

# The Structure of Glycine Dihydrate: Implications for the Crystallization of Glycine from Solution and Its Structure in Outer Space

Wenqian Xu, Qiang Zhu,\* and Chunhua (Tony) Hu\*

**Abstract:** Glycine, the simplest amino acid, is also the most polymorphous. Herein, we report the structure determination of a long unknown phase of glycine, which was first reported by Pyne and Suryanarayanan in 2001. To date, this phase has only been prepared at 208 K as nanocrystals within ice. Through computational crystal-structure prediction and powder X-ray diffraction methods, we identified this elusive phase as glycine dihydrate (GDH), representing the first report on the structure of a hydrated glycine structure. The structure of GDH has important implications for the state of glycine in aqueous solution and the mechanisms of glycine crystallization. GDH may also be the form of glycine that comes to Earth from extraterrestrial sources.

Glycine is not only abundant in proteins, but also found in meteorites<sup>[1]</sup> and comets,<sup>[2]</sup> supporting the hypothesis that the origin of life on Earth might be extraterrestrial. As an essential nutrient, glycine is also widely used in polycrystalline form as a pharmaceutical excipient.<sup>[3]</sup> Glycine exhibits the greatest phase diversity among all of the naturally occurring amino acids, and its polymorphs have been prototypical objects for studies of polymorphism and crystallization.<sup>[4]</sup>

The first crystal structure of glycine, referred to as the  $\alpha$ -phase, was determined in 1939 from a crystal grown from aqueous solution ( $P2_1/n$ ,  $a = 5.10 \text{ \AA}$ ,  $b = 11.96 \text{ \AA}$ ,  $c = 5.45 \text{ \AA}$ ,  $\beta = 111.63^\circ$ ,  $V = 304.56 \text{ \AA}^3$ ,  $Z = 4$ ).<sup>[5a]</sup> The structure of  $\beta$ -glycine ( $P2_1$ ,  $a = 5.077 \text{ \AA}$ ,  $b = 6.268 \text{ \AA}$ ,  $c = 5.380 \text{ \AA}$ ,  $\beta = 113.2^\circ$ ,  $V = 157.36 \text{ \AA}^3$ ,  $Z = 2$ ) was subsequently solved by analysis of a crystal grown from an ethanol–water mixture in

1960<sup>[5b]</sup> whereas that of  $\gamma$ -glycine ( $P3_1$  or  $P3_2$ ,  $a = b = 7.037 \text{ \AA}$ ,  $c = 5.483 \text{ \AA}$ ,  $V = 235.14 \text{ \AA}^3$ ,  $Z = 3$ ) was determined from crystals grown in acidic or basic solution the year after.<sup>[5c]</sup>  $\gamma$ -Glycine is the most thermodynamically stable of these phases whereas the  $\beta$ -phase is the least stable one, readily undergoing a single crystal to single crystal transformation into the  $\alpha$ - and  $\gamma$ -phases in humid environments.<sup>[6]</sup> These transformations were shown to proceed in non-topotactic ways, regardless of the similarity between the  $\alpha$ - and  $\beta$ -structures and the dissimilarity of  $\gamma$ -glycine versus the  $\alpha$ - and  $\beta$ -phases.

The effect of pressure on these polymorphs has also been extensively investigated.  $\alpha$ -Glycine persists to a hydrostatic pressure of 23 GPa<sup>[7a]</sup> whereas  $\beta$ -glycine is readily converted into a new phase at 0.8 GPa, which is called the  $\beta'$ -phase by Boldyreva et al.<sup>[7b,c]</sup> and the  $\delta$ -phase by Parsons et al.<sup>[7d]</sup> in a single crystal to single crystal transition.  $\beta'(\delta)$ -Glycine exists in the monoclinic space group  $P2_1/a$  ( $a = 11.156 \text{ \AA}$ ,  $b = 5.864 \text{ \AA}$ ,  $c = 5.342 \text{ \AA}$ ,  $\beta = 125.83^\circ$ ,  $V = 283.32 \text{ \AA}^3$ ,  $Z = 4$  at room temperature and 1.9 GPa). Likewise, at 1.9 GPa,  $\gamma$ -glycine begins to undergo a single crystal to polycrystal transition into another phase,<sup>[7e]</sup> referred to as the  $\delta$ -phase by Boldyreva et al.<sup>[7f]</sup> and the  $\epsilon$ -phase by Parsons et al.<sup>[7d,g]</sup> The crystal structure of the  $\delta(\epsilon)$ -phase was solved using powder X-ray diffraction data. It belongs to the monoclinic space group  $Pn$  ( $a = 4.889 \text{ \AA}$ ,  $b = 5.754 \text{ \AA}$ ,  $c = 5.442 \text{ \AA}$ ,  $\beta = 116.68^\circ$ ,  $V = 136.78 \text{ \AA}^3$ ,  $Z = 2$  at room temperature and 4.3 GPa). Upon decompression to 0.2 GPa, the  $\delta(\epsilon)$ -phase is transformed into a new phase,  $\zeta$ , increasing the number of glycine polymorphs to six. As the  $\zeta$ -phase coexists with the  $\gamma$ - and  $\delta(\epsilon)$ -forms, its structure has yet to be determined.

Furthermore, a new phase of glycine has also been prepared by freezing. Quenching an aqueous glycine solution in liquid nitrogen and keeping the solid at 208 K, Pyne and Suryanarayanan<sup>[8a]</sup> discovered a new phase in 2001, which was later called the X-phase by Boldyreva et al.<sup>[8b]</sup> when they revisited this unknown phase in 2012. Using a modified sample preparation method, they collected synchrotron X-ray powder data and deduced a set of cell parameters ( $a = 6.648 \text{ \AA}$ ,  $b = 25.867 \text{ \AA}$ ,  $c = 5.610 \text{ \AA}$ ,  $\beta = 113.12^\circ$ ). However, the coexistence of the X-phase and ice, and the inadequate quality of the diffraction data, prevented the solution of the structure. Herein, we elucidate the structure of the X-phase by complementing high-quality powder diffraction data with crystal-structure prediction methods. We have called this new phase glycine dihydrate (GDH; Scheme 1).

Instead of pouring a glycine solution onto a cold surface or into liquid nitrogen, we applied a simpler method called flash cooling, which is routinely used in single-crystal diffraction

[\*] W. Xu

X-ray Science Division, Advanced Photon Source  
Argonne National Laboratory  
Argonne, IL 60439 (USA)

Q. Zhu

Department of Physics and Astronomy  
High Pressure Science and Engineering Center  
University of Nevada Las Vegas  
Las Vegas, NV 89154 (USA)

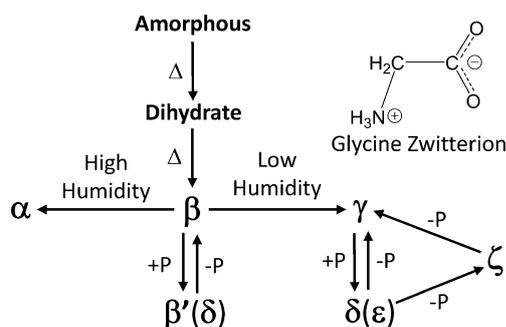
and

Department of Geosciences, Stony Brook University  
Stony Brook, NY 11794 (USA)  
E-mail: qiang.zhu@unlv.edu

C. Hu

Department of Chemistry, New York University  
100 Washington Square East, New York, NY 10003 (USA)  
E-mail: chunhua.hu@nyu.edu

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:  
<http://dx.doi.org/10.1002/anie.201610977>.



**Scheme 1.** The solid phases of glycine and the relations between these phases.  $\Delta$  denotes heating. +P and -P represent increasing and decreasing pressure, respectively. The alternative names are given in parentheses.

experiments. A 20% w/w glycine aqueous solution (ca. 5  $\mu\text{L}$ ) was sealed into a 0.7 mm glass capillary for measurements. A plastic card, shielding the stream of cold nitrogen (controlled at 100 K) from flowing onto the capillary, was rapidly removed. The clear glycine solution in the capillary instantaneously froze and became highly light-scattering. The two-dimensional (2D) diffraction data of the frozen solid collected at 100 K revealed that only diffraction peaks attributed to the ice  $I_h$  phase were observed (see the Supporting Information, Figure S1). Glycine did not crystallize. The earlier reported thermal analysis data<sup>[8a]</sup> had revealed a glass transition event, suggesting the formation of an amorphous solid. The capillary was then warmed up to 209 K,<sup>[9]</sup> and new diffraction peaks emerged. These peaks did not correspond to any known structure of glycine, and thus belonged to the unknown X-phase.<sup>[8a]</sup> At 250 K, the peaks from the unknown phase disappeared, and  $\beta$ -glycine formed. Finally, at 266 K,  $\alpha$ -glycine was observed in the ice–water mixture. Similar experiments were performed with 15, 10, 5, and 2% w/w glycine solutions; these gave similar results while GDH could not be detected with the 1% solution (Figure S2). Flash cooling of the 20% solution to 200 K instead of 100 K did not affect the outcome (Figure S3). In contrast, flash cooling to 209 K and 222 K directly resulted in the X-phase (Figure S4A,B) whereas flash cooling to 225 K yielded the  $\beta$ -phase directly (Figure S4C). As a comparison, a slow-cooling experiment led to the  $\beta$ -phase (Figure S4D); in this experiment, the solution did not freeze until 243 K ( $-30^\circ\text{C}$ ) were reached owing to the strong antifreeze effect of glycine.

The powder data acquired with the single-crystal instrument at the X-ray diffraction facility of New York University (NYU) do not have sufficient angular resolution for peak indexing and structure solution. Therefore, we collected high-resolution diffraction data at the 17-BM beamline of the Advanced Photon Source, Argonne National Laboratory (Figure S5). A 20% w/w glycine aqueous solution in a 0.7 mm glass capillary was first frozen to 173 K by flash cooling and then heated to 208 K and kept at this temperature for two hours while powder data were collected every minute to monitor the phase change. As previously observed in the NYU laboratory, ice was the only crystalline phase present at 173 K. A broad hump centered at  $2\theta = 11.5^\circ$  was also

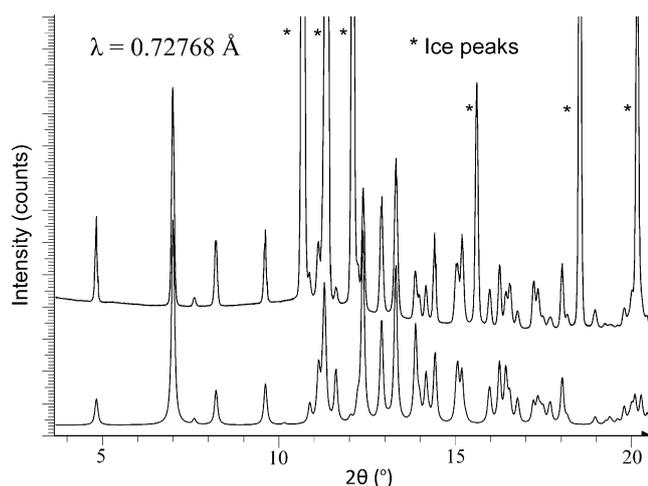
observed underneath the three major peaks of ice (Figure S5). No new peaks emerged during heating. The unknown phase began to appear after 4 min at 208 K. The peak intensities of this phase increased over the next 18 min and then remained unchanged at this temperature. With the growth of the unknown phase, the broad peak at  $11.5^\circ$  decreased in intensity, confirming that the peak is associated with amorphous glycine.

The pattern with the highest intensities for the new peaks was selected for further analysis. The first 20 peaks characterized by the smallest  $2\theta$  angles were manually picked, and their angular positions were accurately determined in the program DASH.<sup>[10a]</sup> These peaks were indexed by the program DICVOL91,<sup>[10b]</sup> which provided a best fit with  $M(20) = 26.1$  and  $F(20) = 92.8$  from a monoclinic unit cell [ $a = 8.964(3) \text{ \AA}$ ,  $b = 8.223(3) \text{ \AA}$ ,  $c = 7.622(3) \text{ \AA}$ ,  $\beta = 104.32(4)^\circ$ ,  $V = 544.42 \text{ \AA}^3$ ]. Systematic absences suggested  $P2_1/c$  as the most probable space group, in which case a dihydrate seemed most likely on account of the unit cell volume of  $544 \text{ \AA}^3$  compared to about  $305 \text{ \AA}^3$  for  $\alpha$ -glycine. Alternative options were the space groups  $P2_1$  and  $Pc$  with  $Z' = 3$  for anhydrous glycine.

The Pawley method did not provide a good fit owing to the ice peaks (Figure S6) but we could still obtain 105 reflections for structure solution. We attempted to solve the crystal structure using the simulated annealing algorithm in DASH with the three space groups  $P2_1/c$ ,  $Pc$ , and  $P2_1$ , but failed to find any appropriate solutions. Therefore, we turned to computational crystal structure prediction (CSP) methods.<sup>[11]</sup>

To solve the unknown crystal structure, we performed a systematic search based on evolutionary algorithms, as implemented in the USPEX code.<sup>[12a,b]</sup> The most significant feature of this approach is that only molecular geometry is used as input. The number of molecules per asymmetric unit and the choice of space groups, specified by the user, define the extent of the crystal structure search. Alternatively, one can predetermine the unit cell if the cell parameters are known. The GULP<sup>[12c]</sup> and DFTB + <sup>[12d]</sup> codes were used to perform the structure relaxations within USPEX (see the Supporting Information for details). We initially conducted a search for anhydrous crystal structures with  $Z' = 1$  at ambient pressure for the 30 most common space groups, which successfully returned all three known glycine polymorphs (namely  $\alpha$ ,  $\beta$ , and  $\gamma$ ). Subsequently, we selected the 100 lowest-energy structures from the initial CSP search and re-optimized their geometries and energy ranking by the VASP code<sup>[12e]</sup> at the optB88 level of theory.<sup>[12f]</sup> Not surprisingly, all polymorphs experimentally observed under ambient conditions appeared at the very first places in the energy ranking (Figure S7).

Using the cell parameters from DICVOL, we performed two independent structure searches, namely  $Z' = 3$  for glycine in the space groups  $P2_1$  and  $Pc$  and  $Z' = 1$  for glycine-2H<sub>2</sub>O (GDH) in  $P2_1/c$ . The fixed-cell searches immediately yielded the lowest-energy structure of GDH, and its simulated powder pattern matched the experimental pattern well (Figure 1), while the search on  $Z' = 3$  for glycine failed to return any meaningful solutions. We then repeated the prediction for the same space group without specifying cell

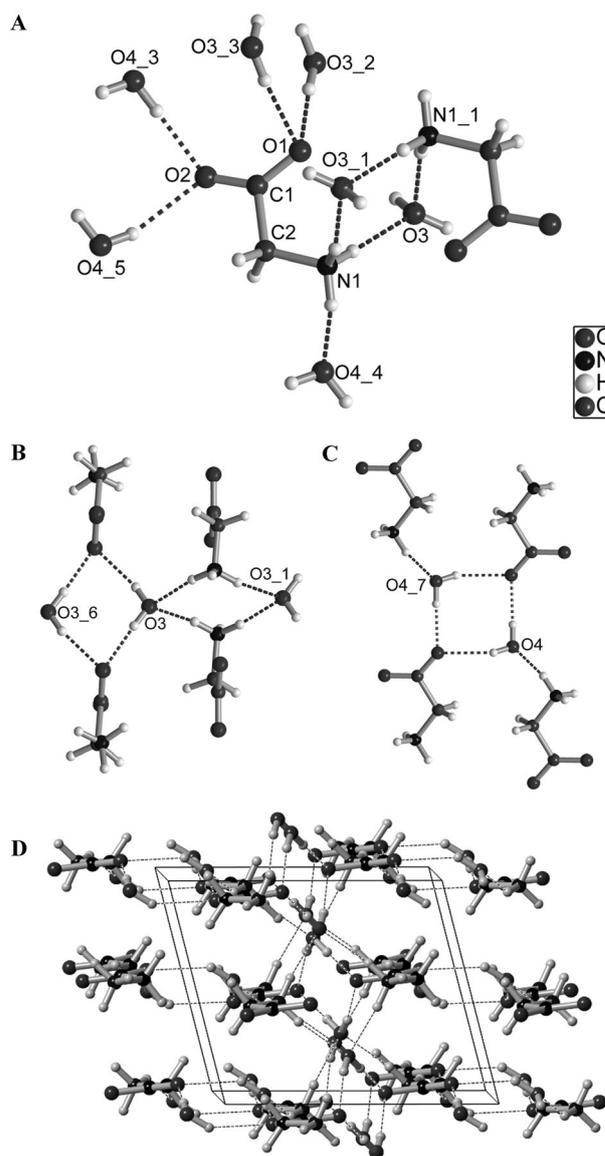


**Figure 1.** The synchrotron X-ray powder pattern of the unknown phase in ice (top) and the calculated pattern of the predicted structure of GDH (bottom).

parameters for the GDH search with  $Z' = 1$ . The same structure was identified as the global energy minimum at the optB88 level of theory, confirming that the results obtained from fixed-cell searches are indeed the best solution for the given stoichiometry.

The predicted crystal structure was used as the starting model for Rietveld refinement against the synchrotron powder data. The structure model of ice was taken from ICSD (No. 27873). The program GSAS<sup>[13a,b]</sup> was used for two-phase refinement. The refined structure is almost identical to the predicted one (Figure S8) in terms of the non-hydrogen atoms. A restrained refinement on the H atoms failed to yield meaningful hydrogen-bonding information owing to the resolution limit of the powder data. In the final refinement, the H atoms were fixed to the most probable positions according to the model from the prediction. The resulting lattice parameters are  $a = 8.9585(9)$  Å,  $b = 8.2166(8)$  Å,  $c = 7.6142(6)$  Å,  $\beta = 104.262(7)^\circ$ , and  $V = 543.20(8)$  Å<sup>3</sup>. Refinement details are provided in the Supporting Information.<sup>[23]</sup>

The asymmetric unit in the GDH structure contains one glycine and two water molecules. Each glycine molecule is surrounded by seven water molecules via three N–H···O and four O···H–O hydrogen bonds (Figure 2A). The glycine molecules are well separated by water and do not strongly interact with each other. The shortest distance between two adjacent glycine molecules related by an inversion center is 2.74 Å from O1 to H1A on N1\_1, that is, the hydrogen atom that forms the hydrogen bond with O3 (Figure 2A), indicating the absence of any significant interactions. The two independent water molecules in the asymmetric unit are surrounded only by glycine molecules. One of the water molecules is surrounded by four glycine molecules via two O–H···O and two O···H–N hydrogen bonds (Figure 2B) while the other one is surrounded by three glycine molecules via two O–H···O and one O···H–N hydrogen bond (Figure 2C). The first water molecule (O3) forms two types of  $R_4^2(8)$  hydrogen-bonding patterns,<sup>[14]</sup> one of which has four hydrogen atoms from two water molecules as donors and two



**Figure 2.** The closest neighbors of glycine (A) and two water molecules (B, C) in the GDH crystal structure as well as the molecular packing along the  $b$  axis (D). The same Figure is shown in color as Figure S16. Closest contacts: N1···O3 2.88 Å, N1···O3\_1 2.87 Å, N1···O4\_4 2.71 Å, O1···O3\_2 2.69 Å, O1···O3\_3 2.84 Å, O2···O4\_3 2.72 Å, O2···O4\_5 2.85 Å. Symmetry operations:  $_1$ :  $1-x, 1-y, 2-z$ ;  $_2$ :  $1-x, 0.5+y, 1.51-z$ ;  $_3$ :  $x, 1.5-y, 0.5+Z$ ;  $_4$ :  $x, 0.5-y, 0.51-z$ ;  $_5$ :  $-x, 0.5+y, 1.5-z$ ;  $_6$ :  $-x, 0.5+y, 1.5-z$ ;  $_7$ :  $-x, 1-y, 1-z$ .

oxygen atoms from two glycine molecules as acceptors while the other one has four hydrogen atoms from two NH<sub>3</sub> groups of two glycine molecules as donors and two water oxygen atoms as acceptors. The second water molecule (O4) also forms an  $R_4^2(8)$  hydrogen-bond motif with contributions from four water hydrogen atoms as donors and two glycine oxygen atoms as acceptors. These hydrogen bonds connect the glycine and water molecules and build a three-dimensional network (Figure 2D). The absence of strong interactions between glycine molecules in GDH is unique among the twelve crystal structures of other amino acid hydrates published thus far (Figure S9).<sup>[4,15a,b]</sup>

The discovery of GDH prompted us to rethink the crystallization of glycine from solution. An earlier study based on *ab initio* calculations<sup>[16a]</sup> had suggested that a glycine molecule in aqueous solution is most probably surrounded by six water molecules in its first solvation shell, in which three water molecules bind with  $\text{NH}_3^+$  and three with  $\text{COO}^-$ , including a cyclic four-center  $\text{COO}^- \cdots \text{H}_2\text{O}$  interaction (Figure S10). Later, it was argued that two extra water molecules could be accommodated near  $\text{COO}^-$  via two-center  $\text{O} \cdots \text{H} \cdots \text{O}$  interactions (Figure S11).<sup>[16b]</sup> However, the authors also suggested that the cyclic four-center  $\text{COO}^- \cdots \text{H}_2\text{O}$  hydrogen bond would be unfavorable with respect to competing effects from water $\cdots$ water interactions, supporting the existence of the glycine $\cdot$ 7 $\text{H}_2\text{O}$  species such as those in the crystal structure of GDH. These calculations implied that individual glycine molecules are isolated by water in solution and exist as monomers, which is consistent with the experimental observations from the freezing-point depression measurements<sup>[17]</sup> and the molecular dynamics simulation study.<sup>[18]</sup> Accordingly, we propose that glycine crystals that grow from aqueous solution begin as amorphous nuclei of glycine hydrate that consist of glycine $\cdot$ 7 $\text{H}_2\text{O}$  species. Furthermore, such glycine $\cdot$ 7 $\text{H}_2\text{O}$  species might also be present in comets, where glycine and water coexist in a cold environment.

The successive transitions from amorphous to GDH,  $\beta$ -, and  $\alpha$ -glycine happen only in one direction, and there is no evidence for the reverse transformations. This irreversibility demonstrates the relative stability of these phases, that is,  $\text{GDH} < \beta < \alpha$ . We applied three different van der Waals inclusive methods implemented in the VASP code, namely optB88,<sup>[12f]</sup> optPBE,<sup>[12f]</sup> and PBE + D3,<sup>[19]</sup> to confirm the energy ranking of these phases. All three methods yielded the same stability sequence  $\gamma \approx \alpha > \beta$ , which agrees with the experimental reports (Table 1).<sup>[20]</sup> With this encouragement, we turned to GDH. The enthalpy differences computed from three methods all suggest that the dehydration process from GDH to  $\beta$  and ice ( $I_h$ ) is energetically favorable, confirming that GDH is metastable.

The crystal structures of GDH and  $\beta$ -glycine have no similarity so that the  $\text{GDH} \rightarrow \beta$  dehydration has to be destructive and non-topotactic. Such a phase transformation can best be interpreted as a recrystallization process as for the  $\beta \rightarrow \alpha$  and  $\beta \rightarrow \gamma$  transformations<sup>[7]</sup> and other examples described by Mnyukh.<sup>[21]</sup> The diffraction data confirmed the absence of any intermediate phases aside from ice, GDH, and

$\beta$ -glycine during the transformation, suggesting that the  $\text{GDH} \rightarrow \beta$ -glycine dehydration is a recrystallization process with simultaneous loss of two crystalline water molecules.

The fact that only the metastable GDH and  $\beta$ -phases rather than stable  $\alpha$ - and  $\gamma$ -glycine crystallize in the ice matrix is in agreement with the observation that  $\beta$ -glycine favorably grows in nano-confined spaces.<sup>[22a-c]</sup> Not surprisingly, peak profile analysis of the X-ray powder patterns revealed that the average crystallite sizes of GDH and  $\beta$ -glycine are 110(5) and 90(5) nm, respectively. Indeed, the flash cooling process generated nanometer-sized spaces in the ice. Upon dehydration, the grain size decreased by approximately 18%. This small size change from GDH to  $\beta$ -glycine hints that only one nucleus might form in each nanospace during the dehydration (recrystallization) process. If there were many nucleation sites, the grain size would be much smaller.

In summary, we have established the structure of the unknown phase of glycine as glycine dihydrate, the first hydrated form of glycine. Computational crystal-structure prediction played a key role in solving this crystal structure. Given the tremendous importance of both water and glycine to life and their possible abundance in the universe, this glycine dihydrate structure, albeit originally discovered in pharmaceutical engineering processes, might have vital implications for planetary science.

## Acknowledgements

Work at UNLV is supported by the National Nuclear Security Administration under the Stewardship Science Academic Alliances program through DOE Cooperative Agreement DE-NA0001982. Work at NYU is supported by the Materials Research Science and Engineering Center (MRSEC) program of the National Science Foundation (NSF) under Award Numbers DMR-0820341 and DMR-1420073. Calculations were performed on XSEDE facilities and on the cluster of the Center for Functional Nanomaterials, Brookhaven National Laboratory under Contract No. DE-AC02-98CH10086 supported by the DOE-BES. This research used Beamline 17-BM of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. We thank B. Kahr, A. Shtukenberg, M. Tan, and the reviewers for suggestions on improving the manuscript.

## Conflict of interest

The authors declare no conflict of interest.

**Keywords:** amino acids · crystal structures · glycine · phase transitions · X-ray diffraction

**Table 1:** The enthalpy differences ( $\Delta H$ ) relative to  $\beta$ -glycine in  $\text{kJ mol}^{-1}$  per formula unit. The calculated values were obtained from the simulation in the VASP code with different vdW inclusive methods at the PBE level of theory. The experimental values were taken from the literature.

Solid-phase transitions	$\Delta H(\text{optB88})$	$\Delta H(\text{optPBE})$	$\Delta H(\text{PBE-D3})$	$\Delta H(\text{exp})^{[20]}$
$\beta \rightarrow \alpha$	1.74	1.50	1.01	0.327
$\beta \rightarrow \gamma$	1.34	1.33	0.40	0.595
$\text{GDH} \rightarrow$	6.36	3.60	8.39	— <sup>[a]</sup>
$\beta + 2\text{H}_2\text{O}$ (s)				

[a] The process was determined to be exothermic by Pyne and Suryanarayanan,<sup>[8a]</sup> but the  $\Delta H$  value was not provided in the paper.



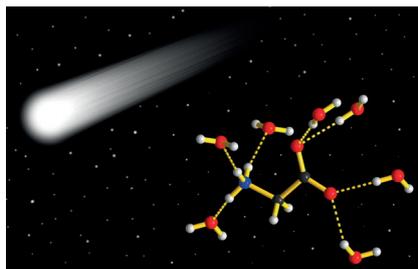
## Communications



## Amino Acids

W. Xu, Q. Zhu,\* C. Hu\* ——— ■■■–■■■

The Structure of Glycine Dihydrate:  
Implications for the Crystallization of  
Glycine from Solution and Its Structure in  
Outer Space



The **crystal structure** of glycine dihydrate was predicted by computational methods and confirmed by Rietveld refinement of synchrotron X-ray powder diffraction data. The discovery of this elusive phase improves our understanding of glycine crystallization from aqueous solution while it is suggested that this is the phase adopted by glycine in comets.