-Asn | 14 | 162 | 176 |
| Thr-Gln | 11 | 51 | 62 |
| Tyr-Asn | 4 | 64 | 68 |
| Tyr-Gln | 6 | 39 | 45 |

**Table S21.** QTAIM analysis for the PDB files showing representative ambifunctional HBs between Ser and Asn, and Tyr and Gln in Figure 8 of the main text. The analysis was performed using B3LYP-D3/6-31G*-geometry-optimized PDB structures with heavy atoms constrained. At most, one of the N···O and O···O HB distances for these PDB structures falls outside the HB distance criteria for single HBs, and hence, one of the HBs in these ambifunctional HBs is weak, as corroborated by the BCP HB energy estimates. The HB distance and angle information are given in Å and °, respectively. The energies of N-H···O HB ($E_{HB}$ N-H···O) and O-H···O HB ($E_{HB}$ O-H···O), which are estimated as half of the potential energy at the BCP obtained from QTAIM topology analysis, are given in kcal/mol.

| PDB ID | Chain Name | Res 1 | Res 2 | *d*(N···O) (Å) | *d*(O···O) (Å) | N-H···O angle (°) | O-H···O angle (°) | $E_{HB}$ N-H···O (kcal/mol) | $E_{HB}$ O-H···O (kcal/mol) |
|---|---|---|---|---|---|---|---|---|---|
| 1sfs | A | N66 | S43 | 3.09 | 3.40 | 153.5 | 130.9 | -4.92 | -1.20 |
| 2vov | A | N330 | S93 | 2.70 | 2.58 | 138.1 | 129.9 | -9.02 | -9.77 |
| 1o4y | A | N107 | S105 | 3.36 | 2.77 | 130.8 | 160.2 | -1.47 | -9.30 |
| 3epw | B | Y150 | Q119 | 2.85 | 3.43 | 162.7 | 132.9 | -9.16 | -1.17 |
| 3bvu | A | Y892 | Q866 | 2.87 | 2.75 | 141.6 | 147.1 | -6.64 | -8.76 |
| 4b9f | B | Q51 | Y43 | 3.22 | 2.68 | 124.5 | 164.2 | -1.79 | -12.20 |

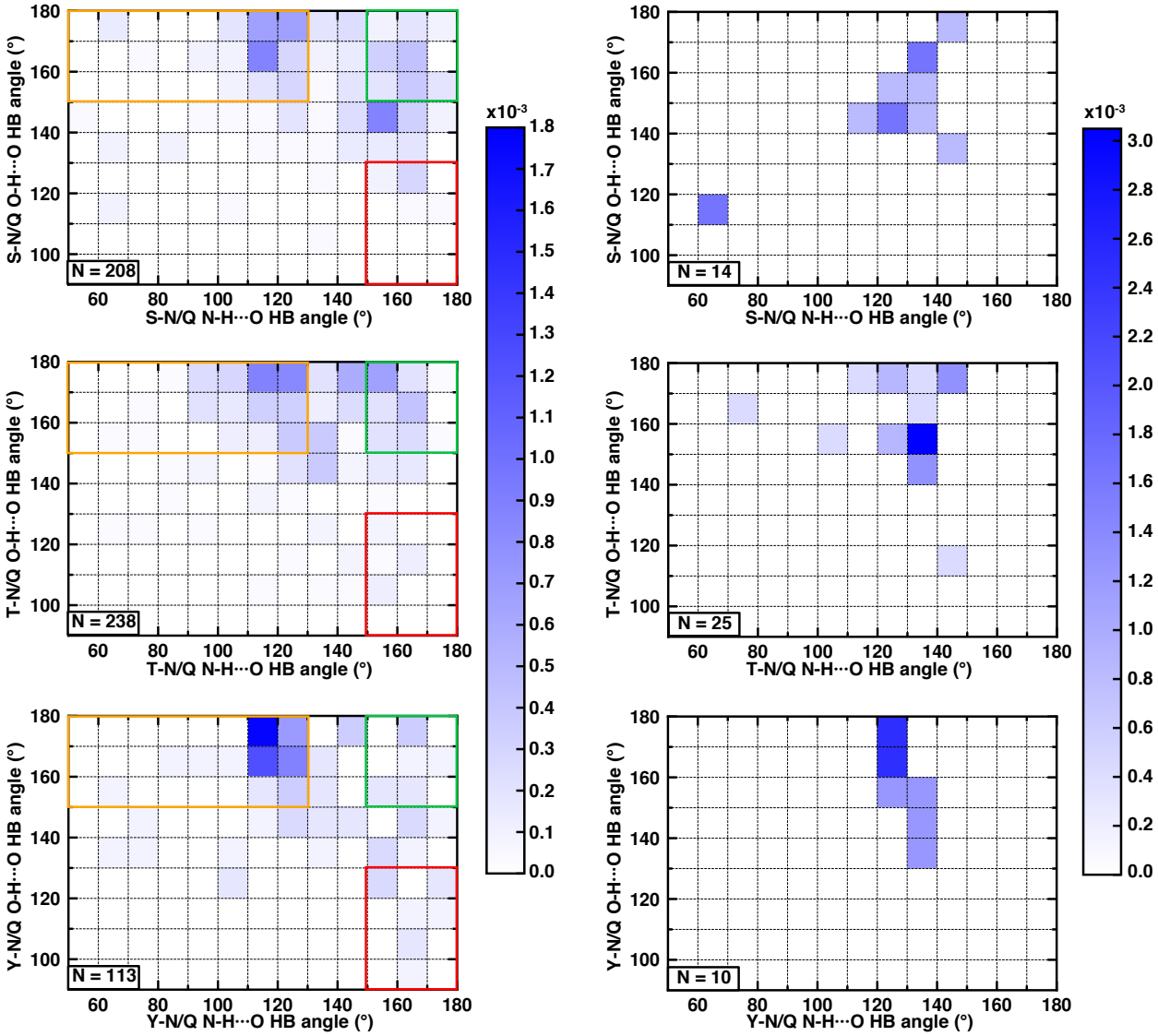


**Figure S23.** (Left) Normalized 2D histograms of O–H···O HB angle (in °) vs N–H···O HB angle (in °) of residue pairs that show ambifunctional HBs in the high-resolution crystal structure data set. The green square indicates the most favorable O–H···O and N–H···O HB angles. The orange rectangle indicates the regions with most favorable O–H···O HB angles but poor N–H···O HB angles while the red rectangle indicates the regions with most favorable N–H···O HB angles but poor O–H···O HB angles. (Right) Normalized 2D histograms of O–H···O HB angle (in °) vs. N–H···O HB angle (in °) of residue pairs that show strong ambifunctional HBs with both moderate O···O and N···O HB distances in the high-resolution crystal structure data set. On both the left and the right, the residue pairs shown are Ser-Asn/Gln (top), Thr-Asn/Gln (middle), and Tyr-Asn/Gln (bottom). The residue counts (N) are shown in the bottom left corners of the plots. The bin width along each axis is 10°, and the frequency is indicated by the colorbar to the right of each graph.

**Table S22.** Mean N–H···O angle and O–H···O angle of all ambifunctional HBs (columns 2 and 3), strong ambifunctional HBs with both moderate HB distances (columns 4 and 5), and moderately strong ambifunctional HBs with one moderate and the other longer HB distance (columns 6 and 7) in Ser-Asn, Ser-Gln, Thr-Asn, Thr-Gln, Tyr-Asn, and Tyr-Gln residue pairs in the protein data set.

| Residue pair | Mean N–H···O HB angle | Mean O–H···O HB angle | Mean N–H···O HB angle (strong ambi.) | Mean O–H···O HB angle (strong ambi.) | Mean N–H···O HB angle (moderately strong ambi.) | Mean O–H···O HB angle (moderately strong ambi.) |
|---|---|---|---|---|---|---|
| Ser-Asn | 135.8° | 157.3° | 126.2° | 146.0° | 136.2° | 157.7° |
| Ser-Gln | 136.4° | 154.4° | 118.2° | 145.4° | 138.8° | 155.5° |
| Thr-Asn | 131.3° | 160.5° | 128.5° | 163.7° | 131.5° | 160.2° |
| Thr-Gln | 135.3° | 159.0° | 130.7° | 152.3° | 136.3° | 160.4° |
| Tyr-Asn | 131.7° | 156.7° | 129.0° | 155.6° | 131.8° | 156.8° |
| Tyr-Gln | 127.1° | 158.8° | 130.0° | 159.6° | 126.6° | 158.7° |

**Table S23.** Interaction energies (Int. E, in kcal/mol) obtained from DLPNO-CCSD(T)/CBS calculations using the MP2/6-31G*-optimized geometries of ambifunctional HB configurations of formamide–methanol, acetamide–methanol, propanamide–methanol, and acetamide–ethanol. The interaction energy is obtained as the difference in energy of the dimer from the isolated molecules, all evaluated on MP2/6-31G*-optimized geometries. The two-point extrapolation formula based on the aug-cc-pVDZ and aug-cc-pVTZ energies is used to extrapolate to the complete basis set (CBS) limit following Refs.[6-8] for DLPNO-CCSD(T). Here, Tight refers to default thresholds of TCutPairs = $10^{-5}$, TCutPNO = 1.00 x $10^{-7}$, TCutMKN = $10^{-3}$. The interaction energies of models where Asn/Gln is modeled by formamide, acetamide, and propanamide differ by no more than 0.34 kcal/mol. The interaction energies of models where Ser/Thr is modeled by methanol and ethanol differ by no more than 0.18 kcal/mol. While formamide is computationally cheaper to model than acetamide and propanamide, the latter would be more accurate for modeling Asn or Gln. Propanamide is computationally expensive and gives fairly comparable results to acetamide (within 0.06 kcal/mol), and hence, acetamide has been used to model both Asn and Gln.

| Representative Models | Int. E (kcal/mol) |
|---|---|
| Formamide–methanol ambifunctional HB | -10.30 |
| Acetamide–methanol ambifunctional HB | -10.64 |
| Propanamide–methanol ambifunctional HB | -10.58 |
| Acetamide–ethanol ambifunctional HB | -10.82 |

**Table S24.** Key distances (in Å) and angles (in °) in MP2-optimized geometries of an acetamide–methanol ambifunctional HB configuration using different basis sets.

| Basis | O–H···O=C (Å) | H–O···H–N (Å) | O···O (Å) | N···O (Å) | ∠OH···O (°) | ∠NH···O (°) |
|---|---|---|---|---|---|---|
| 6-31G* | 1.90 | 2.01 | 2.81 | 2.88 | 152.6 | 141.8 |
| 6-31++G** | 1.90 | 2.02 | 2.80 | 2.88 | 150.5 | 140.5 |
| cc-pVDZ | 1.85 | 1.98 | 2.77 | 2.85 | 153.9 | 140.7 |
| cc-pVTZ | 1.83 | 1.99 | 2.75 | 2.85 | 155.7 | 140.1 |
| aug-cc-pVDZ | 1.87 | 2.01 | 2.78 | 2.88 | 154.1 | 140.5 |
| aug-cc-pVTZ | 1.84 | 2.01 | 2.76 | 2.86 | 154.7 | 140.1 |

**Table S25.** Comparison of interaction energies (Int. E, in kcal/mol) obtained using MP2/6-31G*-optimized geometry or B3LYP-D3/6-31G*-optimized geometries of the acetamide-methanol ambifunctional HB configuration evaluated at several levels of theory. The CBS limit was extrapolated based on a two-point formula. Differences in structures lead to differences in single point energy-evaluated interaction energies of no more than 0.1 kcal/mol for each level of theory.

| Method/Basis | B3LYP-D3 struct. Int. E (kcal/mol) | MP2 struct. Int. E (kcal/mol) | Col. 1-2 (kcal/mol) |
|---|---|---|---|
| DLPNO-CCSD(T) / aug-cc-pVDZ (Normal) | -10.6 | -10.7 | 0.1 |
| DLPNO-CCSD(T) / aug-cc-pVDZ (Tight) | -10.9 | -10.9 | 0.0 |
| CCSD(T) / aug-cc-pVDZ | -11.1 | -11.2 | 0.1 |
| DLPNO-CCSD(T) / aug-cc-pVTZ (Normal) | -10.5 | -10.5 | 0.0 |
| DLPNO-CCSD(T) / aug-cc-pVTZ (Tight) | -10.7 | -10.7 | 0.0 |
| CCSD(T) / aug-cc-pVTZ | -10.9 | -10.9 | 0.0 |
| DLPNO-CCSD(T)/CBS (Tight) | -10.7 | -10.6 | -0.1 |
| CCSD(T)/CBS | -10.9 | -10.9 | 0.0 |

**Table S26.** Comparison of interaction energies (Int. E, in kcal/mol) obtained using MP2/6-31G*-optimized geometries for all four acetamide–methanol and acetamide–*p*-cresol HB configurations evaluated at B3LYP-D3/aug-cc-pVTZ and DLPNO-CCSD(T)/aug-cc-pVTZ (Tight) levels of theory. Differences in single point energy-evaluated interaction energies of both the levels of theory are no more than 0.8 kcal/mol for each configuration. Here, Tight refers to default thresholds of TCutPairs = $10^{-5}$, TCutPNO = 1.00 x $10^{-7}$, TCutMKN = $10^{-3}$.

| HB configuration | B3LYP-D3/aug-cc-pVTZ Int. E (kcal/mol) | DLPNO-CCSD(T)/ aug-cc-pVTZ (Tight) Int. E (kcal/mol) | Col. 1-2 (kcal/mol) |
|---|---|---|---|
| acetamide–methanol | | | |
| *syn* N–H···O | -6.6 | -7.1 | 0.5 |
| *anti* N–H···O | -6.0 | -6.6 | 0.6 |
| O–H···O | -7.8 | -8.0 | 0.2 |
| ambifunctional | -10.5 | -10.7 | 0.2 |
| acetamide–*p*-cresol | | | |
| *syn* N–H···O | -5.8 | -6.4 | 0.6 |
| *anti* N–H···O | -5.7 | -6.5 | 0.8 |
| O–H···O | -11.0 | -11.4 | 0.4 |
| ambifunctional | -12.0 | -12.5 | 0.5 |

**References**

(1) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminf.* **2012,** *4*, 17.

(2) Halgren, T. A. Merck Molecular Force Field .3. Molecular Geometries and Vibrational Frequencies for MMFF94. *J. Comput. Chem.* **1996,** *17*, 553-586.

(3) Nikolaienko, T. Y.; Bulavin, L. A.; Hovorun, D. M. Bridging QTAIM with Vibrational Spectroscopy: The Energy of Intramolecular Hydrogen Bonds in DNA-Related Biomolecules. *Phys. Chem. Chem. Phys.* **2012,** *14*, 7441-7447.

(4) Bader, R. F. W. A Quantum-Theory of Molecular-Structure and Its Applications. *Chem. Rev.* **1991,** *91*, 893-928.

(5) Bader, R. F. Atoms in Molecules. *Acc. Chem. Res.* **1985,** *18*, 9-15.

(6) Zhong, S. J.; Barnes, E. C.; Petersson, G. A. Uniformly Convergent N-Tuple-Zeta Augmented Polarized (Nzap) Basis Sets for Complete Basis Set Extrapolations. I. Self-Consistent Field Energies. *J. Chem. Phys.* **2008,** *129*, 184116.

(7) Neese, F.; Valeev, E. F. Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated Ab Initio Methods? *J. Chem. Theory Comput.* **2011,** *7*, 33-43.

(8) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997,** *106*, 9639-9646.

(9) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010,** *132*, 154104.

(10) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988,** *37*, 785-789.

(11) Becke, A. D. Density-Functional Thermochemistry .3. The Role of Exact Exchange. *J. Chem. Phys.* **1993,** *98*, 5648-5652.

(12) 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. *J. Phys. Chem.* **1994,** *98*, 11623-11627.

(13) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004,** *25*, 1157-1174.