SUPPORTING INFORMATION

Determination of Hydrogen Bond Structure in Water Versus Aprotic Environments to Test the Relationship Between Length and Stability

Paul A. Sigala[†,Δ], Eliza A. Ruben[†], Corey W. Liu[‡], Paula M. B. Piccoli[§], Edward G. Hohenstein[#], Todd J. Martínez[#], Arthur J. Schultz[§,≠], and Daniel Herschlag[†,#,*]

[†] Dept. of Biochemistry, Stanford University, Stanford, California 94305
[‡] Stanford Magnetic Resonance Laboratory, Stanford University, Stanford, California 94305
[§] Intense Pulsed Neutron Source, Argonne National Laboratory, Argonne, Illinois 60439
[#] Dept. of Chemistry, Stanford University, Stanford, California 94305
[Δ] Present address: Dept. of Molecular Microbiology, Washington University, St. Louis, Missouri 63110
[≠] Present address: X-ray Science Division, Argonne National Laboratory, Argonne, Illinois 60439
[*] e-mail: herschla@stanford.edu

MATERIALS AND METHODS

Materials. All reagents were of the highest purity commercially available. Substituted salicylates, substituted phenols, substituted pyridine-N-oxides, triethylamine, 1,8,9-trihydroxyanthracene, and N,N,N',N'-tetramethyl-1,8-naphthalenediamine (proton-sponge®) were purchased from Sigma-Aldrich. All solutions were prepared with reagent-grade solvents or better.

Structure determination by single crystal X-ray and neutron diffraction. *Crystallization.* Single crystals of 1:1 complexes of 2,4-dinitrophenol•4-(dimethyl)aminopyridine-N-oxide; 2,4-dinitrophenol•4-methoxypyridine-N-oxide; 2,4-dinitrophenol•4-methylpryridine-N-oxide, 4-nitrophenol•4-methoxypyridine-N-oxide; 4-nitrophenol•pyridine-N-oxide; and 3,5-dinitrosalicylate•proton-sponge were prepared by slow evaporation (7-14 days) of equimolar solutions in acetone or acetonitrile. Crystals were screened for quality and hydrogen bond formation by polarized light microscopy and IR spectroscopy (KBr pellets).

X-ray crystallography. X-ray structures were determined for 2,4-dinitrophenol•4-(dimethyl)aminopyridine-N-oxide; 2,4-dinitrophenol•4-methylpryridine-N-oxide; 4-nitrophenol•4-methoxypyridine-N-oxide; and 3,5-dinitrosalicylate•proton-sponge. Single crystals were mounted on a glass fiber and optically aligned on a Bruker APEX-II CCD X-ray diffractometer using APEX-II (v 1.0-22) software.[1] A sphere of three-dimensional data was collected using graphite monochromated Mo Kα radiation from a sealed tube and monocapillary collimator. Data collection included four unique series of exposures of 600 frames with an exposure time of 30s per frame and frame widths of 0.3˚ in ω.

The determination of corrected integrated intensities and global refinement of lattice parameters were performed using Saint+ (v 7.09) with a narrow-frame integration algorithm.[2] A semi-empirical absorption correction was subsequently applied with the program SADABS.[3] SHELXTL (v 6.14) software was used for space group determination (XPREP), direct methods structure solution (XS), and least-squares refinement (XL).[4] Structures were solved via direct methods and refined by full-matrix least squares on the basis of F^2. The final refinement included anisotropic displacement parameters for all non-hydrogen atoms. Hydrogen atoms were located in the difference maps and were included in the refinement. X-ray crystallographic data collection and refinement details are given in Table S4. Crystallographic Information Files (CIF) for all X-ray structures reported herein have been deposited in the Cambridge Structural Database (accession codes are given in Table S4).

Neutron Crystallography. Neutron diffraction data for each of the six co-crystals listed above were obtained at the Intense Pulsed Neutron Source (IPNS) at Argonne National Laboratory using the time-of-flight Laue single-crystal diffractometer (SCD),[5] as previously described.[6] Briefly, each crystal was wrapped in aluminum foil and glued to an aluminum pin that was mounted on the cold stage of a closed-cycle helium refrigerator. The crystal was then cooled to 30 ± 1 K. For each setting of the diffractometer angles, data were stored in three-dimensional histogram form with coordinates *x,y,t* corresponding to horizontal and vertical detector positions and the time-of-flight, respectively. An auto-indexing algorithm[7] was used to obtain an initial orientation matrix from the peaks in three preliminary histograms measured for approximately 30 minutes each. For intensity data collection, runs of 3–6 hours per histogram were initiated for the data set. Settings were arranged at χ and φ values suitable to cover at least one unique portion of reciprocal space. Bragg peaks in the recorded histograms were indexed

and integrated using individual orientation matrices for each histogram, to allow for any misalignment of the sample. Intensities were integrated about their predicted locations and were corrected for the Lorentz factor, the incident spectrum, and the detector efficiency. A wavelength-dependent spherical absorption correction was applied using cross-sections for the non-hydrogen atoms[8] and the hydrogen atoms[9] (see Table S5 for sample-specific absorption corrections). Symmetry related reflections were not averaged since different extinction factors are applicable to reflections measured at different wavelengths.

The GSAS software package was used for structural analysis.[10] The atomic positions of X-ray diffraction structures were used as a starting point in the refinement of the neutron diffraction data. X-ray structures were either experimentally determined for the complexes described above or were available from the Cambridge Structural Database (accession codes: NILZOX [pyridine-N-oxide•4-nitrophenol] and NELTIH [4-methoxypyridine-N-oxide•2,4-dinitrophenol]). The refinement was based on F^2 reflections with a minimum *d*-spacing of 0.5 or 0.7 Å. Weights were assigned as $w(F_o^2) = 1 / [(\sigma(F_o^2) + (0.002 * F_o^2)]^2$ where $\sigma^2(F_o^2)$ is the variance based on counting statistics. In the final refinement all atoms, including hydrogen atoms, were refined with anisotropic displacement parameters. For the complex containing 3,5-dinitrosalicylate and proton-sponge, a low data-to-parameter ratio prevented anisotropic refinement of the atoms, and an isotropic refinement was used instead to obtain the best model. Structures were viewed and analyzed using ORTEP-III[11] and Mercury[12], and the refined hydrogen bond distances are given in Table S1. Data collection and refinement statistics for each structure are given in Table S5. CIF files for all neutron structures reported herein have been deposited in the Cambridge Structural Database (accession codes are given in Table S5).

Refinement of the X-ray and neutron diffraction data for 3,5-dinitrosalicylate (ΔpK_a of 0.5) resulted in final structural models (Table S4 and S5) in which the proton of the intramolecular hydrogen bond was closer to the oxygen atom of the carboxylic acid group than that of the hydroxyl group. This configuration, which differs from the more common observation in most substituted salicylate structures (with ΔpK_a values >1) that the bridging proton is closer to the oxygen atom of the hydroxyl group than the carboxylate group,[13] presumably reflects the very similar proton affinities of the carboxylic acid and hydroxyl groups in 3,5-dinitrosalicylate (see Table S1) which are expected to result in a double-minimum potential energy surface with a low barrier to proton transfer between the hydrogen-bonded oxygens. Our structures are consistent with a prior X-ray diffraction study that reported multiple, independently refined structures of 3,5-dinitrosalicylate and observed a mixture of proton positions (closer either to the oxygen atom of the carboxylic acid or hydroxyl group) that varied idiosyncratically with the particular counter-ion present in each co-crystal.[14]

Analysis of low-temperature neutron structures of O-H•••O hydrogen-bonded complexes. We analyzed the small molecule structures of O-H•••O hydrogen-bonded complexes determined by low temperature (<130 K) single crystal neutron diffraction that were compiled by Steiner and Saenger in their highly cited 1994 study.[15] Coordinates for all available structures (27 out of 32) were retrieved from the Cambridge Structural Database[16], version 5.34 (November 2012), and analyzed using the Conquest software for O-H•••O hydrogen bonds with O•••O distances <3.0 Å and H•••O distances ≤2.0 Å. Structures containing polymers (BUVSEQ03 and SIGHOF) or donor groups coordinated to metal ions (SOGGEA and KIDREU01) were excluded from the analysis, resulting in a chemically diverse data-set of 23 structures and 62 hydrogen bonds.

Accession codes for these 23 structures and the O•••O, O-H, and H•••O distances and O-H•••O angles for the 62 O-H•••O hydrogen bonds are given in Table S1. Aqueous pK_a values were assigned to each donor and acceptor oxygen using the sources indicated in Table S1. These values reflect experimental pK_a values[17,18] determined for the identical molecule where available, experimental pK_a values determined for a chemically similar compound or group, or computational estimates using the ACD/Labs I-Lab 2.0 pK_a Prediction Module.

Calculation of ΔpK_a values for substituted salicylates. The pK_a and ΔpK_a values used in this study are a proxy for the relative proton affinities of the hydrogen-bonded donor and acceptor groups. The "intrinsic" pK_a values of salicylate OH and COOH groups represent the pK_a value of each group expected from the inductive and resonance effects of ring substituents in the absence of hydrogen bonding to each other, and ΔpK_a represents the difference in intrinsic pK_a values of the hydrogen bond donor and acceptor groups.[19] We therefore calculated intrinsic pK_a values for both the COOH and OH groups in order to obtain the ΔpK_a value for the intra-molecular hydrogen bond in each substituted salicylate monoanion. Any errors in calculated ΔpK_a values do not affect the conclusions of the paper, as the chemical shifts of the same compounds in multiple solvents are directly compared.

Intrinsic pK_a values for each group were estimated using the pK_a of that group in an analogous compound with the OH group *para* to the COOH group instead of *ortho* to avoid any convoluted effects from hydrogen bonding between the groups. Any error introduced by this approximation has a constant effect on the series of compounds studied and therefore does not affect the slope of the plots herein. For unsubstituted salicylic acid, the intrinsic pK_a values of the COOH and OH groups were estimated from the pK_a values of 4-hydroxybenzoic acid (4.6) and

4-hydroxybenzoate (9.4), respectively.[17] For substituted salicylates, substituent effects were factored in by first calculating the pK_a difference between benzoic acid (4.2) and the corresponding substituted benzoic acid and the pK_a difference between phenol (10.0) and the corresponding substituted phenol, using literature pK_a values.[17] These differences were then subtracted from the pK_a values of 4.6 (COOH) and 9.4 (OH) calculated above to give the intrinsic pK_a value for the COOH and OH group of each substituted salicylate, as summarized below:

$$pK_a^{COOH,\,int} = \underset{4.6}{HO\text{-}C_6H_4\text{-}COOH} - \left(\underset{4.2}{C_6H_5\text{-}COOH} - \underset{X}{X\text{-}C_6H_4\text{-}COOH} \right)$$

$$pK_a^{OH,\,int} = \underset{9.4}{HO\text{-}C_6H_4\text{-}COO^-} - \left(\underset{10.0}{C_6H_5\text{-}OH} - \underset{X}{X\text{-}C_6H_4\text{-}OH} \right)$$

$$\Delta pK_a = pK_a^{OH,\,int} - pK_a^{COOH,\,int}$$

NMR spectroscopy. ^1H spectra of substituted salicylates and the 1,8,9-trihydroxyanthracene monoanion were acquired at the Stanford Magnetic Resonance Laboratory on an 800 MHz (^1H frequency) Varian ^{UNITY}INOVA spectrometer running VNMR v6.1C and equipped with a 5-mm, triple-resonance, gradient $^1H(^{13}C/^{15}N)$ probe. Probe temperatures were calibrated using a sealed 100% methanol low-temperature calibration standard sample. NMR samples consisted of 25 mM substituted salicylic acid and triethylamine (TEA) in 100% water (containing 5% D_2O), acetone-d_6 (with 0.05% tetramethylsilane, TMS), or $CDCl_3$ (with 0.05% TMS). The aqueous sample of 3,5-dinitrosalicylate was dissolved in 10% DMSO-d_6/water. Samples of 1,8,9-

trihydroxyanthracene contained 10 mM of this compound and TEA in DMSO-d_6 (with 0.05% TMS) or 10% DMSO-d_6/water (the compound was insoluble in 100% water), conditions which lead to formation of the 1,8,9-trihydroxyanthracene monoanion.[20] Spectra of the HD isotopologues were acquired using samples dissolved in 10% DMSO-d_6/54% D_2O/36% H_2O (0.6 mole fraction of exchangeable deuterons) or samples that had first been dissolved in 50% MeOH/MeOD (0.5 mole fraction of exchangeable deuterons), dried by vacuum centrifugation, and then dissolved in DMSO-d_6.

^{1}H NMR spectra were acquired at 4 ˚C, except for the samples of 3,5-dinitrosalicylate in 10% DMSO-d_6/water and 1,8,9-trihydroxyanthracene in DMSO-d_6, which were acquired at -3 ˚C and 20 ˚C, respectively. Spectra for samples dissolved in acetone-d_6, chloroform-d_6, and DMSO-d_6 were acquired with 100-200 scans using the s2pul pulse sequence, and chemical shifts were referenced internally to TMS (0 ppm). Spectra for aqueous samples were acquired with 100-200 scans using the 1331 binomial pulse sequence[21] to suppress the water signal, a spectral width of 30 ppm (carrier frequency set on the water resonance), an excitation maximum of 14–17 ppm, and a baseline correction applied over the peak of interest. Chemical shifts in water were referenced internally to the water resonance (5.0 ppm at 4 ˚C) and externally to a sample of sodium 3-trimethylsilylpropionate-2,2,3,3-d_4 (0 ppm). All spectra were processed using a 10 Hz line broadening.

^{2}H NMR spectra of substituted salicylates were acquired at 0 ˚C at the Stanford Magnetic Resonance Laboratory on a 600 MHz (^{1}H frequency) Varian ^{UNITY}INOVA spectrometer running VNMR v6.1C and equipped with a 10 mm broadband probe. Samples contained 25 mM of each substituted salicylate and TEA in 95% D_2O/5% DMSO, acetone (containing 1% acetone-d_6 for referencing), or chloroform (containing 1% $CDCl_3$ for referencing). Acetone and chloroform

samples were prepared by first dissolving samples in MeOD to exchange labile protons for deuterons, drying them in a vacuum centrifuge, and dissolving the dried sample in the respective organic solvent. Spectra for acetone and chloroform samples were acquired with 1000 scans using the s2pul pulse sequence. Spectra for aqueous samples were acquired with 1000-4000 scans using the 1331 binomial pulse sequence[21] to suppress the water signal, a spectral width of 30 ppm (carrier frequency set on the water resonance), an excitation maximum of 13–17 ppm, and a baseline correction applied over the peak of interest. Chemical shifts were referenced to the solvent peak in each spectrum (water = 4.8 ppm, acetone = 2.1 ppm, chloroform = 7.3 ppm). All spectra were processed using a 10 Hz line broadening function.

Estimation of hydrogen bond distances from ^{1}H NMR chemical shifts. Hydrogen bond distances were estimated from the measured ^{1}H NMR chemical shifts for the detected salicylate hydrogen-bonded protons using references[22-25] and the following correlation functions, which are based on empirical analysis of solid state ^{1}H NMR spectra and crystallographic studies of diverse small molecule O-H•••O hydrogen bonds:

1) O•••O distance (Å) = 5.040 Å – 1.160*ln(chemical shift) + 0.044*(chemical shift)

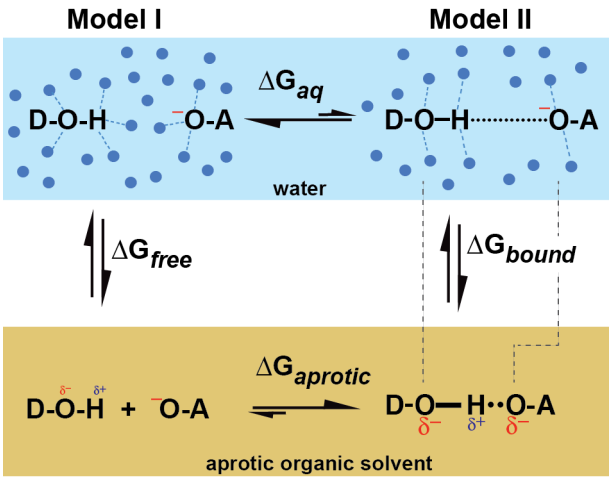
2) H•••O distance (Å) = 2.380 Å – 0.055*(chemical shift)

Quantum mechanical calculations. *Ab initio* electronic structure calculations for the water-hydroxide and formic acid-formate complexes and 5-methylsalicylate were performed with the Hartree-Fock (HF); Becke, three-parameter, Lee-Yang-Parr (B3LYP); or Møller-Plesset (MP2) methods and the 6-31G(d), 6-31G**, 6-31+G(d), or 6-311++G(d,p) basis sets using the Gaussian software package.[26] Solvent dielectric was varied using the polarizable continuum model

(PCM)[27], with the set of atomic radii from Pauling. Calculations were carried out at continuum dielectric values of 4.9 (chloroform), 20.7 (acetone), 46.7 (DMSO), and 78.4 (water). Calculations of the water-hydroxide ion complex with two solvating water molecules (Fig. 5B) were performed at the B3LYP/6-311++G(d,p) level using a continuum dielectric value of 4.9 (chloroform) by adding two water molecules that were within hydrogen bonding distance of the donor and acceptor oxygen atoms, respectively, of the energy-minimized water-hydroxide ion structure in chloroform, followed by an additional round of energy minimization. For calculation of one-dimensional potential energy surfaces for proton transfer between the hydrogen-bonded oxygen atoms of the water-hydroxide dimer at the B3LYP level and the 6-311++G(d,p) basis set, the O-O distance was fixed as indicated in Fig. S7, the proton was displaced between the oxygen atoms in the given increment, and the energy of the system was calculated relative to the minimal-energy proton displacement. Energy minimization and geometry optimization of the isolated 1,8,9-trihydroxyanthracene monoanion were performed with the B3LYP density functional method and the 6-31+G* basis set.

Text S1. Thermodynamic cycle of potential contributions of Models I and II to overall hydrogen bond formation energies (ΔG_f) in water versus aprotic solvents. The potential contributions of Models I and II to the different hydrogen bond ΔG_f values in water and aprotic organic solvents can be represented as a closed thermodynamic cycle (see below). The pertinent question for this study is why ΔG_{aq} differs from $\Delta G_{aprotic}$. Model I focuses on (potential) differences between the free species (vertical equilibrium on the left, ΔG_{free}), whereas Model II focuses on (potential) differences between the complexed species (vertical equilibrium on the right, ΔG_{bound}). The difference between these two ΔG values, $\Delta G_{free} - \Delta G_{bound}$, is identical to the

difference in the hydrogen bond equilibria, ΔG_{aq} - $\Delta G_{aprotic}$. Thus, any differences in the nature of the hydrogen-bonded species (Model II, ΔG_{bound}) would impact ΔG_{aq} - $\Delta G_{aprotic}$, since ΔG_{aq} - $\Delta G_{aprotic}$ = ΔG_{free} - ΔG_{bound}.



Text S2. Literature references and quotes that provide a basis for Model 2.

1) Kreevoy MM and Liang TM. *J. Am. Chem. Soc.* (1980) 102: 3315-3322. **"It may be that water, by interacting with the acceptor groups, weakens and lengthens the AHA-hydrogen bonds, and possibly lengthens the A-A distance."**

2) Gerlt JA, Kreevoy MM, Cleland WW, and Frey PA. *Chem. Biol.* (1997) 4: 259-267. **"Short is strong….the idea that hydrogen bonds become stronger as the donor-acceptor distance becomes shorter is generally accepted and leads to a rationalization of a large number of bond lengths."**

"short, strong hydrogen bonds probably do not occur in dilute aqueous solution."

3) Gao J, Bosco DA, Powers ET, and Kelly JW. *Nat. Str. and Mol. Biol.* (2009) 16: 684-690: **"Thus, hydrogen bond lengths seem to decrease as the environment becomes more nonpolar, suggesting that hydrogen bonds are stronger in nonpolar environments."**

4) Bowie JU. *Curr. Opin. Str. Biol.* (2011) 21: 42-49: **"backbone hydrogen bond lengths vary significantly as a function of the environment polarity."**

5) Cleland WW and Kreevoy MM. *Science* (1994) 264: 1887-1890: **"The bond strength seems well correlated with the distance between the heteroatoms, with the shortest bonds being the strongest."**

6) Hibbert F and Emsley J. *Adv. Phys. Org. Chem.* (1990) 26: 255-379: **"It is taken for granted that as a hydrogen bond becomes shorter it becomes stronger."**

7) Lin IJ, Gebel EB, Machonkin TE, Westler WM and Markley JL. *Proc. Natl. Acad. Sci.* (2005) 102: 14581-14586: **"Shorter H-bonds are stronger than longer H-bonds…. for small changes in the H-bond length, the energy changes in a roughly linear fashion."**

8) Warshel A and Papazyan A. *Proc. Natl. Acad. Sci.* (1996) 93: 13665-13670: **"calculations confirm that the short [hydrogen] bond in vacuum is indeed strong and that in water a longer OHB [ordinary hydrogen bond] is the most stabilized form."**

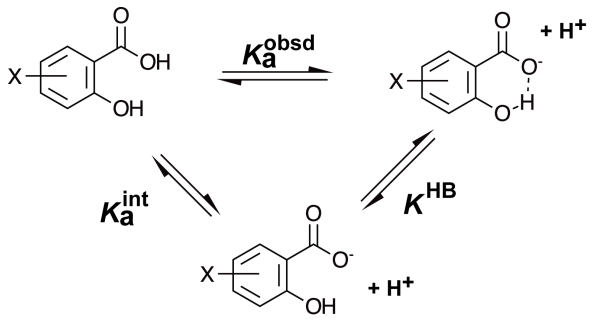
"when a gas-phase LBHB [low barrier hydrogen bond] is placed in a polar environment, it becomes an ordinary double-well HB [hydrogen bond] due to solvation effects."

9) Feierberg I and Aqvist J. *Biochemistry* (2002) 41: 15728-15735: **"[For] the H-bond in a vacuum….the donor-acceptor (O-O) distance is 2.49 Å. In water, on the other hand, the H-bond has a donor-acceptor distance of 2.8 Å."**

10) Perrin C. *Science* (1994) 266: 1665-1668. **"the importance of the local environment in determining the symmetry of the hydrogen bond is demonstrated by the fact that the symmetry seen in crystals or nonpolar solvents is broken in aqueous solution….This is because a polar environment stabilizes a concentrated negative charge more than it does a delocalized one, as in a symmetric structure."**

11) Mildvan AS, Harris TK, and Abeygunawardana C. *Meth. Enzymol.* (1999) 308: 219-245. **"As O-H•••O hydrogen bond lengths decrease from 2.54 to 2.45 Å, their strengths ($\Delta E_{formation}$) increase rapidly from -7.8 to -32 kcal/mol, presumably due to exponentially increasing overlap of the proton and oxygen wave functions."**

Text S3. Hydrogen bond formation free energies (ΔG_f) for substituted salicylates. As discussed in detail in a prior publication,[19] the free energy change for O-H•••O^- hydrogen bond formation (ΔG_f) in substituted salicylate monoanions reflects the equilibrium interaction between the OH group and COO^- group after ionization of the COOH moiety. Hydrogen bond ΔG_f values were determined as previously described[19] by first calculating the intrinsic pK_a of the COOH group in each substituted salicylate, measuring the observed pK_a of that salicylic acid in the given solvent conditions, and then using these two values to determine the equilibrium constant of hydrogen bond formation for that system, as depicted in the scheme below.



$$K^{HB} = K_a^{obsd}/K_a^{int}$$

Text S4. Calculated hydrogen bond distances in a vacuum versus a polarizable continuum. For the water-hydroxide dimer, the formic acid-formate dimer, and 5-methylsalicylate, *ab initio* calculations indicated changes on the scale of 0.03-1 Å or larger when comparing distances calculated at dielectric values ≥5 to distances calculated at dielectric values of 1.0 (vacuum) or 1.4 (argon gas) (Table S3). The physical transition from a vacuum to solvent (including polarizable continuum models) is complex, and we have on-going theoretical work to understand both the structural changes in this very low dielectric region and their relationship to formation energy. We have focused in the present manuscript on the region spanned by dielectric values of 5-80 as this region corresponds to the solution regime in question (i.e. chloroform to water) and in which large differences in hydrogen bond formation free energies (ΔG_f) are observed experimentally. Focus on this region allows us to achieve our goal of testing and understanding whether these large differences in ΔG_f values (between aprotic organic solvents and water) are accompanied by proportionately large changes in hydrogen bond distances.

Text S5. The effect of varying continuum dielectric on the potential energy surface of an O-H•••O$^-$ hydrogen bond. Calculations of one-dimensional potential energy surfaces for proton transfer between the oxygen atoms of the water-hydroxide dimer indicated very similar double minima and overall shapes for dielectric values between 5 and 80. When calculations were

performed using the energy-minimized O-O distance determined at each dielectric (which differed by 0.04 Å between 5 and 80), a difference of 1.5 kcal/mol was observed between the heights of the central barrier at dielectric values of 5 and 80 (Fig. S7A). To determine whether this difference was predominantly due to the differing continuum dielectric or to the slight difference in O-O distances used in the calculations performed at each dielectric, we fixed the O-O distance at the energy-minimized value determined at a dielectric of 5 (2.533 Å) or 80 (2.572 Å) and repeated the calculations. When a constant O-O distance was used at all dielectrics, the difference in central barrier heights reduced to 0.4-0.5 kcal/mol and the absolute height of the barrier increased with increasing O-O distance (Fig. S7, B and C), indicating that differences in the height of the central barriers in Fig. S7A arose predominantly from the small differences in O-O distance rather than differential interactions with the polarizable continuum. We conclude that the overall shape of the one-dimensional potential energy surface for the water-hydroxide hydrogen bond shows little sensitivity to varying the continuum dielectric between 5 and 80, in agreement with the minimal differences in $^{1}H - {}^{2}H$ NMR chemical shift observed for salicylates in chloroform and water in Figure 3 of the main text.

Text S6. The steeper dependence of hydrogen bond formation energy (ΔG_f) on ΔpK_a in aprotic organic solvents compared to water. The results herein clarify the origin of the steeper dependence of hydrogen bond ΔG_f on ΔpK_a in aprotic organic solvents compared to water. Based on the model that hydrogen bond length is independent of solvent and determined only by ΔpK_a, the differential sensitivity of ΔG_f arises from the differential solvation energy of the separated hydrogen bond donor and acceptor groups in protic versus aprotic environments. Across a series of hydrogen bond donor and acceptor groups, as ΔpK_a approaches zero there is increased charge

density on the acceptor oxygen atom and/or the donor hydrogen. According to Model I, this enhanced atomic charge localization results in great destabilization of the dissociated donor and acceptor groups (for a given hydrogen-bonded complex) and thus a greater thermodynamic driving force for complexation and hydrogen bond formation in organic solvents relative to water. Indeed, this behavior is qualitatively predicted by simple electrostatic models that account only for differences in the stability of the dissociated donor/acceptor groups in different solvents without invoking differences in the structural properties of hydrogen bonds formed upon complexation.[19,28,29]

As ΔpK_a approaches zero, stronger interactions are expected between donor and acceptor groups, and these strengthened interactions are presumably reflected in the shorter hydrogen bonds. But as this shortening is common to all of the solvents studied herein, either the interaction energies within the complexes are the same across the solvents or stronger interaction energies in non-aqueous solvents do not result in significant hydrogen bond shortening. Regardless, the widely touted connection of hydrogen bond length and formation free energy is not observed, and hydrogen bond length should not be considered as the primary driving force for hydrogen bond stabilities.

Text S7. The structural properties of intra- and intermolecular hydrogen bonds. Intra- and intermolecular hydrogen bonds display a similar overall variation in O•••O (Fig. S3A) and H•••O (Fig. S3B) distance as a function of ΔpK_a. Nevertheless, we note that subtle structural differences between intra- and intermolecular complexes are apparent in Fig. S3C. At O-O distances ≥ ~2.6 Å, intramolecular hydrogen bonds appear to have slightly longer H•••O distances than intermolecular hydrogen bonds, despite very similar O-H lengths, indicating that intramolecular

OHO angles in this region are slightly less (by 10-20˚) than intermolecular hydrogen bonds with the same O-O distance (See Table S1 for OHO angles). Thus, while there are some small differences apparent in Fig. S3C, the overall behavior of intra- and inter-molecular hydrogen bonds appears to be very similar: as ΔpK_a decreases, the O-O and H•••O distances shorten and the O-H bond elongates.

To further investigate potential differences between intra- and intermolecular hydrogen bonds, we collated X-ray diffraction data in the Cambridge Structural Database for a range of crystalline salicylates. Analysis of this data also indicate a 0.017 Å/pK_a unit change in O•••O distance that is nearly identical to that estimated by solution NMR for salicylates and that determined by neutron diffraction for diverse intra- and intermolecular hydrogen-bonded complexes (Fig. S2). Thus, despite general constraints on atomic positions conferred by the covalently bonded architecture of salicylates, the hydrogen-bonded oxygen atoms retain sufficient mobility on the ~0.1 Å scale to adjust their relative distance with changes in ΔpK_a (over the range of ΔpK_a values explored herein).

We also note that the O-H bond length and the position of the bridging proton are not substantially affected by the covalent structure of salicylates and the intramolecular nature of the hydrogen bond. Thus, even without considering a change in O•••O distance in salicylate hydrogen bonds, Model II would still have predicted a shorter O-H bond in water (and smaller change in O-H distance with ΔpK_a) compared to aprotic solvents, in order to maximize charge localization and the interaction energy with solvent water molecules. These shorter O-H distances would have been expected to result in more upfield ^{1}H NMR chemical shifts (and a more shallow change in chemical shift with ΔpK_a) than the corresponding values in chloroform or acetone, according to Model II. Our data provide strong evidence that salicylate hydrogen

bonds maintain a very similar structure in all three protic and aprotic solvent environments.

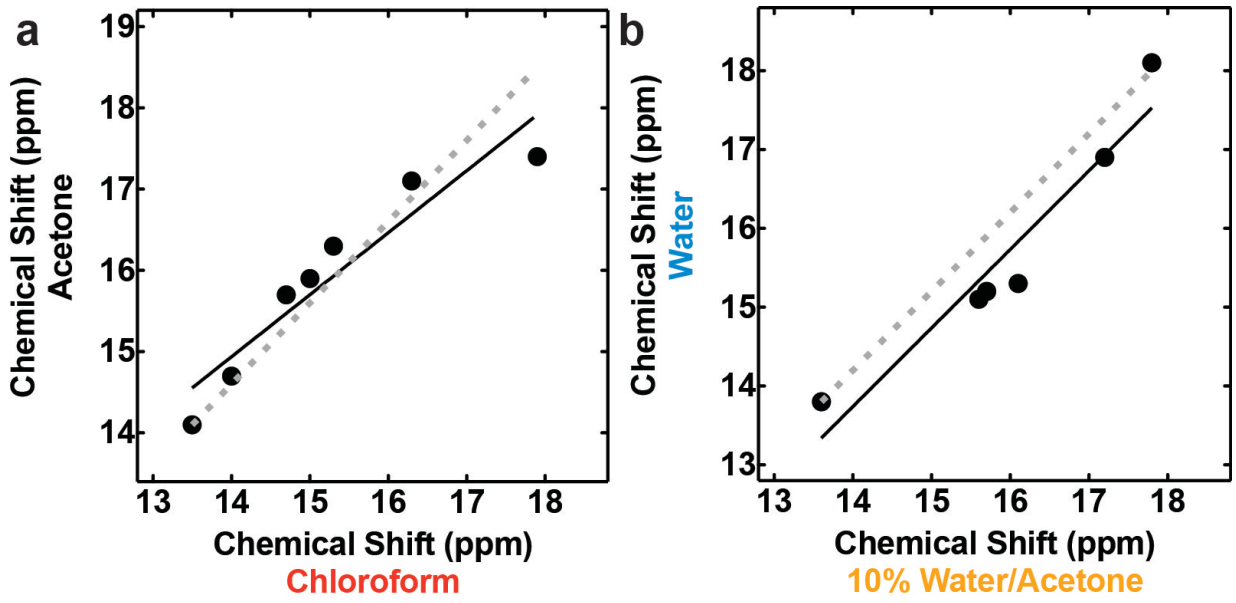


Fig. S1. Plots of the observed ^{1}H NMR chemical shift for the hydrogen-bonded proton in salicylate monoanions. Comparison of chemical shifts in acetone versus chloroform (slope = 0.8, R^2=0.89) **(a)** and water versus 10% water/acetone (slope = 0.9, R^2=0.92) **(b)**. The data for 10% water/acetone are from a prior publication[30] and were acquired at -50 ˚C. The dashed gray lines in **(a)** and **(b)** have slopes of 1.

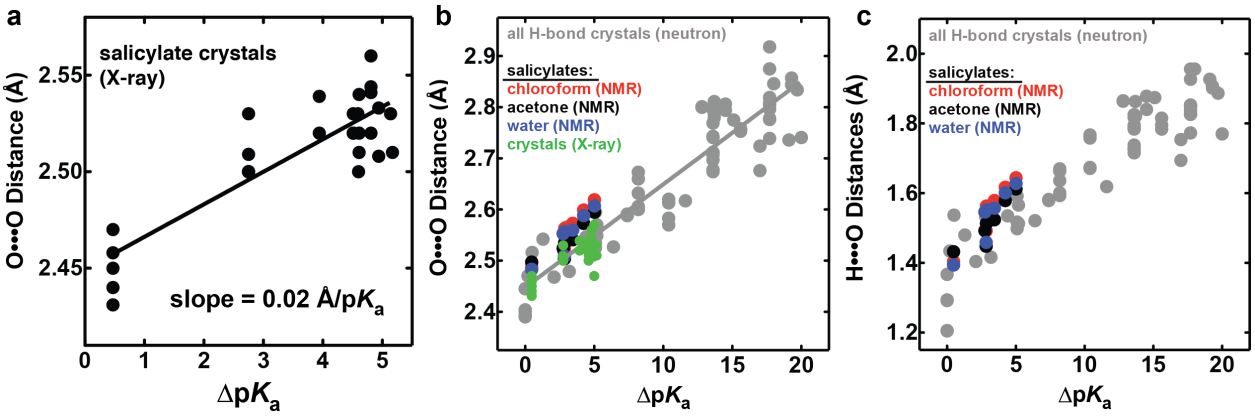


Fig. S2. Structural comparison of hydrogen bond lengths for salicylates and diverse crystalline complexes. (a) Hydrogen bond O•••O distances determined by X-ray diffraction for substituted salicylate crystals (data from Table S6). The solid black line in (a) is from a linear fit and has a slope of 0.017 Å/pK_a unit, which is very similar to the slope of 0.021 Å/pK_a unit obtained from a linear fit (gray line) of the aggregate neutron diffraction data in (b) Comparison of hydrogen bond O•••O **(b)** and H•••O **(c)** distances determined for diverse intra- and intermolecular small-molecule crystals by low-temperature neutron diffraction (data from Table S1), distances determined for substituted salicylates in chloroform, acetone, and water using solution ^{1}H NMR and previously published correlation functions (see main text), and distances determined for substituted salicylate crystals by X-ray diffraction.

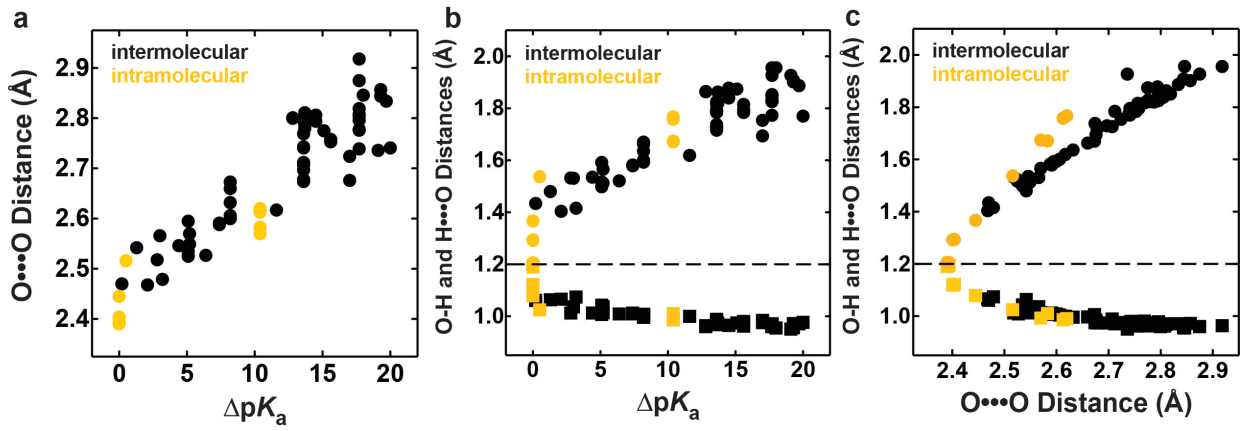


Fig. S3. Comparison of inter- versus intra-molecular O-H•••O hydrogen bond distances determined by low-temperature neutron diffraction. O•••O distance versus ΔpK_a **(a)**. O-H and H•••O distances versus ΔpK_a **(b)**. O-H and H•••O distances versus O•••O distance **(c)**. Data are from Table S1.

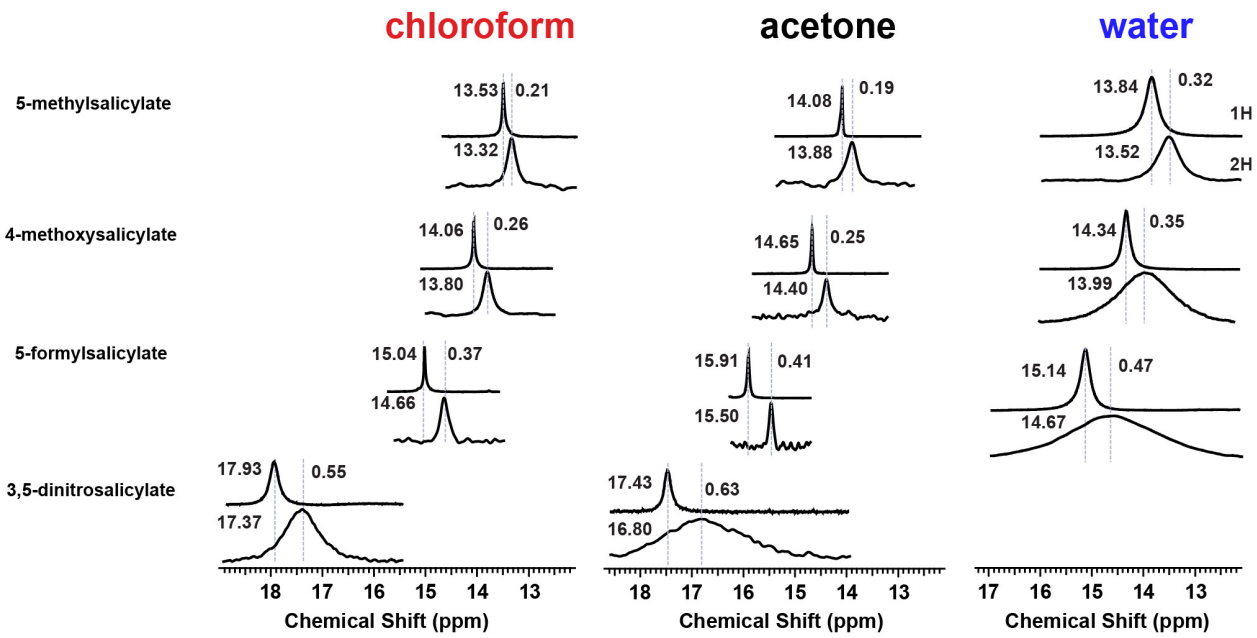


Fig. S4. ^{1}H (upper) and ^{2}H (lower) NMR spectra for substituted salicylates in chloroform, acetone, and water.

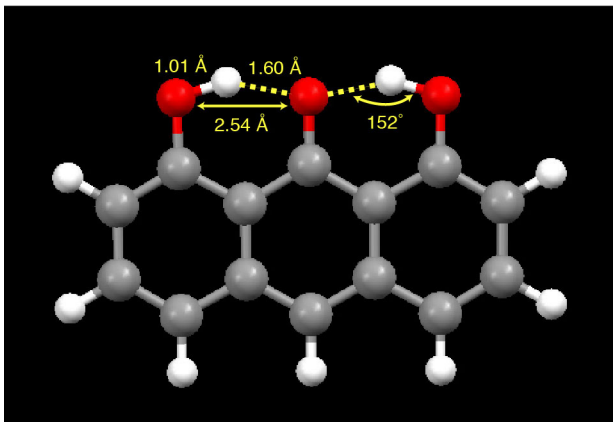


Fig. S5. Energy minimized gas-phase structure of the 1,8,9-trihydroxy anthracene monoanion, calculated at the B3LYP level with the 6-31+G* basis set.

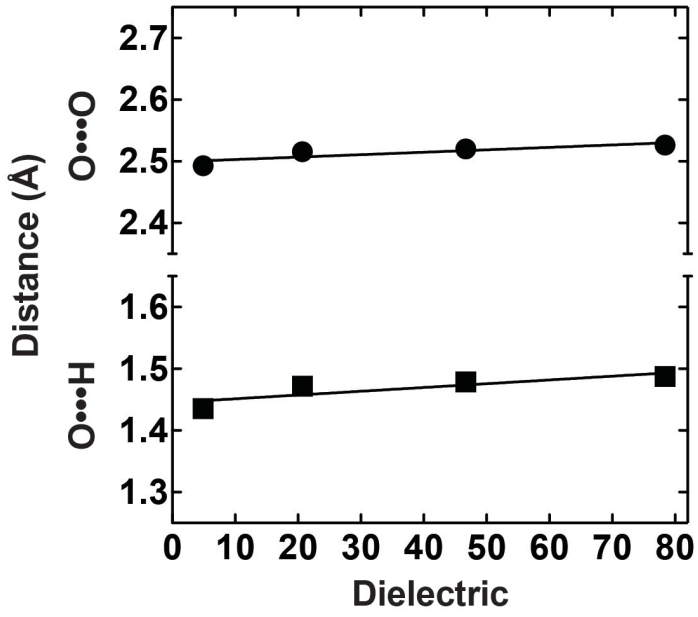


Fig. S6. Calculated effect of varying the continuum dielectric on the structure of the formic acid-formate hydrogen bond. Calculations were performed at the B3LYP level using the 6-311++G(d,p) basis set.

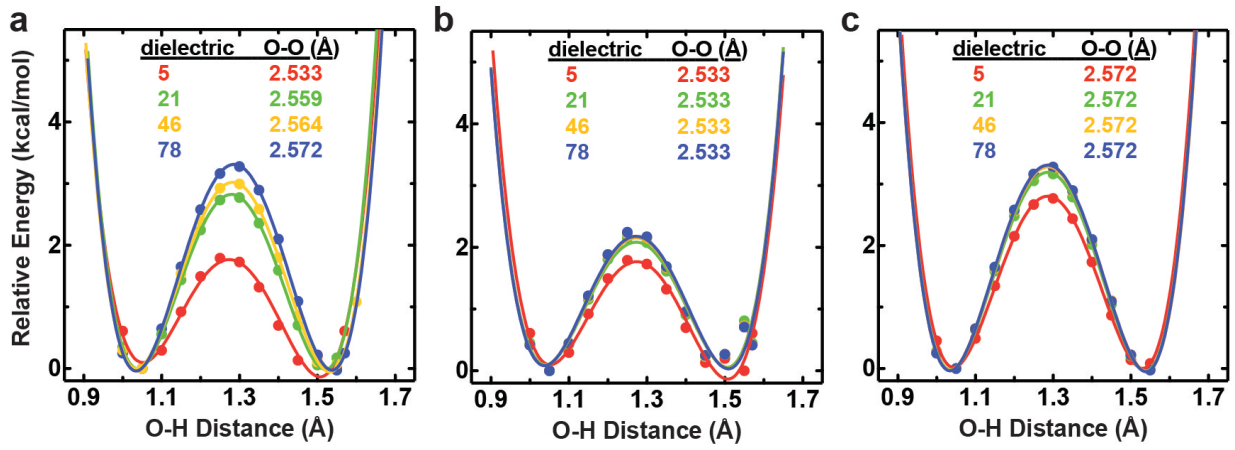


Fig. S7. QM potential energy surface calculations. One-dimensional potential energy curves for displacement of the hydrogen-bonded proton between the donor and acceptor oxygen atoms of the water-hydroxide dimer using **(a)** the energy-minimized O-O distance calculated at each dielectric value, **(b)** a constant O-O distance corresponding to the energy-minimized distance at a dielectric of 5, and **(c)** a constant O-O distance corresponding to the energy-minimized distance at a dielectric of 78. All calculations were performed at the B3LYP level using the 6-311++G(d,p) basis set. The O-O distance was kept constant in each calculation, the bridging proton was displaced between the oxygen atoms in the given increments, and the energy of the system was then calculated relative to the minimal energy configuration (which was set to zero kcal/mol). The data were empirically fit with a fourth order polynomial.

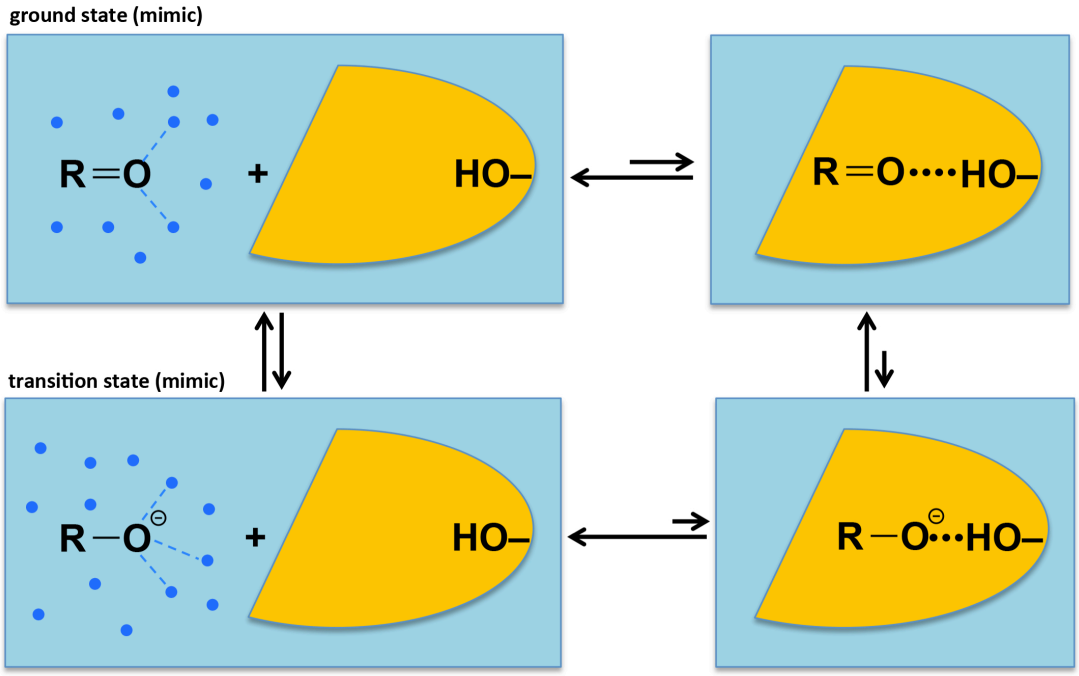


Fig. S8. Schematic depiction of hydrogen bond formation to a substrate or transition state bound within a hydrophobic enzyme active or in aqueous solution. The hydrogen-bonded state would be expected to be less stable in a hydrophobic or low-dielectric environment (yellow) than in water (blue), in the absence of additional stabilizing features such as additional hydrogen bonds, pre-positioned hydrogen bonding groups, and/or favorably oriented or rearranging surrounding enzymatic dipolar and charged groups. With increased charge localization in the transition state (bottom), association with the hypothetical hydrophobic enzyme would become less, rather than more, favorable. Although there could be additional favorable hydrophobic interactions with the R group, these interactions would be the same for R=O and R-O^- and thus not change the effect from charge localization. For, simplicity only negative charge localizes in the transition state is shown, and counter ions are not depicted.

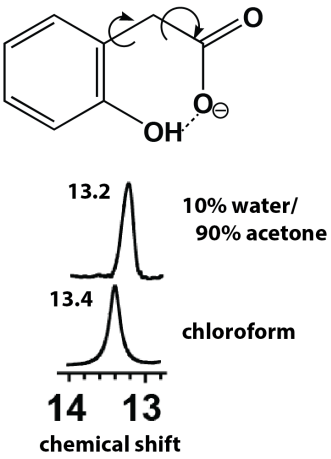


Fig. S9. ^{1}H NMR spectra for the hydrogen-bonded proton in 2-hydroxyphenylacetate in chloroform and 10% water/acetone.

Table S1. Analysis of O-H•••O hydrogen bonds from low-temperature neutron structures compiled previously[15] or determined as part of this study. Sources for estimating pK_a values: 1 = ref.[17], 2 = ref.[18], 3 = ACD/Labs I-Lab 2.0 pK_a Prediction Module, MM = as explained in Materials and Methods in the main text.

		Estimated pK_a Values					Distance (Å)				
Description	**accession code**	**donor pK_a**	**pK_a source**	**acceptor pK_a**	**pK_a source**	**ΔpK_a**	**O-H**	**H---O**	**O-O**	**OHO Angle (°)**	**type**
arabinose	ABINOR04	11.6	3	-4.0	3	15.6	0.980	1.784	2.753	169.290	inter
		11.6	3	-2.0	3	13.6	0.989	1.726	2.707	171.600	inter
		11.6	3	-2.0	3	13.6	0.981	1.817	2.769	162.980	inter
		11.6	3	-2.0	3	13.6	0.979	1.825	2.796	170.860	inter
	ABINOS02	11.6	3	-4.0	3	15.6	0.974	1.795	2.758	169.260	inter
		11.6	3	-2.0	3	13.6	0.975	1.738	2.674	163.680	inter
		11.6	3	-2.0	3	13.6	0.972	1.795	2.741	159.830	inter
adenosine	ADENOS01	11.6	3	-2.0	2	13.6	0.980	1.787	2.743	163.980	inter
phosphorylethanolamine	AEPHOS02	5.9	3	1.5	3	4.4	1.013	1.535	2.546	176.060	inter
asparagine•H2O	ASPARM06	15.7	1	2.0	1	13.7	0.968	1.821	2.780	170.620	inter
		15.7	1	2.0	1	13.7	0.969	1.863	2.811	165.280	inter
naphthazarin	DHNAPH17	7.4	3	-3.0	2	10.4	0.991	1.767	2.620	141.980	**intra**
		7.4	3	-3.0	2	10.4	0.993	1.674	2.570	147.800	**intra**
		7.4	3	-3.0	2	10.4	1.010	1.671	2.583	147.950	**intra**
		7.4	3	-3.0	2	10.4	0.986	1.758	2.613	142.860	**intra**
ferrocenedicarboxylic acid	FEROCA12	4.2	1	-4.0	2	8.2	1.005	1.601	2.606	179.120	inter
		4.2	1	-4.0	2	8.2	0.997	1.663	2.660	178.410	inter
		4.2	1	-4.0	2	8.2	1.008	1.593	2.600	176.010	inter
		4.2	1	-4.0	2	8.2	0.996	1.636	2.632	178.380	inter
succinate	KACBEV01	5.5	3	4.2	3	1.3	1.064	1.480	2.542	176.340	inter
K oxalate perhydrate	KOXPHY11	11.6	1	4.2	1	7.4	1.012	1.579	2.591	178.370	inter
N-acetyl-L-cysteine	NALCYS02	3.2	3	-2.0	2	5.2	1.036	1.513	2.549	178.370	inter
Li hydrogen phthalate	LIHPAL01	15.7	1	-2.0	2	17.7	0.970	1.837	2.804	174.660	inter
		3.0	3	3.0	3	0.0	1.122	1.294	2.404	168.870	**intra**
		15.7	1	-2.0	2	17.7	0.967	1.826	2.778	167.480	inter
		15.7	1	3.0	3	12.8	0.964	1.864	2.801	163.490	inter
		3.0	3	3.0	3	0.0	1.195	1.205	2.394	171.880	**intra**
	LIHPAL02	15.7	1	-2.0	2	17.7	0.962	1.848	2.808	175.450	inter
		15.7	1	-2.0	2	17.7	0.963	1.827	2.776	167.770	inter
		15.7	1	3.0	3	12.8	0.960	1.865	2.800	163.940	inter
		3.0	3	3.0	3	0.0	1.119	1.292	2.401	168.900	**intra**
		3.0	3	3.0	3	0.0	1.190	1.206	2.390	171.950	**intra**
sodium maleate	NAHMAL01	15.7	1	-4.0	2	19.7	0.971	1.887	2.834	164.480	inter
		15.7	1	-2.0	2	17.7	0.971	1.851	2.819	174.330	inter
		15.7	1	-2.0	3	17.7	0.960	1.927	2.875	168.820	inter
		15.7	1	-2.0	3	17.7	0.970	1.773	2.739	173.160	inter
		2.0	1	2.0	3	0.0	1.079	1.367	2.445	176.070	**intra**
		15.7	1	2.0	3	13.7	0.973	1.820	2.788	172.300	inter
		15.7	1	-2.0	3	17.7	0.964	1.956	2.918	175.340	inter
Na oxalate perhydrate	NAOXAP11	11.6	1	4.2	1	7.4	1.009	1.582	2.588	174.380	inter
sodium oxalate	NHOXAL14	15.7	1	1.2	1	14.5	0.965	1.878	2.794	157.320	inter
		4.2	1	1.2	1	3.0	1.036	1.531	2.566	176.750	inter
		15.7	1	1.2	1	14.5	0.966	1.840	2.806	177.860	inter
ribosyluronic acid•uracil	NRURAM11	13.1	3	-6.0	2	19.1	0.951	1.927	2.736	141.390	inter
		16.0	2	-2.0	2	18.0	0.955	1.956	2.846	154.220	inter
		3.5	3	-1.7	1	5.2	1.014	1.566	2.570	169.810	inter
		15.7	1	-2.0	2	17.7	0.967	1.853	2.796	164.090	inter
		13.1	3	-2.0	2	15.1	0.974	1.874	2.775	152.550	inter
putrescine diphosphate	PUTRDP11	7.2	1	2.1	1	5.1	1.007	1.592	2.595	173.920	inter
		7.2	1	2.1	1	5.1	1.015	1.517	2.525	171.160	inter
succinic acid	SUCACB03	4.2	3	-4.0	2	8.2	1.006	1.670	2.673	174.270	inter
urea•hydrogen peroxide	UREXPO11	11.6	1	0.0	2	11.6	1.000	1.619	2.617	174.530	inter
urea•oxalic acid	UROXALO1	1.2	3	-2.0	2	3.2	1.074	1.416	2.479	168.920	inter
tetracyclodecane-2-one	VUYYUJ	15.3	3	-4.0	2	19.3	0.972	1.902	2.857	166.590	inter
		15.3	3	-4.0	2	19.3	0.956	1.908	2.844	165.330	inter
L-xylopyranose	XYLOSE03	11.6	3	-4.0	3	15.6	0.962	1.815	2.756	165.110	inter
		11.6	3	-2.0	3	13.6	0.971	1.784	2.712	158.780	inter
		11.6	3	-2.0	3	13.6	0.974	1.716	2.678	168.710	inter
		11.6	3	-2.0	3	13.6	0.974	1.729	2.697	171.690	inter
methylpregnene-3,20-diol methanolate	MPRGOM01	18.0	1	-2.0	2	20.0	0.977	1.770	2.741	171.720	inter
		15.0	1	-2.0	2	17.0	0.976	1.754	2.724	171.640	inter
		15.0	3	-2.0	2	17.0	0.984	1.694	2.676	175.640	inter
STRUCTURES DETERMINED AS PART OF THIS STUDY											
2,4-dinitrophenol•4-(dimethylamino)pyridine-N-oxide	978290	4.1	1	3.9	1	0.2	1.061	1.434	2.470	163.520	inter
2,4-dinitrophenol•4-methoxypyridine-N-oxide	978287	4.1	1	2.1	1	2.1	1.066	1.404	2.468	175.710	inter
2,4-dinitrophenol•4-methylpyridine-N-oxide	978288	4.1	1	1.3	1	2.8	1.013	1.532	2.518	162.640	inter
4-nitrophenol•4-methoxypyridine-N-oxide	978289	7.1	1	2.1	1	5.1	1.043	1.498	2.535	171.530	inter
4-nitrophenol•pyridine-N-oxide	978291	7.1	1	0.8	1	6.4	1.009	1.521	2.527	174.200	inter
3,5-dinitrosalicylate	980683	3.6	MM	3.1	MM	0.5	1.025	1.537	2.516	157.990	**intra**

Table S2. Comparison of hydrogen bond distances determined by NMR and neutron diffraction for 3,5-dinitrosalicylate. The neutron structure of 3,5-dinitrosalicylate was determined herein (Table SI).

Environment	Method	O•••O Distance (Å)	H•••O Distance (Å)
crystal	neutron diffraction	2.516	1.537
chloroform	1H NMR	2.487	1.404
acetone	1H NMR	2.497	1.432
water	1H NMR	2.483	1.394

Table S3. QM calculations of hydrogen bond O•••O, O-H, and H•••O distances (Å) for the water-hydroxide dimer, formic acid-formate dimer, and 5-methylsalicylate.

		Calculated Distances (Å)								
		HF								
		6-31G(d)			6-31+G(d)			6-311++G(d,p)		
Complex	Dielectric	O•••O	O-H	H•••O	O•••O	O-H	H•••O	O•••O	O-H	H•••O
water-hydroxide	1.0	2.525	1.040	1.490	2.614	1.005	1.618	2.578	1.002	1.582
	1.4	2.489	1.055	1.437	2.625	0.999	1.63	2.594	0.995	1.603
	4.9	2.587	1.005	1.583	2.673	0.986	1.689	2.643	0.980	1.665
	20.7	2.612	0.996	1.616	2.699	0.980	1.719	2.676	0.974	1.703
	46.7	2.609	0.996	1.612	2.705	0.979	1.726	2.682	0.973	1.710
	78.4	2.610	0.996	1.614	2.707	0.979	1.728	2.685	0.972	1.714
		B3LYP								
		6-31G(d)			6-31+G(d)			6-311++G(d,p)		
	Dielectric	O•••O	O-H	H•••O	O•••O	O-H	H•••O	O•••O	O-H	H•••O
	1.0	2.468	1.236	1.237	2.514	1.106	1.409	2.446	1.224	1.223
	1.4	2.464	1.234	1.231	2.533	1.083	1.45	2.497	1.089	1.409
	4.9	2.492	1.107	1.385	2.565	1.052	1.513	2.535	1.051	1.484
	20.7	2.511	1.083	1.429	2.587	1.039	1.548	2.559	1.037	1.522
	46.7	2.514	1.080	1.434	2.591	1.037	1.554	2.564	1.035	1.529
	78.4	2.518	1.076	1.443	2.598	1.035	1.563	2.572	1.030	1.542
		MP2								
		6-31G(d)			6-31+G(d)			6-311++G(d,p)		
	Dielectric	O•••O	O-H	H•••O	O•••O	O-H	H•••O	O•••O	O-H	H•••O
	1.0	2.464	1.235	1.233	2.514	1.106	1.409	2.430	1.217	1.213
	1.4	2.475	1.227	1.249	2.589	1.054	1.54	2.492	1.071	1.422
	4.9	2.514	1.083	1.432	2.620	1.035	1.586	2.529	1.039	1.490
	20.7	2.533	1.066	1.467	2.636	1.026	1.611	2.542	1.029	1.514
	46.7	2.533	1.064	1.469	2.639	1.025	1.615	2.543	1.028	1.517
	78.4	2.536	1.062	1.473	2.640	1.025	1.616	2.545	1.027	1.520
formic acid-formate		HF								
		6-31G(d)			6-31+G(d)			6-311++G(d,p)		
	Dielectric	O•••O	O-H	H•••O	O•••O	O-H	H•••O	O•••O	O-H	H•••O
	1.0	2.545	1.002	1.552	2.575	0.991	1.598	2.538	0.990	1.559
	4.9	2.611	0.986	1.625	2.642	0.982	1.660	2.606	0.979	1.627
	20.7	2.640	0.982	1.658	2.670	0.978	1.692	2.639	0.974	1.665
	46.7	2.643	0.981	1.661	2.676	0.977	1.698	2.647	0.973	1.673
	78.4	2.655	0.980	1.675	2.686	0.977	1.710	2.660	0.972	1.688
		B3LYP								
		6-31G(d)			6-31+G(d)			6-311++G(d,p)		
	Dielectric	O•••O	O-H	H•••O	O•••O	O-H	H•••O	O•••O	O-H	H•••O
	1.0	2.430	1.170	1.260	2.472	1.101	1.372	2.433	1.127	1.306
	4.9	2.493	1.071	1.422	2.535	1.051	1.484	2.493	1.058	1.435
	20.7	2.514	1.057	1.458	2.557	1.040	1.517	2.516	1.044	1.472
	46.7	2.518	1.055	1.463	2.562	1.038	1.523	2.520	1.042	1.479
	78.4	2.524	1.052	1.473	2.570	1.036	1.534	2.527	1.039	1.487
		MP2								
		6-31G(d)			6-31+G(d)			6-311++G(d,p)		
	Dielectric	O•••O	O-H	H•••O	O•••O	O-H	H•••O	O•••O	O-H	H•••O
	1.0	2.472	1.102	1.372	2.509	1.077	1.433	2.413	1.133	1.280
	4.9	2.530	1.053	1.479	2.570	1.041	1.529	2.468	1.057	1.411
	20.7	2.550	1.044	1.508	2.590	1.034	1.556	2.487	1.045	1.442
	46.7	2.554	1.042	1.513	2.593	1.032	1.561	2.490	1.043	1.447
	78.4	2.555	1.042	1.515	2.595	1.032	1.563	2.491	1.042	1.450
5-methylsalicylate		HF								
		6-31G**								
	Dielectric	O•••O	O-H	H•••O						
	1.0	2.520	0.980	1.610						
	4.9	2.550	0.970	1.660						
	78.4	2.550	0.970	1.670						
		B3LYP								
		6-31G**								
	Dielectric	O•••O	O-H	H•••O						
	1.0	2.420	1.100	1.350						
	4.9	2.460	1.040	1.470						
	78.4	2.470	1.040	1.490						

Table S4. Data collection and refinement statistics for X-ray diffraction experiments. Formulas for extinction and weighting can be found in the cif files deposited with the Cambridge Structural Database.

	3,5-dinitrosalicylate N,N,N',N'-tetramethyl-1,8-naphthalenediamine	2,4-dinitrophenol 4-methylpyridine-N-oxide	4-nitrophenol 4-methoxypyridine-N-oxide	2,4-dinitrophenol 4-dimethylamino-pyridine-N-oxide
formula	$C_{21}H_{22}N_4O_7$	$C_{12}H_{11}N_3O_6$	$C_{12}H_{12}N_2O_5$	$C_{13}H_{14}N_4O_6$
fw	442.43	293.24	264.24	322.28
temperature, K	293(2)	173(2)	173(2)	273(2)
crystal system	*monoclinic*	*orthorhombic*	*triclinic*	*orthorhombic*
space group	$P2_1/n$	$P2_12_12_1$	P-1	$P2_12_12_1$
a, Å	9.9407(5)	6.5577(8)	6.6545(4)	6.8204(13)
b, Å	13.6994(7)	7.6142(10)	6.7300(4)	7.4543(14)
c, Å	15.9341(8)	25.315(3)	14.5060(9)	28.371(5)
α, ϒ	90	90	87.7810(10)	90
β, ϒ	102.3890(10)	90	80.5520(10)	90
γ, ϒ	90	90	70.1870(10)	90
V, Å^3	2119.40(19)	1264.0(3)	602.80(6)	1442.4(5)
Z	4	4	2	4
d_{calc}, g cm^{-3}	1.387	1.541	1.456	1.484
dimensions, mm	0.3 x 0.2 x 0.1	0.4 x 0.2 x 0.2	0.25 x 0.15 x 0.1	0.3 x 0.3 x 0.1
radiation	Mo(Kα), 0.71073 Å			
μ(Mo), cm^{-1}	0.106	0.126	0.115	0.12
extinction parameter	0.0026(13)	0.0033(17)	0.000(6)	0.0018(10)
no. of measured reflns	20979	21207	10936	12774
no. of unique reflns	3487	4263	4453	2073
no. of reflns ($I > 2\sigma(I)$)	2510	3131	3303	1841
no. of parameters refined/restraints	293 / 0	235 / 0	221 / 0	265 / 0
refinement method	Full-matrix least-squares on F^2			
R1	0.0531	0.0428	0.051	0.0288
wR2	0.143	0.0948	0.1333	0.0672
goodness-of-fit	1.08	1.025	1.027	1.007
CCDC deposition #	980682	978285	978286	978284

Table S5. Data collection and refinement statistics for neutron diffraction experiments.

	3,5-dinitrosalicylate N,N,N',N'-tetramethyl-1,8-naphthalenediamine	2,4-dinitrophenol 4-methylpyridine-N-oxide	4-nitrophenol 4-methoxypyridine-N-oxide	2,4-dinitrophenol 4-dimethylamino-pyridine-N-oxide	4-nitrophenol pyridine-N-oxide	2,4-dinitrophenol 4-methoxypyridine-N-oxide
formula	$C_{21}H_{22}N_4O_7$	$C_{12}H_{11}N_3O_6$	$C_{12}H_{12}N_2O_5$	$C_{13}H_{14}N_4O_6$	$C_{11}H_{10}N_2O_4$	$C_{12}H_{11}N_3O_7$
fw	442.43	293.235	264.24	322.28	234.21	309.23
temperature, K	30(1)	30(1)	30(1)	30(1)	30(1)	30(1)
crystal system	*monoclinic*	*orthorhombic*	*triclinic*	*orthorhombic*	*monoclinic*	triclinic
space group	$P2_1/n$	$P2_12_12_1$	P-1	$P2_12_12_1$	$P2_1$/c	P -1
a, Å	9.544(2)	6.55(2)	6.666(2)	6.860(1)	7.725(3)	6.6850(10)
b, Å	13.845(2)	7.45(2)	6.632(2)	7.189(2)	6.013(3)	7.4440(10)
c, Å	15.765(3)	25.18(7)	14.486(5)	28.087(7)	22.338(5)	14.3080(20)
α, ϒ	90	90	87.095(2)	90	90	88.163(13)
β, ϒ	101.81(2)	90	80.761(3)	90	94.39(2)	89.038(13)
γ, ϒ	90	90	70.299(2)	90	90	114.820(12)
V, Å^3	2039.0(7)	1228.7(6)	595.1(3)	1385.2(5)	1034.6(7)	645.55(17)
Z	4	4	2	4	4	2
d_{calc}, g cm^{-3}	1.441	1.541	1.475	1.545	1.504	1.591
dimensions, mm	3 x 1.5 x 1	0.8 x 3 x 4	4 x 2 x 1	4 x 2 x 1	3 x 1 x 1	2 x 1 x 0.8
radiation	neutrons					
data collection technique	time-of-flight Laue					
$\mu(\lambda)$, cm^{-1}	1.266 + 0.837λ	1.118 + 0.678	1.203 + 0.782λ	1.247 + 0.788λ	1.189 + 0.752λ	1.030 + 0.662λ
max, min transmission	0.8422, 0.4383	0.8155, 0.3893	0.8081, 0.3673	0.8101, 0.4099	0.8717, 0.5280	0.8926, 0.5837
extinction parameter	9.5(5) x 10^{-6}	3.5(1) x 10^{-5}	4.1(1) x 10^{-05}	9.8(3) x 10^{-06}	7.4(4) x 10^{-06}	7.6(2) x 10^{-06}
d_{min}, Å	0.7	0.5	0.5	0.5	0.5	0.5
no. of reflns	2204	4493	4038	4116	3187	2961
no. of reflns (I> 3s(I))a	1562	3094	3044	2601	1789	1844
no. of parameters refined	245	311	324	356	270	342
refinement method	Full-matrix least-squares on F^2					
R indices $R_w(F^2)$[b], $R(F^2)$[c]	0.165, 0.165	0.154, 0.182	0.137, 0.153	0.129, 0.151	0.152, 0.153	0.116, 0.141
goodness-of-fit	1.78	1.93	1.50	1.57	1.32	1.49
CCDC deposition #	980683	978288	978289	978290	978291	978287

[a] Outliers with $\left| F_o^2 / F_c^2 \right| > 2$, $\left| F_c^2 / F_o^2 \right| > 2$ were rejected.

[b] $R_w(F^2) = (\sum[w(F_o^2 - F_c^2)^2] / \sum[w(F_o^2)^2])^{1/2}$

[c] $R(F^2) = \Sigma|F_o^2 - F_c^2| / \Sigma|F_o^2|$

Table S6. Analysis of salicylate X-ray structures deposited with the Cambridge Structural Database (CSD). Indicated ΔpK_a values were calculated as described in Materials and Methods.

Salicylate	ΔpK_a	CSD	O•••O (Å)
3,5-dinitro	0.47	ABELUP	2.44
3,5-dinitro	0.47	ABUJUE	2.45
3,5-dinitro	0.47	ABUKAL	2.47
3,5-dinitro	0.47	AHEROU	2.43
3,5-dinitro	0.47	AJEBIA	2.46
5-nitro	2.75	MIPRIM	2.51
5-nitro	2.75	ZIVROL	2.50
5-nitro	2.75	ACACUE	2.50
5-nitro	2.75	AFUJES	2.50
5-nitro	2.75	FAQVAV	2.53
5-nitro	2.75	FASJEQ	2.50
5-SO3-	3.94	CASTAR	2.54
5-SO3-	3.94	ETIZOW	2.52
5-I	4.51	ACACAK	2.52
5-I	4.51	ACACEO	2.52
5-I	4.51	CELXEY	2.53
5-Br	4.59	CELXAU	2.53
6-F	4.60	CELWUN	2.50
5-Cl	4.61	CELYAV	2.51
5-Cl	4.61	CELYEZ	2.52
5-Cl	4.61	ENUFOJ	2.54
H	4.81	APALUZ	2.56
H	4.81	CAQMEM	2.52
H-	4.81	BAKYES	2.54
H-	4.81	EDATUS	2.54
5-OH	4.94	EWINAZ	2.51
5-OH	4.94	RACBED	2.53
5-F	5.14	ACABUD	2.53
5-OMe	5.17	ACACIS	2.51

REFERENCES

(1) Bruker *APEX2 User Manual*, 2004.
(2) Bruker *Saint-Plus*, 2004.
(3) Bruker *SADABS*, 2004.
(4) Sheldrick, G. M. *Acta Crystallogr. A* **2008**, *64*, 112-122.
(5) Schultz, A. J.; De Lurgio, P. M.; Hammonds, J. P.; Mikkelson, D. J.; Mikkelson, R. L.; Miller, M. E.; Naday, I.; Peterson, P. F.; Porter, R. R.; Worlton, T. G. *Physica B* **2006**, *385-86*, 1059-1061.
(6) Piccoli, P. M. B.; Koetzle, T. F.; Schultz, A. J.; Zhurova, E. A.; Stare, J.; Pinkerton, A. A.; Eckert, J.; Hadzi, D. *J. Phys. Chem. A* **2008**, *112*, 6667-6677.
(7) Jacobson, R. A. *J. Appl. Crystallogr.* **1976**, *9*, 115-118.
(8) Sears, V. F. *In Methods of Experimental Physics, Vol. 23, Neutron Scattering, Part A*; Academic Press: Orlando, FL, 1986.
(9) Howard, J. A. K.; Johnson, O.; Schultz, A. J.; Stringer, A. M. *J. Appl. Cryst.* **1987**, *20*, 120-122.
(10) Larson, A. C.; Von Dreele, R. B. *General Structure Analysis System (GSAS)*, 2004.
(11) Burnett, M. N.; Johnson, C. K. *Oak Ridge National Laboratory Report ORNL-6895* **1996**.
(12) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. *J. Appl. Crystallogr.* **2008**, *41*, 466-470.
(13) Montis, R.; Hursthouse, M. B. *Crystengcomm.* **2012**, *14*, 7466-7478.
(14) Smith, G.; Wermuth, U. D.; Healy, P. C.; White, J. M. *J Chem. Crystallogr.* **2011**, *41*, 1649-1662.
(15) Steiner, T.; Saenger, W. *Acta Crystallogr., Sect. B: Struct. Sci.* **1994**, *50*, 348-357.
(16) Allen, F. H. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 380-388.
(17) Jencks, W. P.; Regenstein, J. In *Handbooks of Biochemistry and Molecular Biology*; Fasman, G. D., Ed.; CRC Press: Cleveland, **1976**, p 305-351.
(18) Lowry, T. H.; Richardson, K. S. *Mechanism and Theory in Organic Chemistry*; Harper & Row, 1987.
(19) Shan, S. O.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14474-14479.
(20) Czerwinska, M.; Sikora, A.; Szajerski, P.; Zielonka, J.; Adamus, J.; Marcinek, A.; Piech, K.; Bednarek, P.; Bally, T. *J. Org. Chem.* **2006**, *71*, 5312-5319.
(21) Turner, D. L. *J. Magn. Reson.* **1983**, *54*, 146-148.
(22) Jeffrey, G. A.; Yeon, Y. *Acta Crystallogr., Sect. B: Struct. Sci.* **1986**, *42*, 410-413.
(23) Harris, T. K.; Mildvan, A. S. *Proteins* **1999**, *35*, 275-282.
(24) McDermott, A.; Rydenour, C. F. In *Encyl. NMR*; Grant, D. M., Harris, R. K., Eds.; John Wiley and Sons (Sussex, England): 1996, p 3820-3825.
(25) Harris, R. K.; Jackson, P.; Merwin, L. H.; Say, B. J.; Hagele, G. *J. Chem. Soc. Faraday Trans. I* **1988**, *84*, 3649-3672.
(26) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A *Gaussian 03*; Revision C.02 ed.; Gaussian, Inc.: Wallingford CT, 2004.
(27) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999-3093.
(28) Stahl, N.; Jencks, W. P. *J. Am. Chem. Soc.* **1986**, *108*, 4196-4205.
(29) Hine, J. *J. Am. Chem. Soc.* **1972**, *94*, 5766-5771.

(30) Mock, W. L.; Morsch, L. A. *Tetrahedron* **2001**, *57*, 2957-2964.