Two Concomitant Polymorphs of a Supramolecular Model of the Asp···His···Ser Catalytic Triad

ARTICLE *in* CRYSTAL GROWTH & DESIGN · JULY 2005

Impact Factor: 4.89 · DOI: 10.1021/cg050092+

3 AUTHORS, INCLUDING:

John C Macdonald
Worcester Polytechnic Institute
46 PUBLICATIONS **3,739** CITATIONS

CRYSTAL GROWTH & DESIGN
2005
VOL. 5, NO. 6
2248−2255

Two Concomitant Polymorphs of a Supramolecular Model of the Asp···His···Ser Catalytic Triad

John C. MacDonald,* Mehmet V. Yigit, and Kyle Mychajlonka

Department of Chemistry & Biochemistry, Worcester Polytechnic Institute, Worcester, Massachusetts 01609-2280

Received March 13, 2005; Revised Manuscript Received May 31, 2005

ABSTRACT: The supramolecular chemistry and crystal structures of two concomitant polymorphs (**1** and **2**) of a 1:2:2 complex between betaine, imidazole, and picric acid are reported. Cocrystallization of these complexes gives discrete aggregates composed of five different molecules joined by four N−H···O and four C−H···O hydrogen bonds. These interactions define a supramolecular unit with a unique composition, structure, and connectivity between the molecular components that is common to both polymorphs despite considerable differences in crystal packing. The network of hydrogen bonds between betaine, imidazole, and picric acid mimics the connectivity present in the Asp···His···Ser catalytic triad and serves as a structural model for hydrogen bonding in the active site of serine proteases.

Introduction

Serine proteases are a ubiquitous class of proteins present in all organisms. These proteins are comprised of the trypsin and the subtilisin families, which independently have evolved a similar catalytic mechanism to cleave peptide bonds.[1−5] This mechanism utilizes aspartate, histidine, and serine residues, which form the hydrogen-bonded network illustrated in Figure 1. The functional importance of the Asp···His···Ser catalytic triad in catalytic cleavage of peptide bonds has been well established.[6,7] The nucleophilic oxygen on the Ser residue promotes hydrolysis of peptide bonds by attacking the carbonyl carbon of the peptide on a substrate, yielding a tetrahedral intermediate that is stabilized by neighboring residues in the active site. The His residue plays a dual role as a proton acceptor as well as a proton donor in separate steps during hydrolysis. The Asp residue is believed to serve a structural role by aligning His in the proper orientation to facilitate nucleophilic attack of the substrate by the Ser residue. Until recently, the positions of hydrogen atoms involved in hydrogen bonding in the Asp···His···Ser triad could not be resolved in the crystal structures of serine proteases. Consequently, the positions of hydrogen atoms were inferred from the distances and geometry of heavy atoms in the residues. A recent report of the crystal structure of serine protease proteinase K determined at high resolution (0.98 Å) using synchrotron radiation and crystals grown under microgravity conditions clearly established the positions of hydrogen atoms.[8] This study identified sites of protonation on the oxygen of serine and on the nitrogen of the histidine ring as depicted on the top left in Figure 1. This study established that, in the absence of a peptide substrate, the serine OH group and the imidazole ring of histidine remain neutral with the OH group hydrogen bonded to the basic nitrogen of histidine. In addition, the N−H group on the imidazole ring bonds to a negatively charged carboxylate group on the aspartate residue. During catalytic cleavage of a peptide substrate, proton transfer from serine to histidine generates a histidinium cation that remains bonded to the nucleophilic oxygen of serine in the resulting tetrahedral intermediate, as shown in the bottom left in Figure 1. While the hydrogen-bonding interactions have been characterized thoroughly in the crystal structure of proteinase K for the His and Ser residues in the neutral state, hydrogen-bonding interactions between the charged residues in the active site have yet to be examined by high-resolution X-ray diffraction. Studies of small model systems with charged residues have been carried out.[9−14] For example, structural studies of model systems for the Asp···His diad by Zimmerman and Gandour have revealed that imidazole forms a stronger hydrogen bond when bonded to a carboxylate group in a syn orientation compared to an anti orientation.[15−17] This model system suggests that a carboxylate group causes a larger increase in the pK_a of a proximate imidazole when bonded in the syn orientation, making the imidazole a stronger hydrogen-bonding acceptor. An earlier study of a model system featuring a carboxylate···imidazolium couple bonded to a molecule of water provided some details of hydrogen bonding between charged components in the

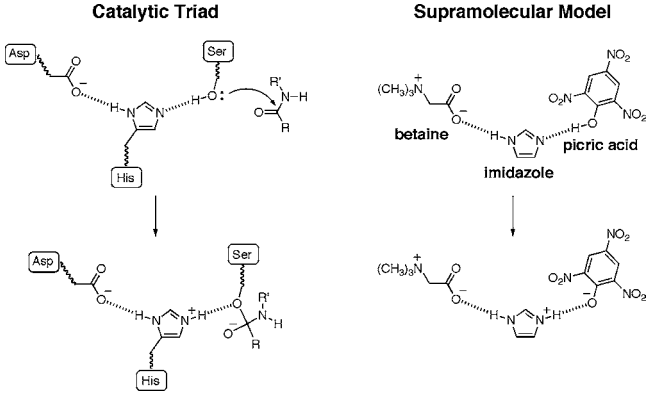

Figure 1. The catalytic triad of Asp···His···Ser residues present in the active site of serine proteases (left). A supramolecular model of the catalytic triad composed of betaine, imidazole and picric acid (right).

* To whom correspondence should be addressed. E-mail: jcm@wpi.edu.

10.1021/cg050092+ CCC: $30.25
Published on Web 07/26/2005

Asp···His···Ser triad.[18] A significant limitation for modeling hydrogen bonding in most of these model systems arises because two or more residues are joined covalently on backbones that are rigid or that have limited conformational mobility. Consequently, conformation constraints can influence the orientation and geometry of the hydrogen-bonding interactions significantly.

The present study describes a supramolecular model of the Asp···His···Ser catalytic triad with charged residues. The molecular components for each residue have no covalent tethers and thus are able to form hydrogen-bonding interactions with idealized geometry that are free of conformational restrictions. This strategy utilizes hydrogen bonding as the primary driving force for the assembly of three different molecular components into an arrangement with intermolecular connectivity similar to that in the Asp···His···Ser triad. An important advantage of this approach is that the individual hydrogen-bonding interactions can form selectively and reflect the energetically most favorable interactions between the three different functional groups. Cocrystallization of two different molecules that have complementary sets of functional groups for hydrogen bonding is well established as a method to form binary cocrystals with two components.[19−44] Etter has shown that cocrystallization is more likely to occur between molecules that contain more than one set of complementary hydrogen-bonding donors and acceptors when the best donor and acceptor reside on different molecules.[19] The concept of matching the best donor and acceptor can be extended to systems with three different components. In principle, a ternary complex should form when the best donor (D_1) and acceptor (A_1) reside on the first and second components and the second best donor (D_2) and acceptor (A_2) reside on the second and third components. Inclusion of two donors and two acceptors on three different molecular components can arise from several different combinations as follows: $A_1\cdots D_1-D_2\cdots A_2$ or $D_1\cdots A_1-A_2\cdots D_2$ or $A_1\cdots D_1-A_2\cdots D_2$ or $D_1\cdots A_1-D_2\cdots A_2$. Aakeröy used this strategy recently to prepare ternary cocrystals that contain isonicotinamide and two different carboxylic acids.[45] Despite this success, formation of ternary cocrystals has been difficult to achieve in practice except in a few select systems. Our model system utilizes betaine, imidazole, and picric acid as molecular components to represent Asp, His, and Ser residues, respectively. This model system is shown on the left in Figure 1. Betaine (*N*,*N*′,*N*″-trimethylglycine) exists permanently as a zwitterion and thus provides a negatively charged carboxylate group that mimics Asp at physiological pH. Imidazole serves as a base that can deprotonate the acidic OH group of picric acid to form an imidazolium picrate ion salt.[46] This charged ion pair mimics the charged histidinium residue bonded to the oxygen of serine in the tetrahedral intermediate in the active site, as shown at the bottom of Figure 1.

In this paper, we describe the supramolecular and crystal structures and intermolecular connectivity of two concomitant polymorphs,[47] **1** and **2**, of cocrystals that contain betaine, imidazolium cations, and picrate anions hydrogen bonded in a 1:2:2 ratio. These five molecular components comprise a supramolecular unit with a fixed structure defined by four strong N−H···O interactions and four strong C−H···O interactions that is common to both polymorphs. We demonstrate that betaine, imidazolium cation and picrate anion selectively self-assemble in the absence of covalent tethers to form a hydrogen-bonded network in both polymorphs that mimics the connectivity present in the Asp···His···Ser triad of serine proteases.

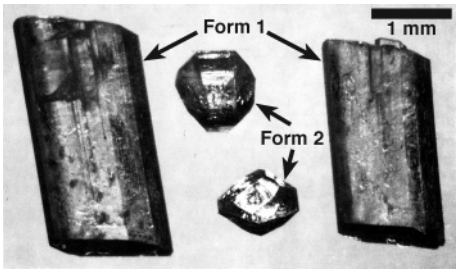

Figure 2. Crystals of polymorph **1** (rhomboids) and polymorph **2** (rounded blocks).

Results

Mixing of equimolar amounts of betaine, imidazole, and picric acid in methanol followed by slow evaporation from dilute solution or cooling of hot concentrated solution gave crystals with two distinct morphologies—rhomboids, **1**, and rounded blocks, **2**—as shown in Figure 2. Crystals of **1** predominated or were the only type of crystal observed in numerous batches of crystals grown using both methods. A few crystals of **2** appeared in solution infrequently and did not appear consistently in different batches of crystals grown under identical conditions. Crystals of **2** always appeared concurrently with crystals of **1**.

Noncrystallographic Characterization of Polymorphs 1 and 2. Infrared spectroscopy was used to verify that proton transfer occurred from picric acid to imidazole and to identify the presence of N−H···O hydrogen bonds. These data complemented the X-ray crystal structures for **1** and **2** by confirming the presence of imidazolium cations and the absence of a phenolic OH group on picric acid. The N−H stretching frequencies of imidazolium cation are sensitive to hydrogen bonding and to the electronic nature of the acceptor atom to which the N−H group is bonded. Neutral imidazole forms self-complementary N−H···N hydrogen bonds in its own crystal structure that give a low-intensity N−H stretching band near 1835 cm^{-1}.[48] After proton transfer, formation of strong N−H···O hydrogen bonds gave a broad band of low intensity near 1940 cm^{-1} in the IR spectra of **1** and **2**. This band appeared in a frequency range similar to that reported in our earlier work for N−H···O stretching bands of imidazolium carboxylate salts[49] and by Johnson and Rumon for pyridinium carboxylate salts.[50]

^{1}H NMR spectra showed that betaine, imidazole, and picric acid were present in a 1:2:2 relative ratio, respectively, in single crystals of **1** and **2**. The presence of imidazole and picric acid in a 1:1 molar ratio in both crystals was expected since those two components are known to form a 1:1 imidazolium picrate salt.[46] Inclusion of 100% molar excess of imidazolium picrate relative to betaine was unexpected, however, considering that equimolar amounts of the three components

were present in solution during crystallization. Crystals of **1** and **2** formed in solution first; crystals of residual betaine formed in solution eventually after several days. To establish whether the ratio of betaine to imidazolium picrate could be controlled in the solid state, cocrystallization was attempted from a series of solutions in which the molar ratio of imidazole and picric acid remained constant at 1:1 but in which the relative molar ratio of betaine was varied systematically. In each case, crystals of **1** and **2** formed in solution first, followed by crystals of pure betaine or imidazolium picrate, depending on which component was in excess.

Crystals of **1** and **2** showed identical melting behavior with melting ranges from 127 to 132 °C. The two crystals showed different thermal stabilities, however, when analyzed by differential scanning calorimetry (DSC). The DSC trace for a crystal of **1** showed a single endothermic peak between 128 and 132 °C that corresponded to the observed melting of the crystals. The DSC trace for a crystal of **2** showed one endothermic peak between 122 and 127 °C followed immediately by a second endothermic peak between 128 and 132 °C. Upon cooling of the sample to room temperature and then reheating, the DSC trace for the same sample of **2** showed a single endothermic peak between 128 and 132 °C. These data suggest that crystals of **1** and **2** are polymorphs and that **2** undergoes a phase transformation to **1** between 122 and 127 °C just prior to melting. The ^{1}H NMR data showed that solvent was not present in crystals of **2**, confirming that the phase transformation was not caused by loss of included solvent during heating and that **2** was not a pseudopolymorph of **1**.[51]

Crystallographic Characterization of Polymorphs 1 and 2. X-ray data were collected, and the crystal structures of **1** and **2** were solved to confirm that the two crystals were polymorphs and to compare the supramolecular structures and hydrogen bonding. The supramolecular structures and hydrogen-bonding interactions of **1** and **2** are compared in Figures 3 and 4. The crystal packing in **1** and **2** is compared in Figure 5. Crystallographic data and hydrogen-bonding geometries are listed in Tables 1 and 2, respectively.

The asymmetric unit in **1** and **2** consists of one molecule of betaine, two imidazolium cations, and two picrate anions, as indicated by the ^{1}H NMR data. Those five components comprise the supramolecular unit—hereafter referred to as SU—that is nearly identical in structure and intermolecular connectivity in **1** and **2**, as shown in Figure 3. The carboxylate group of betaine connects to two imidazolium cations via N−H···O hydrogen bonds at each of the oxygen atoms (B and C in Figure 3a, and B in Figure 4). The imidazolium cations each are bonded to a phenolate anion via a single N−H···O hydrogen bond at the phenolate oxygen (A and D in Figure 3a). The five components in the SU are joined by four C−H···O hydrogen bonds in addition to the four N−H···O hydrogen bonds. The two phenolate anions are bonded by C−H···O hydrogen bonds between a C−H group on one anion and the nitro group on the second anion (E in Figure 3a). One phenolate anion is bonded to the two imidazolium cations via C−H···O hydrogen bonds between the C−H groups on C2 of the cations and the two oxygen atoms of the same nitro group (F and G in Figure 3a). The C−H group on C2 of one imidazolium cation forms a three-centered

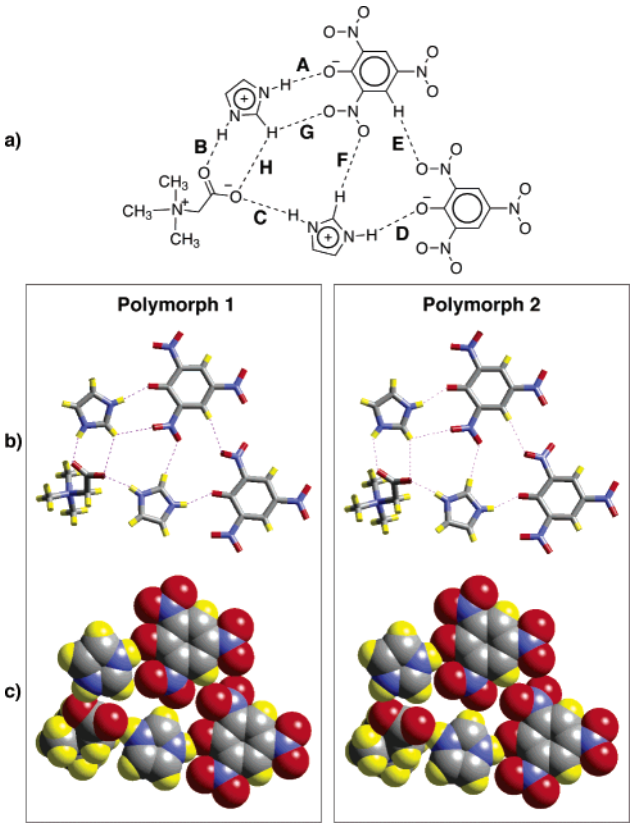

Figure 3. (a) The supramolecular units in polymorphs **1** and **2** contain betaine, imidazolium cations, and picrate anions in a ratio of 1:2:2, respectively, with the five components joined by N−H···O hydrogen bonds (A−D) and C−H···O hydrogen bonds (E−H). (b and c) Stick and space-filling representations that show the similarity in structure of the supramolecular units in **1** and **2**.

C−H···O hydrogen bond by bonding to a second oxygen atom on the carboxylate group of betaine (H in Figure 3a).

Crystal packing of the SUs differs significantly in **1** and **2** despite the similar structures of the SUs themselves, as shown in Figure 5. The SUs in **1** pack in the space group $P\bar{1}$ with $Z = 2$, while those in **2** pack in the space group $C2/c$ with $Z = 8$. The most striking difference in packing between **1** and **2** is that pairs of SUs in **2** form hydrogen-bonded dimers that are not present in the structure of **1**. These dimers arise when N−H groups on two imidazolium cations each form a three-centered hydrogen bond (B and I in Figure 4) with oxygen atoms on two molecules of betaine in adjacent SUs. Two of these three-centered hydrogen bonds generate a hydrogen-bonded ring with graph set notation $R_2^2(4)$ in which the SUs are related by 2-fold rotational symmetry.[53,54] The imidazolium cations in crystals of **1** all form two-centered N−H···O hydrogen bonds with no hydrogen-bonding interactions between SUs. Adjacent SUs in **1** are related by inversion symmetry.

Discussion

Assembly of Three Molecular Components through N−H···O Hydrogen Bonding. Our strategy required that three different molecular components self-assemble to mimic the hydrogen-bonding interactions between residues in the Asp···His···Ser triad. We chose

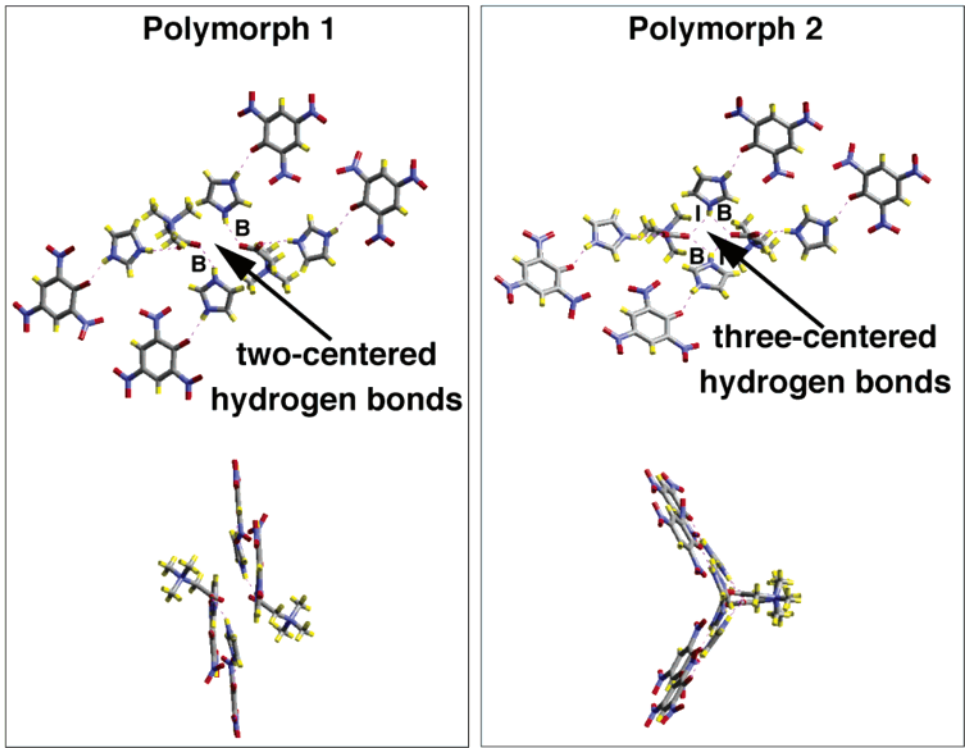

Figure 4. Two different views of a pair of supramolecular units present in **1** and **2**. The central imidazolium cations in **1** form two-centered hydrogen bonds (B), while those in **2** form three-centered hydrogen bonds (B and I).

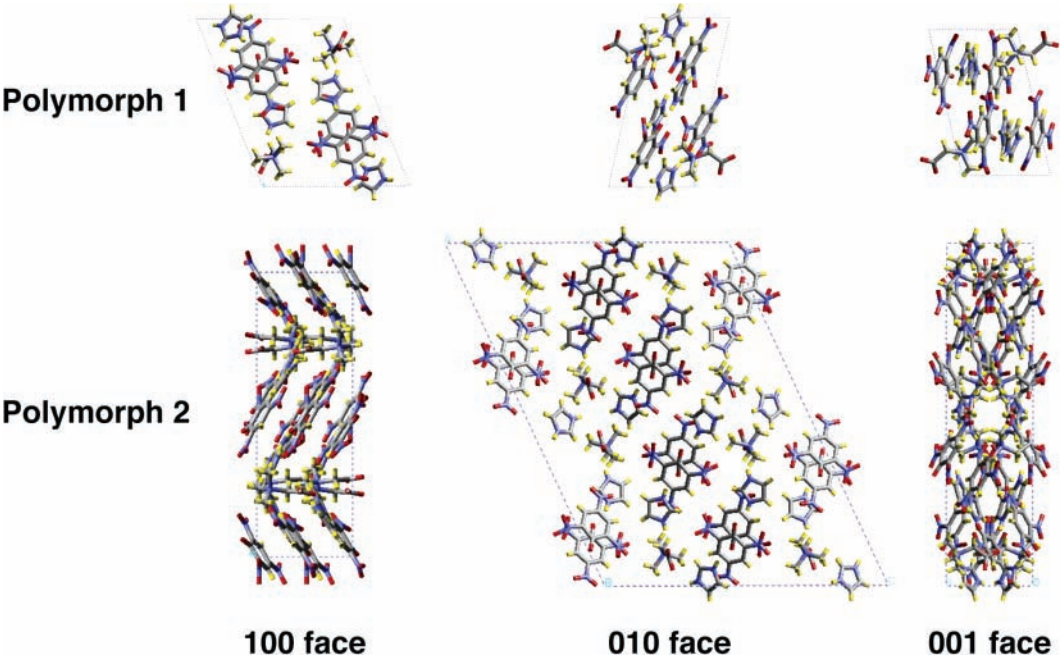

Figure 5. Crystal packing in the unit cells of polymorphs **1** (top) and **2** (bottom) viewed from different directions.

to cocrystallize three components, betaine, imidazole, and picric acid, that were not linked covalently to avoid conformational restrictions imposed by covalent tethers. This model system is a not a true ternary cocrystal consisting of neutral components, however, because proton transfer occurs from picric acid to imidazole. The resulting imidazolium picrate salt ensures that one N−H donor is paired with the picrate acceptor by forming a charge-assisted $^{+}$N−H···O^{-} hydrogen bond between the two ions. The remaining N−H donor on the imidazolium cation forms a second charge-assisted $^{+}$N−H···O^{-} hydrogen bond with the carboxylate acceptor on a molecule of betaine. We expected one of several possible hydrogen-bonding arrangements to occur depending on whether the best acceptor is betaine or picrate anion. For example, picrate anion will make chains of alternating cations and anions (i.e., ···A_1··· D_1−D_2···A_1···D_1−D_2···) if it is the best acceptor (A_1) by forming hydrogen bonds with both imidazolium N−H donors (D_1 and D_2). This arrangement occurs in the crystal structure of imidazolium picrate.[46] Betaine would not participate in hydrogen bonding in this

Table 1. Collection of X-ray and Neutron Diffraction Data and Refinement of Crystals Structures for Polymorphs 1 and 2

polymorph	**1**	**2**	**2**[a]
method of diffraction	X-ray	X-ray	neutron
formula	$C_{23}H_{25}N_{11}O_{16}$	$C_{23}H_{25}N_{11}O_{16}$	$C_{23}H_{25}N_{11}O_{16}$
formula weight	711.54	711.54	711.54
temperature (K)	297(2)	295(2)	28(1)
crystal system	triclinic	monoclinic	monoclinic
space group	*P*$\bar{1}$	*C*2/*c*	*C*2/*c*
a (Å)	7.604(8)	33.908(7)	33.536(5)
b (Å)	13.403(4)	7.765(1)	7.636(1)
c (Å)	16.366(8)	25.389(6)	25.066(4)
α (°)	109.47(3)	90	90
β (°)	98.36(6)	114.74(2)	114.90(1)
γ (°)	99.17(5)	90	90
V (Å^3)	1516.6(18)	6071(2)	5822(2)
Z	2	8	8
D_{calc} (g/cm^3)	1.56	1.56	
R/wR^2 (obs data)	0.0778/0.2249	0.0506/0.1322	0.035/0.046
R/wR^2 (all data)	0.1103/0.2537	0.0735/0.1458	
goodness of fit on F^2	1.056	1.033	1.13

[a] Data published previously that are shown for comparison.[52]

Table 2. Geometry of Hydrogen Bonds

hydrogen bond	interaction	**1** (X-ray)	**2** (X-ray)	**2** (neutron)[a]
A	N···O (Å)	2.668(4)	2.676(3)	2.676(2)
	H···O (Å)	1.82	1.85	1.678(4)
	N−H···O (°)	166.9	161.3	157.2(4)
B	N···O (Å)	2.703(5)	2.894(3)	2.827(2)
	H···O (Å)	1.91	2.23	2.091(1)
	N−H···O (°)	152.4	133.9	127.00(6)
C	N···O (Å)	2.761(4)	2.696(3)	2.685(2)
	H···O (Å)	1.93	1.85	1.637(4)
	N−H···O (°)	161.8	167.2	171.2(4)
D	N···O (Å)	2.611(4)	2.622(3)	2.613(2)
	H···O (Å)	1.79	1.81	1.631(4)
	N−H···O (°)	158.0	157.3	154.5(5)
E	C···O (Å)	3.351(6)	3.336(3)	3.240(2)
	H···O (Å)	2.45	2.43	2.166(1)
	C−H···O (°)	163.3	165.5	166.97(4)
F	C···O (Å)	3.319(6)	3.288(4)	3.296(2)
	H···O (Å)	2.43	2.37	2.208(1)
	C−H···O (°)	160.5	171.3	174.43(4)
G	C···O (Å)	3.033(6)	3.096(4)	3.014(1)
	H···O (Å)	2.50	2.66	2.570(1)
	C−H···O (°)	116.4	109.7	103.77(5)
H	C···O (Å)	3.310(5)	3.433(3)	3.384(2)
	H···O (Å)	2.71	2.65	2.459(2)
	C−H···O (°)	123.3	142.5	143.24(5)
I	N···O (Å)	-	2.968(4)	2.902(1)
	H···O (Å)	-	2.23	2.015(1)
	N−H···O (°)	-	144.4	143.47(6)

[a] Data published previously that are shown for comparison.[52]

situation and would crystallize separately. We observed this behavior in a series of cocrystallization experiments from solutions that contained betaine, imidazole, and para-substituted phenols.[55] A fair comparison cannot be made between the hydrogen-bonding strengths of the components in that series and those in polymorphs **1** and **2**, however, because the pK_a values of the para-substituted phenols were sufficiently high that proton-transfer did not occur to form imidazolium phenolate salts. Alternatively, if betaine is the best acceptor, it will form at least one hydrogen bond with an imidazolium N−H donor (D_1). The remaining N−H donor (D_2) could form a hydrogen bond either to another molecule of betaine ($A_1\cdots D_1-D_2\cdots A_1$) or to a picrate anion ($A_1\cdots D_1-D_2\cdots A_2$). The occurrence of the latter arrangement in **1** and **2** indicates that the carboxylate group of betaine is a better acceptor than the phenolate group of picrate anion. This finding is consistent with an earlier investigation of **2** that we carried out using neutron diffraction (see paragraph that follows) that showed that the imidazolium cation forms a stronger N−H···O hydrogen bond to the carboxylate group (C in Figure 3a) than to the phenolate group.[52] Moreover, the carboxylate group on one molecule of betaine acts as an acceptor twice by forming hydrogen bonds with two different imidazolium cations; each of the phenolate groups forms just one N−H···O hydrogen bond. These results are not surprising considering that carboxylate groups are known to be strong acceptors that often bond to multiple donors. For example, Görbitz has shown that carboxylate groups typically form hydrogen bonds with three different donors and can accommodate as many as five in the solid state.[56,57] The carboxylate groups in **1** and **2** form two and three N−H···O hydrogen bonds, respectively.

We previously reported a neutron diffraction study of hydrogen bonding in the crystal structure of polymorph **2** of the 1:2:2 betaine imidazolium picrate cocrystal.[52] That study was carried out to investigate the electronic structure in our model system of the Asp···His···Ser catalytic triad and to establish whether short N−H···O interactions form low barrier hydrogen bonds with partial covalent character. Formation of a low barrier hydrogen bond can account for many of the special physicochemical observations found in some enzymes, but it is still a matter of controversy whether a low barrier hydrogen bond contributes significantly toward the catalytic activity of the triad.[58] Electron densities calculated from the neutron diffraction data established that the N−H···O bonds in **2** are predominantly electrostatic in nature, but that the short, strong bond between the imidazolium cation and carboxylate group on betaine (C in Figure 3a) is not entirely electrostatic in character. The atomic charges in hydrogen bond C are larger than in the other hydrogen bonds, although still smaller than in typical low barrier hydrogen bonds. It is notable that the hydrogen bond that best mimics the Asp···His structure also is closest to being a low barrier hydrogen bond. This study showed that short, strong hydrogen bonds do not have to be low-barrier, but that if low-barrier hydrogen bonds in fact are present in enzymes, then other factors not active in the model complex must account for the higher charge migration in bond C. This previous work focused primarily on N−H···O interactions in polymorph **2**.

Gandour has argued that the syn lone pairs of electrons on carboxylate groups may be as much as 10^4 times more basic than the corresponding anti lone pairs of electrons.[17] Insofar as basicity reflects the strength of a lone pair to accept a hydrogen bond (and thus the length of the resulting bond), polymorph **1** provides a unique opportunity to examine the stereoelectronic effect of bonding at syn and anti lone pairs of electrons on the carboxylate group of betaine (Figure 6). A comparison of the geometry of syn hydrogen bond B [N···O = 2.703(5) Å, H···O = 1.91 Å, N−H···O = 152.4°] and anti hydrogen bond C [N···O = 2.761(4) Å, H···O = 1.93 Å, N−H···O = 161.8°] shown in Figure 3 and Table 2 reveals that the syn hydrogen bond is shorter than

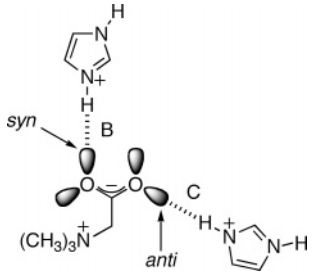

Figure 6. Imidazolium cations form hydrogen bonds at both syn and anti lone pairs of electrons on the carboxylate group of betaine in **1** and **2**.

the anti hydrogen bond by 0.058 Å. This difference in bond length is greater than 3 times the estimated standard deviation of the N···O bond distance (3σ = 0.015 and 0.012 Å for the syn and anti hydrogen bonds, respectively) and, therefore, is experimentally significant. This observation is consistent with Gandour's argument. A similar comparison of the geometry of the respective syn and anti hydrogen bonds in polymorph **2** [syn hydrogen bond B [N···O = 2.894(3) Å, H···O = 2.23 Å, N−H···O = 133.9°] and anti hydrogen bond C [N···O = 2.696(3) Å, H···O = 1.85 Å, N−H···O = 167.2°] shows just the opposite; the syn hydrogen bond is longer than the anti hydrogen bond by 0.198 Å. Again, this difference in bond length is greater than 3 times the estimated standard deviation of the N···O bond distance (3σ = 0.009 Å for both the syn and anti hydrogen bonds) and, therefore, is experimentally significant. Although the data suggest that the anti hydrogen bond is stronger in **2**, comparing syn and anti interactions in **1** and **2** is not meaningful because the N−H donor in **2** forms a three-centered hydrogen bond, while the N−H donor in **1** forms a normal a two-centered hydrogen bond (Figure 4).

Effect of C−H···O Hydrogen Bonds on Supramolecular Structure. The fact that the SU has the same composition and nearly identical structure in **1** and **2** (Figure 3) is remarkable considering that the packing arrangements of the SUs differ so greatly in the two polymorphs. This similarity indicates that factors other than strong N−H···O hydrogen bonds play a significant role both in selecting for the same molecular components (i.e., betaine, imidazole and picric acid in a 1:2:2 ratio, respectively) and in determining their relative positions and orientations within the SU. Taylor and Kennard have demonstrated that nitrogen atoms in aromatic rings and nitro groups substituted on aromatic compounds decrease electron density inductively at adjacent C−H protons, thereby enhancing the ability of C−H protons to act as hydrogen-bonding donors.[59] We reported previously that imidazolium cations function as multidentate proton donors in imidazolium carboxylate salts by forming up to three C−H···O hydrogen bonds in addition to two N−H···O hydrogen bonds.[49] We expected that the formation of C−H···O hydrogen bonds would influence molecular packing in **1** and **2** because all carbon atoms in the imidazolium cations are bonded to one or two nitrogen atoms, all aromatic C−H groups on picric acid are adjacent to two nitro groups, and the components in **1** and **2** have a high ratio of oxygen to C−H groups. Consequently, we examined intermolecular contacts between C−H protons and neighboring oxygen atoms that had H···O distances less than 2.72 Å [the sum of the van der Waals radii for hydrogen (1.20 Å)[60] and oxygen (1.52 Å)[60]] and C−H···O angles greater than 90°. The H···O distances were examined rather than the C···O distances because C−H···O interactions frequently are bent instead of linear. The five molecular components in the SU are joined by four C−H···O hydrogen bonds in addition to four N−H···O hydrogen bonds, as shown in Figure 3. The hydrogen-bonding data in Table 2 show that C−H···O interactions E and F, in particular, are short contacts (e.g., H···O = 2.43 and 2.37 Å and C−H···O = 165.5° and 171.3°, respectively, in **2**) that link one picrate anion to a second picrate anion and to an imidazolium cation in the same SU. This pattern of connectivity suggests that the stronger N−H···O interactions join the five components together serially into a chain, while the weaker C−H···O hydrogen bonds cross-link the components and tie the ends of the chain together to create a more rigid SU with well-defined shape and structure. Although we have not examined the energetic contributions of individual hydrogen bonds to the overall stability of the SU using computations, the eight hydrogen bonds provide a significant enthalpic contribution to the stability of the SU by ensuring that each component is bonded to others by at least two and as many as four hydrogen bonds. The periphery of the SU displays identical topology and exposed functionality along the edges and surfaces in polymorphs **1** and **2**. The occurrence of this unusual building block in two different crystalline forms shows that this hydrogen-bonded arrangement is not fortuitous but rather generates a robust and persistent SU. While it is not clear why betaine, imidazole, and picric acid crystallize selectively in a 1:2:2 ratio from solutions that contain equimolar amounts of the three components, the fact that the ratio remains constant at 1:2:2 in both polymorphs indicates that a SU composed of those five molecules is favored over SUs with different composition during nucleation of crystals.

It should be noted that the H···O distances for C−H···O hydrogen bonds in **1** and **2** were determined based on X-ray data in which all of the C−H hydrogen atoms on imidazolium cations were placed at idealized positions 0.93 Å from the carbon atom to which they were bonded during refinement of the crystal structures. Hydrogen atoms were treated in this manner because the positions of hydrogen atoms generally cannot be determined accurately with X-ray diffraction, especially with data collected at room temperature. Data on C−H bond lengths from studies using neutron diffraction suggest that C−H bond lengths should be fixed at 1.10 Å for organic compounds.[61] Considering that the SHELX-97 refinement program[62] assigns the shorter distance of 0.93 Å for C−H bond lengths on aromatic compounds, the H···O distances in **1** and **2** are at least 0.17 Å shorter than the values from the X-ray data. The neutron diffraction data in Table 2 for the crystal structure of **2** (reported previously)[52] confirms that the H···O distances for contacts E and F, in fact, are shorter (and thus stronger) by 0.26 and 0.16 Å, respectively, than the corresponding contacts in the crystal structure of **2** determined using X-ray diffraction.

Comparison of Polymorphs 1 and 2. While the intermolecular connectivity within SUs is identical in

1 and **2**, the connectivity between SUs differs in the two structures as described in the results section. The absence or presence of hydrogen-bonded dimers between SUs depends on whether the N−H donor on the imidazolium cation participates in forming a two-centered hydrogen bond with a single betaine molecule in the same SU (**1**) or a three-centered hydrogen bond with betaine molecules in two different SUs (**2**). As shown in Figure 4, the imidazolium cation is able to turn slightly and reorient the N−H group to go from a two-centered to a three-centered hydrogen bond while minimally perturbing the structure of the SU. This behavior in a supramolecular structure is analogous to a small conformational change in a covalently bonded molecular structure; therefore, it follows that **1** and **2** can be considered as supramolecular conformational polymorphs. The endothermic peak at 122−127 °C in the DSC trace of **2** just prior to melting at 128−132 °C (melting point of **1**) suggests that crystals of **2** undergo a phase transformation to **1**, although we have not confirmed experimentally that **1** forms prior to melting.

Conclusions

We demonstrate that the strategy of matching complementary sets of hydrogen-bonding donors and acceptors can be utilized to prepare ternary cocrystals composed of three different molecular components—betaine, imidazole, and picric acid. Furthermore, we show that these molecules selectively self-assemble from solution in the absence of covalent tethers to form a hydrogen-bonded network that mimics the connectivity present in the Asp···His···Ser catalytic triad of serine proteases. The fact that these components assemble in this manner suggests that hydrogen bonding plays a significant role in governing the assembly of the Asp, His, and Ser residues in the active site of serine proteases. Although the energies of the individual hydrogen bonds between the residues are small relative to the energies associated with the folding of proteins, nonetheless, the formation of this catalytic triad is more than random chance. In addition to strong N−H···O hydrogen bonds that define the connectivity between betaine, imidazole, and picric acid, C−H···O interactions play a significant role in determining the 1:2:2 composition, structure, and orientation of the five components within the supramolecular unit. Although the contribution of C−H···O hydrogen bonds toward the stability of proteins such as serine proteases is not well understood, these weak interactions may play a more significant role than previously thought.

Experimental Section

General. All chemicals and solvents were purchased from Aldrich or ACROS and used as received without further purification. Spectroscopic grade solvents were used for all experiments involving crystallization. Melting points were determined on a Fisher-Johns apparatus and are uncorrected. Infrared spectra were recorded on a Nicolet 5DXB FTIR spectrometer from Nujol mulls unless otherwise indicated and are reported in cm^{-1}. ^1H NMR spectra were recorded on an IBM NR200AF spectrometer from solutions in DMSO-d_6, and chemical shifts are reported in parts per million (δ) from TMS. DSC traces were recorded on a Mettler FP800 instrument.

Preparation of Cocrystals of Polymorphs 1 and 2. Method 1: Cocrystals of **1** and **2** were prepared by dissolving equimolar amounts (~1 mmol) of betaine, imidazole, and picric acid in approximately 30 mL of methanol at room temperature. The solutions were covered partially with a watch glass and allowed to evaporate slowly at room temperature. Crystals appeared after several days. Different batches of crystals grown using this method contained only crystals of **1** or crystals of **1** predominantly with a few crystals of **2**. Method 2: Cocrystals of **1** and **2** were prepared from hot concentrated solution by adding methanol to the solutes with mild heating (40−50 °C) and stirring until dissolution was complete. The hot solution was removed from the hot plate and covered. Cocrystals appeared either as the solution was cooling to room temperature or soon after the solution finished cooling. Cocrystals were removed from solution and blotted dry on filter paper to prevent epitaxial growth from residual solution. Different batches of crystals grown using this method contained crystals of **1** predominantly with a few crystals of **2**. X-ray data for the crystal structures of **1** and **2** was obtained from crystals grown using this method.

1:2:2 Betaine Imidazolium Picrate (Polymorph 1). Dark yellow rhomboids (Figure 2); mp 127−132 °C; IR (Nujol) 3135, 3080, 3044, (broad, weak), 1634, 1616, 1566, 1514, 1495, 1435, 1396, 1341, 1317, 1277, 1163, 1099, 1080, 1055, 984, 939, 910, 860, 787, 762, 744, 712, 640 and 632 cm^{-1}; ^1H NMR (200 MHz, DMSO-d_6) δ 14.25 (br s, 4 H), 8.77 (s, 2 H), 8.61 (s, 4 H), 7.55 (s, 4 H), 4.01 (s, 2 H) and 3.21 (s, 9 H) ppm. DSC (2.5 °C/min, 100 mV) endotherm 128−132 °C.

1:2:2 Betaine Imidazolium Picrate (Polymorph 2). Light yellow chunky blocks (Figure 2); mp 127−132 °C; IR (Nujol) 3147, 3082, 3043, 1940 (broad, weak), 1634, 1615, 1568, 1547, 1518, 1493, 1435, 1323, 1281, 1167, 1084, 1055, 918, 903, 843, 791, 746, 713 and 630 cm^{-1}; ^1H NMR (200 MHz, DMSO-d_6) δ 13.90 (br s, 4 H), 8.87 (s, 2 H), 8.61 (s, 4 H), 7.58 (s, 4 H), 3.98 (s, 2 H) and 3.22 (s, 9 H) ppm. DSC (2.5 °C/min, 100 mV) endotherm 122−127 °C, endotherm 128−132 °C.

Determination of X-ray Crystal Structures. X-ray data for the crystal structures of **1** and **2** were obtained on an Enraf-Nonius CAD4 diffractometer with graphite monochromated Mo K_α radiation ($\lambda = 0.71069$) using the ω-2θ technique. Lattice parameters were obtained from least-squares analysis of 25 reflections. SHELXS-97 and SHELXL-97 software were used to solve and refine structures.[62] Refinement was based on F^2. All non-hydrogen atoms were refined anisotropically. All hydrogen atoms on heteroatoms were first located in difference maps and then fixed in calculated positions and refined isotropically with thermal parameters based upon the corresponding attached heteroatoms [U(H) = 1.2 Ueq (C)]. The remaining hydrogen atoms were fixed in calculated positions and refined isotropically with thermal parameters based upon the corresponding attached carbon atoms [U(H) = 1.2 Ueq (C)].

Acknowledgment. We thank Doyle Britton at the University of Minnesota for collecting X-ray data for **1** and **2**. We also thank Jacob Overgaard, Birgit Schiøtt, Finn K. Larsen, Arthur J. Schultz, and Bo B. Iversen for determining the neutron structure of **2** previously.[52]

Supporting Information Available: X-ray data for the crystal structures of **1** and **2** is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Blow, D. M. *Acc. Chem. Res.* **1976**, *9*, 145−152.
(2) Bone, R.; Shenvi, A. B.; Kettner, C.; Agard, D. A. *Biochemistry* **1987**, *26*, 7609−7614.
(3) Lange, G.; Betzel, C.; Branner, S.; Wilson, K. S. *Eur. J. Biochem.* **1994**, *224*, 507−518.
(4) Siezen, R. J.; Willem, M. V.; Leunissen, J. A. M.; Dikstra, B. W. *Protein Eng.* **1991**, *4*, 719−737.
(5) Tsukuda, H.; Blow, D. *J. Mol. Biol.* **1985**, *184*, 703−711.
(6) Carter, P.; Wells, J. A. *Nature* **1988**, *332*, 564−568.
(7) Russel, A. J.; Fersht, A. R. *Nature* **1987**, *328*, 496−500.
(8) Betzel, C.; Gourinath, S.; Kumar, P.; Kaur, P.; Perbandt, M.; Eschenburg, S.; Singh, T. P. *Biochemistry* **2001**, *40*, 3080−3088.

(9) Rogers, G. A.; Bruice, C. T. *J. Am. Chem. Soc.* **1974**, *96*, 2473−2581.
(10) Komiyama, M.; Bender, M. L. *Bioorg. Chem.* **1977**, *6*, 13−20.
(11) Roberts, J. D.; Yu, C.; Fanagan, C.; Birdseye, T. R. *J. Am. Chem. Soc.* **1982**, *104*, 3945−3949.
(12) Bruice, C. T.; Sturtevant, J. M. *J. Am. Chem. Soc.* **1959**, *81*, 2860−2870.
(13) D'Souza, V. T.; Bender, M. L. *Acc. Chem. Res.* **1987**, *20*, 146−152.
(14) Mallick, I. M.; D'Souza, V. T.; Yamaguchi, M.; Lee, J.; Chalabi, P.; Gadwood, R. C.; Bender, M. L. *J. Am. Chem. Soc.* **1984**, *106*, 7252−7254.
(15) Zimmerman, S. C.; Cramer, K. D. *J. Am. Chem. Soc.* **1988**, *110*, 5906−5908.
(16) Cramer, K. D.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1990**, *112*, 3680−3682.
(17) Gandour, R. D. *Bioorg. Chem.* **1981**, *10*, 169−176.
(18) Czugler, M.; Angyan, J. G.; Naray-Szabo, G.; Weber, E. *J. Am. Chem. Soc.* **1986**, *108*, 1275−1281.
(19) Etter, M. C. *J. Phys. Chem.* **1991**, *95*, 4601−4610.
(20) Palmore, G. T. R.; MacDonald, J. C. In *The Amide Linkage: Structural Significance in Chemistry, Biochemistry and Materials Science*; Greenberg, A., Breneman, C. M., Liebman, J. F., Eds.; John Wiley and Sons: New York, 2000.
(21) Chin, D. N.; Zerkowski, J. A.; MacDonald, J. C.; Whitesides, G. M. In *Organised Molecular Assemblies in the Solid State*; Whitesell, J. K., Ed.; John Wiley and Sons: New York, 1999.
(22) Zerkowski, J. A.; MacDonald, J. C.; Whitesides, G. M. *Chem. Mater.* **1997**, *9*, 1933−1941.
(23) Palacin, S.; Chin, D. N.; Simanek, E. E.; MacDonald, J. C.; Whitesides, G. M.; McBride, M. T.; Palmore, G. T. R. *J. Am. Chem. Soc.* **1997**, *119*, 11807−11816.
(24) Schwiebert, K. E.; Chin, D. N.; MacDonald, J. C.; Whitesides, G. M. *J. Am. Chem. Soc.* **1996**, *118*, 4018−4029.
(25) Zerkowski, J. A.; MacDonald, J. C.; Whitesides, G. M. *Chem. Mater.* **1994**, *6*, 1250−1257.
(26) Zerkowski, J. A.; MacDonald, J. C.; Seto, C. T.; Wierda, D. A.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994**, *116*, 2382−2391.
(27) MacDonald, J. C.; Whitesides, G. M. *Chem. Rev.* **1994**, *94*, 2383−2420.
(28) Aakeröy, C. B.; Seddon, K. R. *Chem. Soc. Rev.* **1993**.
(29) Chang, Y.-L.; West, M.-A.; Fowler, F. W.; Lauher, J. W. *J. Am. Chem. Soc.* **1993**, *115*, 5991−6000.
(30) Coe, S.; Kane, J. J.; Nguyen, T. L.; Toledo, L. M.; Wininger, E.; Fowler, F. W.; Lauher, J. W. *J. Am. Chem. Soc.* **1997**, *119*, 86−93.
(31) Ducharme, Y.; Wuest, J. D. *J. Org. Chem.* **1988**, *53*, 5787−5789.
(32) Simard, M.; Su, D.; Wuest, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 4696−4698.
(33) Lehn, J.-M.; Mascal, M.; DeCian, A.; Fischer, J. *J. Chem. Soc., Perkin Trans. 2* **1992**, 461−467.
(34) Lehn, J.-M.; Mascal, M.; DeCian, A.; Fischer, J. *J. Chem. Soc., Chem. Commun.* **1990**, 479−481.
(35) Moulton, B.; Zaworotko, M. J. *Chem. Rev.* **2001**, *101*, 1629−1658.
(36) Remenar, J. F.; Morissette, S. L.; Peterson, M. L.; Moulton, B.; MacPhee, J. M.; Guzman, H. R.; Almarsson, O. *J. Am. Chem. Soc.* **2003**, *125*, 8456−8457.
(37) Fleishman, S. G.; Kuduva, S. S.; McMahon, J. A.; Moulton, B.; Bailey Walsh, R. D.; Rodriguez-Hornedo, N.; Zaworotko, M. J. *Cryst. Growth Des.* **2003**, *3*, 909−919.
(38) MacGillivray, L. R.; Papaefstathiou, G. S.; Reid, J. L.; Ripmeester, J. A. *Cryst. Growth Des.* **2001**, *1*, 373−375.
(39) Hamilton, T. D.; Papaefstathiou, G. S.; MacGillivray, L. R. *J. Am. Chem. Soc.* **2002**, *124*, 11606−11607.
(40) Fiscic, T.; Drab, D. M.; MacGillivray, L. R. *Org. Lett.* **2004**, *6*, 4647−4650.
(41) Papaefstathiou, G. S.; Zhong, Z.; Geng, L.; MacGillivray, L. R. *J. Am. Chem. Soc.* **2004**, *126*, 9158−9159.
(42) Swift, J. A.; Ward, M. D. *Chem. Mater.* **2000**, *12*, 1501−1504.
(43) Swift, J. A.; Pivovar, A. M.; Reynolds, A. M.; Ward, M. D. *J. Am. Chem. Soc.* **1998**, *120*, 5887−5894.
(44) Holman, K. T.; Pivovar, A. M.; Swift, J. A.; Ward, M. D. *Acc. Chem. Res.* **2001**, *34*, 107−118.
(45) Aakeröy, C. B.; Beatty, A. M.; Helfrich, B. A. *Angew. Chem., Int. Ed.* **2001**, *40*, 3240−3242.
(46) Soriano-Garcia, M.; Schatz-Levine, M.; Toscano, R. A.; Iribe, R. V. *Acta Crystallogr. C* **1990**, *46*, 1556−1558.
(47) Bernstein, J.; Davey, R. J.; Henck, J.-O. *Angew. Chem., Int. Ed.* **1999**, *38*, 3441−3461.
(48) McMullan, R. K.; Epstein, J.; Ruble, J. R.; Craven, B. M. *Acta Crystallogr. B* **1979**, *35*, 688−691.
(49) MacDonald, J. C.; Dorrestein, P. C.; Pilley, M. M. *Cryst. Growth Des.* **2001**, *1*, 29−38.
(50) Johnson, S. L.; Rumon, K. A. *J. Phys. Chem.* **1965**, *69*, 74−86.
(51) Bernstein, J. *Polymorphism in Molecular Crystals*; Clarendon Press: Oxford, 2002.
(52) Overgaard, J.; Schiøtt, B.; Larsen, F. K.; Schultz, A. J.; MacDonald, J. C.; Iversen, B. B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1239−1242.
(53) Etter, M. C.; MacDonald, J. C.; Bernstein, J. *Acta Crystallogr. B* **1990**, *46*, 256−262.
(54) Bernstein, J.; Davis, R. E.; Shimoni, L.; Chang, N.-L. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1555−1573.
(55) Equimolar amounts of betaine, imidazole, and para-substituted phenols (4-bromophenol, 4-chlorophenol, 4-cyanophenol, 4-methylphenol, 4-nitrophenol, and phenol) were dissolved in a series of solvents (ethyl acetate, chloroform, methanol, or acetone). The resulting solutions were allowed to evaporate slowly until crystals appeared. Crystals were removed from solution and characterized by NMR and IR spectroscopy. In all cases, the phenol cocrystallized with imidazole and betaine crystallized separately.
(56) Görbitz, C. H.; Etter, M. C. *J. Am. Chem. Soc.* **1992**, *114*, 627−631.
(57) Görbitz, C. H.; Etter, M. C. *J. Chem. Soc., Perkin Trans. 2* **1992**, 131−135.
(58) Warshal, A.; Papazyan, A.; Kollman, P. A. *Science* **1995**, *269*, 102−104.
(59) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063−5070.
(60) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441−451.
(61) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: New York, 1991.
(62) Sheldrick, G. M. Göttingen, Germany, 1997.

CG050092+