

**NH●●●O and OH●●●O interactions of glycine derivatives  
with squaric acid**

Journal:	<i>New Journal of Chemistry</i>
Manuscript ID:	NJ-ART-02-2014-000212.R2
Article Type:	Paper
Date Submitted by the Author:	08-Apr-2014
Complete List of Authors:	Aniola, Michalina; Adam Mickiewicz University, Faculty of Chemistry Dega-Szafran, Zofia; Adam Mickiewicz University, Faculty of Chemistry Katrusiak, Andrzej; Adam Mickiewicz University, Faculty of Chemistry Szafran, Mirosław; Adam Mickiewicz University, Faculty of Chemistry

**NJ-ART-02-2014-000212R1 revised****NH $\cdots$ O and OH $\cdots$ O interactions of glycine derivatives with squaric acid<sup>†</sup>**

Michalina Anioła, Zofia Dega-Szafran,\* Andrzej Katrusiak and Mirosław. Szafran

Four new hydrogen-bonded complexes of a simple amino acid glycine (GLY) and its methyl derivatives – sarcosine (N-methylglycine, SAR), dimethylglycine (DMG) and betaine (N,N,N-trimethylglycinium, BET) with squaric acid (3,4-dihydroxy-3-cyclobuten-1,2-dione, H<sub>2</sub>SQ) are synthesized and characterized by X-ray diffraction, FTIR and NMR spectra. The complexes differ in a stoichiometry, interaction with H<sub>2</sub>SQ and hydrogen-bonding system. Methylation of glycine gradually reverses the dissociation of squaric acid in co-crystals. This process is correlated with the number of N-H $\cdots$ O bonds to the squaric acid oxygens and the N-H donors capability of glycine derivatives. The effect of the presence of methyl groups on the proton and carbon-13 chemical shifts is studied. The DFT calculations are performed for one unit of complexes and suggested stable zwitterionic structures of SAR, DMG and BET in the complex with H<sub>2</sub>SQ, except GLY unit, which appears as an uncharged (neutral) form. The experimental and computed vibrational spectra of complexes studied are reported.

---

*Faculty of Chemistry, Adam Mickiewicz University, Umultowska 89b, 61-614 Poznań, Poland;*  
E-mail: [degasz@amu.edu.pl](mailto:degasz@amu.edu.pl). Fax: +48-61-8291555. Tel.: +48-61-8291216

<sup>†</sup> Electronic supplementary information (ESI) available Tables S1- S3; Fig S1, CCDC 966748 (1), 966749 (2), 966750 (3), 966751 (4). For ESI and crystallographic data in CIF or other electronic format see DOI.....

## 1. Introduction

Glycine, GLY, is the simplest protein structure unit and the smallest achiral amino acid that can act as a neuroinhibitor in mammalian center nervous system.<sup>1</sup> Sarcosine, SAR, as N-methyl derivative of glycine belongs to a group of biologically important compounds containing a carbon atom participating in transmethylation reactions. In nature, it occurs in starfish, sea urchins and in the antibiotic actinomycin.<sup>2</sup> N,N-Dimethylglycine, DMG, is a sweet-tasting substance classified as a nutrient. It has been presumed to be a key ingredient of vitamin B-15 or pangamic acid. In the living cell, DMG is present as a product of the metabolism pathway of choline and methionine. DMG has been used to for the treatment of mental disturbance conditions, like autism and epileptic seizures. DMG is a popular ingredient of supplementary diet tablets and energetic drinks for sportsmen.<sup>2,3</sup> Betaine, BET, trimethylglycinium, is a zwitterion containing an anionic carboxylate group and a positively charged quaternary ammonium group. Betaine is a biological oxidation product of choline and it is widely distributed in plants, algae and animal tissues, where it serves as a methyl transfer agent important in amino acid syntheses.<sup>4</sup> GLY,<sup>5</sup> SAR,<sup>6</sup> DMG<sup>7</sup> and BET<sup>8</sup> form crystalline adducts mainly with inorganic acids as well with some organic acids.<sup>7b,9</sup> Some of them display interesting physical properties exhibiting phase transitions leading to ferroelectric, antiferroelectric and ferroelastic phases as well as phases with commensurate and incommensurate superstructures.<sup>10</sup> The N-alkylation of amino acids is known to increase the population of neutral tautomers in solution.<sup>11</sup>

Since amino acids and their complexes are of considerable chemical and biological interest we have undertaken a systematic investigation of several squaric acid complexes with simplest amino acid as glycine and its N-methyl derivatives, sarcosine, dimethylglycine and betaine (Figure 1). Squaric acid, H<sub>2</sub>SQ, is a reagent for chemical synthesis, used for instance to make photosensitive squarate dyes,<sup>12</sup> inhibitors of protein tyrosine phosphatases and

squaraine-modified DNA.<sup>13</sup> The structure of squaric acid is not a perfect square.<sup>14</sup> The high acidity of H<sub>2</sub>SQ with pK<sub>1</sub> = 1.5 for the first proton and pK<sub>2</sub> = 3.4 for the second one<sup>15</sup> is attributable to resonance stabilization of the anion.<sup>16</sup> The conjugate base of squaric acid is the hydrogen squarate anion, HSQ<sup>-</sup>, whereas the conjugate base of the hydrogen squarate anion is divalent squarate anion, SQ<sup>2-</sup> (Figure 1).<sup>17</sup>

The proton transfers belong to most frequent transformations in Nature, however this process is still not well understood. Squaric acid and amino acids are ideal for investigating the proton transfer. Some earlier papers of the groups of Kolev et al. concern complexes of amino acids and their derivatives with H<sub>2</sub>SQ.<sup>18</sup> In the present work we report on the synthesis of four new complexes of H<sub>2</sub>SQ with glycine, GLY, and its methyl derivatives, sarcosine (N-methylglycine, SAR), dimethylglycine (DMG), and betaine (trimethylglycinium, BET). The main aim of this work is investigation of the interaction between the zwitterionic compounds and proton-donor. These complexes differ in their X-ray data, FTIR and NMR spectra. The differences are caused by the presence of different number of acidic protons and methyl groups attached to the nitrogen atoms. From the point of view of material science, the information on the formation of crystal structures of the complexes of glycine and its methyl derivatives with squaric acid is essential for understanding the role of hydrogen bonds and other intermolecular forces in the association of ions in the crystalline state. Quantum chemical calculations are performed for one unit of all complexes to study their structures in the isolated molecules and their IR frequencies.

## 2. Results and discussion

GLY, SAR, DMG and BET form with H<sub>2</sub>SQ stable, crystalline complexes. Their structures are studied by the single diffraction method, while the molecular structures of the isolated unit of complexes, without any interactions with the neighboring molecules, are optimized by the B3LYP/6-311G(d,p) approach. The experimental and calculated bond lengths, bond and

torsion angles are given in Tables S1 and S2, respectively (in Supplementary material), while the selected ones are listed in Table 1. The hydrogen bond geometries, calculated energies (HF) and dipole moments are listed in Table 2.

### 2.1. Crystal (**1**) and optimized (**1a**) structures of the 2:1 complex of glycine with squaric acid

Glycine, GLY, is a non-essential, neutral amino acid without center of chirality.<sup>2,19</sup> In an aqueous solution of pH ~7 GLY exists as zwitterion with the positively charged amino group and the negative carboxylate group. Binding of proton to glycine is shown in Scheme 1.<sup>20</sup> The  $K_1$  and  $K_2$  constants reflect the intrinsic properties of the amino and carboxyl groups and their interactions on protonation ( $\text{pK}(\text{COOH}) = 2.35$  and  $\text{pK}(\text{NH}_3^+) = 9.78$ ).

GLY with  $\text{H}_2\text{SQ}$  forms a crystalline adduct, **1**, at the ratio 2:1. Both OH groups of  $\text{H}_2\text{SQ}$  are engaged in the short, asymmetric and non-equivalent hydrogen bonds,  $\text{O}(5)\text{-H}(3)\cdots\text{O}(3)$  and  $\text{O}(5')\text{-H}(4)\cdots\text{O}(4)$  of 2.444(2) and 2.571(2) Å, respectively. Both O-H protons from  $\text{H}_2\text{SQ}$  are transferred to GLY, generating a squarate dianion,  $\text{SQ}^{2-}$  and two glycinium cations (Fig. 2a, Table 2). Contrary to the crystal of GLY,<sup>21</sup> the glycine units in the complex exist in the cationic forms. A similar arrangement of the base and  $\text{SQ}^{2-}$  is observed in the known complexes of  $\text{H}_2\text{SQ}$  with bases.<sup>17h,22</sup> However, the optimized structure of the 2:1 complex, **1a**, is more symmetric than the crystal one, and  $\text{H}_2\text{SQ}$  interacts with two uncharged (neutral) forms of GLY units, with the  $\text{O}(3)\text{-H}(3)\cdots\text{O}(5)$  and  $\text{O}(4)\text{-H}(4)\cdots\text{O}(5')$  lengths of 2.644 and 2.648 Å, respectively (Fig. 2b, Table 2), The COOH groups in GLY units in **1a** have *trans* conformations. In **1a** there are the  $\text{O}(6)\text{-H}(\text{N}1)\cdots\text{N}(1)$  and  $\text{O}(6')\text{-H}(\text{N}1')\cdots\text{N}(1')$  intramolecular contacts between the carboxylic protons and the nitrogen atoms of GLY moieties, of 2.588 and 2.574 Å, respectively (Table 2), while in the crystal of **1** the N-H atoms form several short N-H $\cdots$ O contacts with the squarate dianions of the neighboring molecules. Aggregates of bonded molecules are shown in Fig 3. The structure of the squarate dianion in

**1** is almost square and planar. The C-C-C bond angles vary from 89.3(1) to 91.2(1)<sup>o</sup>, whereas the C-C-C-C torsion angles are equal  $\pm 0.4(1)^\circ$  (Table 1).

## 2.2. Crystal (2) and optimized (2a) structures of the complex of sarcosine with squaric acid and water

Sarcosine, also known as N-methylglycine, is an intermediate and by-product in glycine synthesis and degradation.<sup>2</sup>

SAR forms a hydrated complex with H<sub>2</sub>SQ at a ratio SAR:H<sub>2</sub>SQ:0.5H<sub>2</sub>O (**2**). One proton from squaric acid is transferred to sarcosine and the hydrogen squarate anion, HSQ<sup>-</sup>, interacts with water (Fig. 4a). The protonation at O(5) is reflected in the C(5)-O(5) bond length of 1.314(1) Å, compared with the C(5)-O(6) bond of 1.201(1) Å. These bond lengths in pure sarcosine are 1.271(2) and 1.239(2) Å.<sup>23</sup> The geometry of the HSQ<sup>-</sup> anion shows that the hydroxyl group is located at the O(4) atom, which is confirmed by; (i) the C(4)-O(4) bond length of 1.308(1) Å is much longer than the other ones and (ii) the C(3)-C(4)-C(1) bond angle of 93.4(1)<sup>o</sup> is much open than the other three ones. The HSQ<sup>-</sup> anion insignificantly deviates from planarity, as illustrated by the O-C-C-O torsion angles varying from -3.1(1) to 3.1(1)<sup>o</sup> (Table 1).

The O(1W) oxygen atom of the water molecule is located on a two-fold axis, hence only H(1W) of the water molecule is symmetry independent. The water molecule links complexes of SARH<sup>+</sup>·HSQ<sup>-</sup> into two cyclamers,  $R_{5/2}^3(16)^{24}$  through four hydrogen bonds: O(5)-H(3)···O(3), O(1W)-H(1W)···O(3), O(4)-H(4)···O(1) and (1)-H(2N)-O(1) of 2.612(1), 2.849(1), 2.528(1) and 2.924(1) Å, respectively, whereas two squaric acids of the neighboring molecules form cyclamer  $R_{2/2}^2(10)^{24}$  through two O(4)-H(4)···O(1) hydrogen bonds of 2.528(2) Å (Fig. 5, Table 2).

The optimized structure at a ratio 1:1:1, SAR:H<sub>2</sub>SQ:H<sub>2</sub>O, by the B3LYP/6-311G(d,p) approach of the hydrated complex of SAR with H<sub>2</sub>SQ, **2a**, is shown in Fig 4b. The optimized

structure of **2a** differs from the crystal of **2**, by the lack of interactions between the neighboring molecules. H<sub>2</sub>SQ interacts with SAR by the O(3)-H(3)···O(5) hydrogen bond of 2.537 Å. The zwitterionic character of sarcosine is confirmed by the C(5)-O(5) and C(5)-O(6) bond lengths of 1.247 and 1.251 Å, respectively (Table 1). The water molecule is engaged in the O(W)-H(1W)···O(6) hydrogen bonds of 2.815 Å with the carboxylate group of the SAR unit. Additionally, the N(1)-H(2N) group forms an intramolecular hydrogen bond to the O(2) atom of 2.722 Å (Table 2).

### 2.3 Linear dimer of the 2:2 complex of dimethylglycine with squaric acid ( **3** and **3a** )

The DMG molecule has an amphoteric character and can exist in an anionic or a cationic form, as well as in two uncharged tautomeric forms, one of which with neutral amino and carboxylic groups, which is more stable in the gas phase and the dipolar zwitterionic form, which is more stable in aqueous solution.<sup>11,25</sup> According to a survey of the Cambridge Structural Database there are only two reports of crystal structures of the hydrogen-bonded complexes of DMG with acids,<sup>7</sup> where DMG exists in the cationic form.

The investigated complex of DMG with H<sub>2</sub>SQ, **3**, consists of two independent 1:1 complexes linked together into a linear dimer by the O(4')-H(4')···O(1) hydrogen bond of 2.624(2) Å (Fig. 6a, Table 2). In one of these complexes DMG is bonded to H<sub>2</sub>SQ through the O(3)-H(3)···O(5) and O(4)-H(4)···O(6) hydrogen bonds of 2.444(2) and 2.527(2) Å, respectively, with both of its oxygen atoms of the carboxylate group engaged. In the second 1:1 complex, the proton transfer occurs from H<sub>2</sub>SQ to DMG, and the DMG cation is joined to the hydrogen squarate anion through the O(5')-H(3')···O(3) hydrogen bond of 2.514(2) Å. The O(6') atom is not involved in any hydrogen bonds. The optimized structure of dimer **3a** (Fig. 6b) differs slightly from the crystal structure of **3**. In **3a** both DMG units exist in the zwitterionic forms and interact with two H<sub>2</sub>SQ molecules by the O-H···O hydrogen bonds longer than in the crystal of **3** (Table 2). The molecules in **3** are extended, whereas in the

optimized molecule of **3a** are bent. The angles between the plane of the squarate ring in **3** and **3a** are 21.37 and 63.18°, respectively.

In the crystal of **3** the dimers are linked to form two cyclamers described by the graph set  $R_2^6(30)$ ,<sup>24</sup> where the main role plays the N-H atoms engaged in the N(1)-H(1)···O(1') and N(1')-H(1')···O(2) hydrogen bonds of 2.870(2) and 2.803(2) Å, respectively (Fig. 7, Table 2).

#### 2.4 Cyclic dimer of betainium hydrogen squarate (**4**, **4a**)

Name betaine (BET) is used when the bonding proton is not transferred and betainium (BETH)<sup>+</sup>, when the proton is transferred from the acid to betaine.<sup>8c</sup>

BET and H<sub>2</sub>SQ form a crystalline complex **4** at the ratio 2:2 (Fig. 8a). In crystal **4** one proton from each squaric acid molecule is transferred to two BET molecules, generating hydrogen squarate anions, HSQ<sup>-</sup> and protonated betaine cations, BETH<sup>+</sup>. The HSQ<sup>-</sup> anions and the BETH<sup>+</sup> cations are linked through the asymmetric O(5)-H(4)···O(4) and O(5')-H(4')···O(4') hydrogen bonds of 2.559(4) and 2.546(4) Å, respectively. Two 1:1 complexes are joined into a cyclic dimer by two non-equivalent O(3)-H(3)···O(2') and O(3')-H(3')···O(2) hydrogen bonds of 2.503(4) and 2.501(4), respectively Å (Table 2). The formation of  $\alpha$ -dimer is one of the most common ways of aggregation of HSQ<sup>-</sup> anions, which are linked by two short O-H···O hydrogen bonds of 2.47 – 2.54 Å.<sup>17a,b,d</sup> All structures containing  $\alpha$ -dimers of HSQ<sup>-</sup> ions have the interesting property of being completely anhydrous, in spite of the fact that most of them crystallized from water.<sup>17c,d,e,g,j,26</sup> The crystal packing of dimers **4** in the unit cell along the [010] axis is shown in Fig. 9.

The optimized structure of the dimer, **4a**, by the B3LYP/6-311G(d,p) approach is much more symmetric than that of crystal **4** (Fig. 8b). No protons are transferred from squaric acid to betaine and complexes form a centrosymmetric dimer. The calculated dipole moment is 0.003 D. In the optimized complex **4a**, the O(4)-H(4)···O(5) hydrogen bonds of 2.473 Å, which linked H<sub>2</sub>SQ and BET molecules, are shorter than these in complex **4**, whereas the



O(3)-H(3)···O(2) bonds of 2.591 Å, which joined 1:1 complexes into a centrosymmetric dimer are longer than in **4** (Table 2). The presence of H<sub>2</sub>SQ molecules in dimer **4a** is confirmed by the C(2)-C(3)-C(4) and C(3)-C(4)-C(1) bond angles of 92.74 and 91.39, respectively, whereas in **4** the C(2)-C(3)-C(4) bond angles in both HSQ<sup>-</sup> are ca. 93°, while the C(3)-C(4)-C(1) ones are 89.0(3) and 89.1(3)°. A similar centrosymmetric dimer has been found in the crystal and optimized structures of the complex of quinuclidine betaine with squaric acid.<sup>26</sup>

In summary the differences between the crystal forms and optimized structures arise from the lack of N-H···O interactions in all isolated glycines - squaric acid complexes.

### **2.5. Correlation between the number of N<sup>+</sup>H donors and the number of the methyl groups in glycine derivatives with the ionization of squaric acid**

The increasing presence of N<sup>+</sup>H-donors in the glycine derivatives, (in sequence BET → DMG → SAR → GLY) leads to systematic changes in the H-bonding patterns, the molecular arrangement, hydration and, most importantly, in the dissociation of protons transferring from squaric acid to glycine derivatives (Fig. 10). We found that the presence of three N<sup>+</sup>H donors (in **1**) results in the full dissociation of squaric acid. The dissociation reduces through two N<sup>+</sup>H donors, when even a neutral squaric acid molecule was obtained. When N<sup>+</sup>H donors are absent (in **4**), half-dissociated squaric acid anions (HSQ<sup>-</sup>) are observed and they aggregate into doubly O-H···O bonded dimers. This process cannot be reproduced by calculations involving one isolated structural unit only, as it does not include N<sup>+</sup>H···O bonds to the squaric acid oxygen atoms. Fig. 10 shows the correlation between the frequency of N<sup>+</sup>H···O hydrogen bonds as a function of the number of methyl groups in glycine derivatives (no methyl in GLY, one methyl in SAR, two methyls in DMG and three methyl groups in BET). The squarate dianion, SQ<sup>2-</sup>, has eight H-acceptor sites in 2GLY<sup>+</sup>·SQ<sup>2-</sup>, six of them are used in N<sup>+</sup>H···O bonds and two remaining in O-H···O bonds (8 H-donors to 8-acceptors). The

average of the numbers of protons in the symmetry independent H<sub>2</sub>SQ molecule and HSQ<sup>-</sup> anion equal to 1.5 is plotted for DMG structure **3**. It can be concluded that each N-H...O hydrogen bonds somewhat reduce the double-bond character of the squaric acid carbonyls and in this way contribute to the dissociation of H<sub>2</sub>SQ molecules. This systematically happens for one and two methyls substituting ammonium protons and this relation is valid for GLY, SAR and DMG, however, this relation is reversed for BET, when no ammonium protons are available.

## 2.6. <sup>1</sup>H and <sup>13</sup>C NMR spectra

Both proton and carbon-13 NMR spectra of complexes of glycine and its methyl derivatives with squaric acid are very simple, however the effect of the number of methyl groups attached to the quaternary nitrogen atom on the proton and carbon-13 chemical shifts is significant. With increasing number of methyl groups in glycines the resonance signals of the CH<sub>2</sub> and CH<sub>3</sub> protons and methylene and methyl carbon-13 atoms are linearly shifted to lower magnetic field, whereas the resonance signal of the COO carbon atom undergoes the upfield shift (Fig. S1). The proton and carbon-13 chemical shifts are given in the Experimental Section. The single <sup>13</sup>C resonance signal of the hydrogen squarate carbon atoms appear in the lower magnetic field and it is observed between 198.64 and 198.86 ppm. The C=O and C-OH groups can form similarly strong H-bonds to the surrounding water molecules and carboxylate groups, leading to the  $\pi$ -conjugation of their C-C and C-O bonds.<sup>17a</sup>

## 2.7. Infrared spectra

Amino acids exhibit a great deal of hydrogen bonding as evidenced by the presence of many broad bands in their infrared spectra especially in the region 3000-2000 cm<sup>-1</sup>.<sup>8b,c,27</sup> The solid-state FTIR spectra of complexes **1**, **2**, **3** and **4** are shown in Fig. 11a,c,e,g, while the frequencies computed by the B3LYP/6-311G(d,p) approach for optimized structures **1a**, **2a**,

**3a** and **4a** are shown as vertical lines in Fig. 11b,d,f,h, respectively. The proposed assignments of bands were made using the Gauss View molecular visualization program.<sup>28</sup>

In the FTIR spectrum of **1** (Fig. 11a) the  $\nu\text{NH}$  band appears at  $3160\text{ cm}^{-1}$ , while the  $\nu\text{OH}$  vibration is assigned to the broad band in the  $3000\text{-}2000\text{ cm}^{-1}$  region. The calculated asym. and sym. stretching, bending in-plane and bending out-of-plane NH vibrations for GLY units in **1a** are observed at  $3614$  and  $3603$ ,  $3538$  and  $3507$ ,  $1663$  and  $1643$ ,  $984$  and  $935\text{ cm}^{-1}$ , respectively. The bands at  $3367$  and  $3344\text{ cm}^{-1}$  are attributed to the  $\nu\text{O}(6)\text{-H}\cdots\text{N}$  vibrations, while these at  $3212$  and  $3188\text{ cm}^{-1}$  to the  $\nu\text{O}(3,4)\text{-H}\cdots\text{O}(5)$  modes (Fig. 11b).

In the FTIR spectrum of **2** (Fig. 11c) the bands at  $3350$ ,  $2830$  and  $2470\text{ cm}^{-1}$  are assigned to the  $\nu\text{O-H}$  mode in water,  $\nu\text{N-H}\cdots\text{O}$  and  $\nu\text{O-H}\cdots\text{O}$  vibrations, respectively. The calculated  $\nu\text{O-H}$  in water,  $\nu\text{N-H}\cdots\text{O}$  and  $\nu\text{O-H}\cdots\text{O}$  vibrations for **2a** appear at  $3640$ ,  $3235$  and  $2715\text{ cm}^{-1}$ , respectively (Fig. 11d).

The experimental infrared spectrum of **3** (Fig. 11e) shows broad bands at  $2720$  and  $2480\text{ cm}^{-1}$  attributed to  $\nu\text{N-H}$  and  $\nu\text{O-H}$  modes, respectively. However, in the calculated spectrum of **3a** (Fig. 11f) there are several intensive bands at  $3365$  and  $3243\text{ cm}^{-1}$  assigned to the  $\nu\text{O}(3)\text{-H}\cdots\text{O}(5)$  and  $\nu\text{O}(4)\text{-H}\cdots\text{O}(6)$  modes, at  $3133\text{ cm}^{-1}$  to the  $\nu\text{O}(4')\text{-H}\cdots\text{O}(1)$  mode, at  $2878\text{ cm}^{-1}$  to the  $\nu\text{O}(3')\text{-H}\cdots\text{O}(5')$  vibration and at  $2471\text{ cm}^{-1}$  to the  $\nu\text{N-H}$  mode of atoms engaged in the strong intramolecular hydrogen bonds (Table 2).

The FTIR spectrum of **4** (Fig. 11g) is characterized by a broad absorption centered at  $2450\text{ cm}^{-1}$ , which is attributed to the  $\nu\text{O-H}\cdots\text{O}$  vibration. A similar absorption is observed in the spectra of betaine complexes with mineral and organic acids,<sup>8b,c,9</sup> and its shifts to lower or higher wavenumbers depend on the hydrogen bond strength. The calculated frequencies for **4a** at  $3160$  and  $2260\text{ cm}^{-1}$  are assigned to the stretching vibrations of the O-H groups engaged in the longer  $\text{O}(3)\text{-H}\cdots\text{O}(2)$  and shorter  $\text{O}(4)\text{-H}\cdots\text{O}(5)$  hydrogen bonds, respectively (Fig. 11h, Table 2).

The broadening of the IR absorption in the spectra of the complexes investigated is also observed in the region below  $1500\text{ cm}^{-1}$ . A similar broad band in this region was observed in the infrared spectrum of  $\text{H}_2\text{SQ}$ <sup>29</sup> and in the spectra of squaric acid complexes with various bases.<sup>17j,30</sup> This broad absorption overlaps the deformation vibrations of the methyl and methylene groups, hydroxyl groups, and ring and skeletal modes. However, the bands in the  $1800\text{-}1500\text{ cm}^{-1}$  region are attributed to the  $\nu\text{C}=\text{O}$ ,  $\nu\text{C}=\text{C}$  and  $\nu\text{COO}$  vibrations. Two bands at  $1759$  and  $1723\text{ cm}^{-1}$  are assigned to  $\nu\text{C}=\text{O}$  of glycinium moieties in the spectrum of **1**. Similarly, the  $\nu\text{C}=\text{O}$  band at  $1740\text{ cm}^{-1}$  was observed in the spectrum of glycinium monophenylphosphate.<sup>5a</sup> In the spectrum of **2** the  $\nu\text{C}=\text{O}$  band at  $1729\text{ cm}^{-1}$  of sarcosinium cation is observed and it appears in a similar frequency as in the spectrum of sarcosinium iodide.<sup>6d</sup> In the spectrum of the complex of **3**, both  $\nu_{\text{as}}\text{COO}$  and  $\nu\text{C}=\text{O}$  bands at  $1674$  and  $1731\text{ cm}^{-1}$  are observed and they are assigned to dimethylglycine and dimethylglycinium cation, respectively. The band at  $1740\text{ cm}^{-1}$  in the spectrum of **4** confirms the presence of the betainium cation in the complex with squaric acid. In the spectrum of the proton-transfer complex of BET with 2,6-dichloro-4-nitrophenol the  $\nu\text{C}=\text{O}$  band appears at  $1720\text{ cm}^{-1}$ , while in the spectrum of the molecular complex formed between BET and pentachlorophenol the  $\nu_{\text{as}}\text{COOH}$  band appears at  $1664\text{ cm}^{-1}$ .<sup>9b</sup>

The frequencies of the squaric acid moiety are much better recognized in the calculated IR spectra than in the experimental ones, where the  $\nu\text{C}=\text{O}$  vibrations appear as a single band in the  $1815\text{-}1790\text{ cm}^{-1}$  region. The calculated frequencies from the  $1885\text{-}1865\text{ cm}^{-1}$  range are assigned to the  $\nu_{\text{s}}\text{C}=\text{O}$ , while these in the  $1790\text{-}1745\text{ cm}^{-1}$  region are attributed to the  $\nu_{\text{as}}\text{C}=\text{O}$  modes in the  $\text{H}_2\text{SQ}$  moiety. The spectral features of the squarate units are similar to those observed earlier.<sup>17f,j,18e,27a,31</sup> The experimental and calculated spectra are slightly different, because the experimental spectrum is recorded for the complex in the solid state, while the computations are made for the vibrations of the isolated molecule in the gas phase, and the

calculations are based on harmonic frequencies, while the experimental ones on the anharmonic frequencies.

The experimental FTIR spectra confirm the crystal structure of the complex investigated, whereas the calculated frequencies correspond well to the structures of the isolated molecules.

The IR data confirm the key information about the proton positions determined in the crystalline complexes by X-ray diffraction.

### 3. Conclusions

In the crystals of complexes of squaric acid, H<sub>2</sub>SQ, with glycine, GLY, **1**, sarcosine, SAR, **2**, dimethylglycine, DMG, **3** and betaine, BET, **4**, the proton from H<sub>2</sub>SQ is transferred to amino acid molecules. The exception is the 2:2 complex of **3** in which one molecule of DMG is protonated, while the other exists in the zwitterionic form. In the optimized structures of one unit of complexes investigated (**1a**, **2a**, **3a**, **4a**) by the B3LYP/6-311G(d,p) approach, the amino acid moieties are in their zwitterionic forms with the exception of **1a**, in which GLY exists in an uncharged (neutral) form. The HSQ<sup>-</sup> and amino acid molecules are linked by the O-H...O hydrogen bonds with the O...O distances from 2.444 to 2.624 Å. Depending on methylation of glycine in its co-crystals with squaric acid, its dianion, monoanion, neutral and mixed form are stabilized. Methylation of glycine gradually reverses the dissociation of squaric acid in crystals. The resonance signals of the proton and carbon-13 of the methylene and methyl groups are shifted to lower magnetic field with increasing number of methyl groups attached to the amino group, while the resonance signal of the carboxylate group is linearly shifted to higher magnetic field.

### 4. Experimental section

#### 4.1. General procedure

Amino acids, GLY, SAR and DMG were purchased from Aldrich Chemical Co, whereas BET from Sigma. FTIR spectra were measured in Nujol and Fluorolube suspensions between KBr

plates using a Bruker IFS 66v/S instrument, with the resolution of  $2\text{ cm}^{-1}$ . Each spectrum was accumulated of 64 scans. The NMR spectra were recorded using a Varian VNMRs-400 spectrometer operating at 402.6435 and 101.2440 MHz for  $^1\text{H}$  and  $^{13}\text{C}$ , respectively. The spectra were measured in  $\text{D}_2\text{O}$  relative to an internal standard of 3-(trimethylsilyl)propionic- $\text{d}_4$  acid sodium salt. Elemental analysis was made using an Elemental Model Vario EL III.

## 4.2. Preparations

### 4.2.1 Synthesis of diglycinium squarate (1).

0.85 g of GLY and 0.64 g of  $\text{H}_2\text{SQ}$  were dissolved in 2 mL of water. The solvent was evaporated and the residue was recrystallized from a mixture of methanol-water (2:1), m.p.  $180\text{ }^\circ\text{C}$ . Elemental analysis calculated for  $\text{C}_8\text{H}_{10}\text{N}_2\text{O}_8$ : %C, 36.65; %H, 3.84; %N, 10.68; found: %C, 36.33; %H, 3.75; %N, 10.43.  $^1\text{H}$  NMR ( $\text{D}_2\text{O}$ , ppm)  $\delta$  3.89 ( $\text{CH}_2$ );  $^{13}\text{C}$  NMR ( $\text{D}_2\text{O}$ , ppm)  $\delta$  43.07 ( $\text{CH}_2$ ), 173.03 (COO), 198.85 (C=O,  $\text{H}_2\text{SQ}$ ).

### 4.2.2 Synthesis of sarcosinium hydrogen squarate hemihydrate (2)

To 0.60 g of SAR in 2 mL of methanol-water mixture (2:1), 0.76 g of  $\text{H}_2\text{SQ}$  in 1 mL water was added. The solvents were evaporated and the residue was recrystallized from a mixture of methanol-water (2:1), m.p.  $163\text{-}164\text{ }^\circ\text{C}$ . Elemental analysis calculated for  $\text{C}_{14}\text{H}_{20}\text{N}_2\text{O}_{13}$ : %C, 39.63; %H, 4.75; %N, 6.60; found: %C, 39.47; %H, 4.74; %N, 6.45.  $^1\text{H}$  NMR ( $\text{D}_2\text{O}$ , ppm)  $\delta$  2.80 ( $\text{CH}_3$ ), 3.93 ( $\text{CH}_2$ );  $^{13}\text{C}$  NMR ( $\text{D}_2\text{O}$ , ppm)  $\delta$  35,67 ( $\text{CH}_3$ ), 51.96 ( $\text{CH}_2$ ), 172.28 (COO), 198.86 (C=O,  $\text{H}_2\text{SQ}$ ).

### 4.2.3. Synthesis of the 2:2 complex of dimethylglycine with squaric acid (3)

0.40 g of DMG and 0.44 g of  $\text{H}_2\text{SQ}$  were dissolved in 1 mL of water. The crystals were filtered off and recrystallized from a mixture of methanol-water (3:1), m.p.  $182\text{-}183\text{ }^\circ\text{C}$ . Elemental analysis calculated for  $\text{C}_8\text{H}_{11}\text{N}_2\text{O}_6$ : %C, 44.24; %H, 5.10; %N, 6.45; found: %C, 44.15; %H, 5.03; %N, 6.10.  $^1\text{H}$  NMR ( $\text{D}_2\text{O}$ , ppm)  $\delta$  2.99 ( $\text{CH}_3$ ), 4.06 ( $\text{CH}_2$ );  $^{13}\text{C}$  NMR ( $\text{D}_2\text{O}$ , ppm)  $\delta$  46.70 ( $\text{CH}_3$ ), 60.58 ( $\text{CH}_2$ ), 171.41 (COO), 198.64 (C=O,  $\text{H}_2\text{SQ}$ ).

#### 4.2.4. Synthesis of betainium hydrogen squarate (4)

0.67 g of BET and 0.46 g of H<sub>2</sub>SQ were dissolved in 2 mL of water. The crystals were filtered off and recrystallized from methanol, m.p. 191 °C. Elemental analysis calculated for C<sub>9</sub>H<sub>13</sub>NO<sub>6</sub>: %C, 46.76; %H, 5.67; %N, 6.06; found: %C, 46.64; %H, 5.61; %N, 6.02. <sup>1</sup>H NMR (D<sub>2</sub>O, ppm) δ 3.32 (CH<sub>3</sub>), 4.21 (CH<sub>2</sub>); <sup>13</sup>C NMR (D<sub>2</sub>O, ppm) δ 56.73 (CH<sub>3</sub>), 66.88 (CH<sub>2</sub>), 170.30 (COO), 198.73 (C=O, H<sub>2</sub>SQ).

#### 4.3. Single crystal X-ray diffraction

X-ray diffraction data of complexes are collected on a KUMA KM-4 CCD diffractometer.<sup>32</sup> The structure was solved by direct methods using SHELXS-97 and refined on F<sup>2</sup> by the full-matrix least-squares with SHELXL-97.<sup>33</sup> The crystal data, details of data collection and structure refinement are given in Table 3 and the final atomic coordinates in Table S3 of Supplementary Material. The crystallographic and structural data in CIF format are available from the Cambridge Crystallographic Database Centre (CCDC 966748 for **1**, CCDC 966749 for **2**, CCDC 966750 for **3**, CCDC 966751 for **4**).

#### 4.4. Computation details

The DFT calculations were performed with the GAUSSIAN 03 program package.<sup>34</sup> The calculations employed the B3LYP exchange-correlation functional, which combines the hybrid exchange functional of Becke<sup>35</sup> with the gradient-correlation functional of Lee et al.<sup>36</sup> and the split-valence polarized 6-311G(d,p) basis set.<sup>37</sup> The X-ray geometry was used as a starting point for the calculations. The calculated IR frequencies are positive and confirmed that the optimized structures were in the states of minimum energy.

#### Acknowledgement

The computations were performed at the Poznań Supercomputing and Networking Center.

## References

1. (a) J.H. Caldwell, N.W. Daw and H.J. Wyatt, *J. Physiol.* 1978, **276**, 277-298; (b) L. Rajendra, J.W. Lynch and P. Schofield, *Pharmacol. Ther.* 1997, **73**, 121-146; (c) P. Schofield, *Adv. Neurol.* 2002, **89**, 263-274.
2. (a) H.D. Jakubke and H. Jeschkeit, *Aminosäuren, Peptide, Proteine*, Academic-Verlag, Berlin, 1982; H.-L. Lane, C.-L. Huang, P.-L. Wu, Y.-C. Liu, Y.-C. Chang, P.-y. Lin, P.-W. Chen and G. Tsai, *Biol. Psych.* 2006, **60**, 645-649.
3. (a) R. N. Girandola, R.A. Wiswell and R. Bulbulian, *Biochem. Med.* 1980, **24**, 218-222; (b) W.M. Bolman and J.A. Richmond, *J. Autism Dev. Disord.* 1999, **29**, 191-194; (c) M.E. Tonda and L.L. Hart, *Ann. Pharmacother* 1992, **26**, 935-937.
4. (a) X. Domingo, *Betaines in Amphoteric Surfactants*, Ed. E.G. Lomax, 1996, Vol. **59**, 75-190, Marcel Dekker Inc, New York, 1996; (b) N. Van Thoai, *Nitrogenous Bases in Comprehensive Biochemistry*, Eds. M. Florkin and E.H. Stotz, 1965, Vol. **6**, 208-252, Elsevier Publ. Comp. Amsterdam, 1965; (c) G. Blunden and S.M. Gordon, *Prog. Phycol. Res.* 1986, **4**, 39-80; (d) W.A. Rees, T.D. Yager, J. Korte and P.H. von Hippel, *Biochem.* 1993, **32**, 137-144.
5. (a) T. Głowiak and A.W. Szemik, *J. Cryst. Spectrosc. Res.* 1986, **16**, 79-89; (b) S. Natarajan, K. Ravikumar and S.S. Rajan, *Z. Kristallogr.* 1984, **168**, 75-82; (c) A.R. Al-Karaghoul, F.E. Cole, M.S. Lehmann, C.F. Miskell, J.J. Verbist and T.F. Koetzle, *J. Chem. Phys.* 1975, **63**, 1360-1366; (d) J. Baran, K. Łukaszewicz, A. Pietraszko and M. Śledź, *J. Mol. Struct.* 2002, **611**, 155-168; (e) D. Tran Qui, B. Lambert-Andron, and J.X. Boucherle, *Acta Crystallogr. C*, 1987, **43**, 907-909; (f) D. Tran Qui, J. Vicat and A. Durif, *Acta Crystallogr. C*, 1984, **40**, 181-184; S. Natarajan, *Z. Kristallogr.* 1983, **163**, 305-306; (g) T. Głowiak and I. Wnęk, *J. Cryst. Spectrosc. Res.* 1985, **15**, 157-171.
6. (a) S.C. Bhattacharyya and N.N. Saha, *J. Cryst. Mol. Struct.* 1978, **8**, 105-113, 209-215; (b) R.V. Krishnakumar and S. Natarajan, *Cryst. Res. Technol.* 1995, **30**, 825-830; (c) M.T. Averbuch-Pouchot, *Z. Kristallogr.* 1988, **183**, 285-291, 1993, **207**, 149-150; (d) V.V. Ghazaryan, M. Fleck and A.M. Petrosyan, *J. Mol. Struct.* 2013, **1032**, 35-40; (e) R.V. Krishnajumar, M.P. Ramella and S. Natarajan, *Cryst. Res. Technol.* 1996, **31**, 203-207; (f) A. Mostad and S. Natarajan, *Cryst. Res. Technol.* 1996, **31**, 295-300.



7. (a) B.D. Santarsiero and R. E. Marsh, *J. Cryst. Spectrosc. Res.* 1983, **11**, 245-251; (b) V.H. Rodrigues, J.A. Paixão, M.M.R.R. Costa and A.M. Matos Beja, *Acta Crystallogr. C*, 2001, **57**, 417-420
8. (a) J. Baran, M. Drozd, T. Głowiak, M. Śledź and H. Ratajczak, *J. Mol. Struct.* 1995, **372**, 131-144; (b) M. Iczyszyn, D. Godzisz and M.M. Ilczyszyn, *J. Mol. Struct.* 2003, **611**, 103-118; (c) D. Godzisz, M.M. Ilczyszyn and M. Ilczyszyn, *J. Mol. Struct.* 2002, **606**, 123-137; (d) W.H. Yip, R.-J. Wang and T.C.W. Mak, *Acta Crystallogr. C*, 1990, **46**, 717-179; (e) J. Baran, M. Drozd, T. Lis, M. Śledź, A.J. Barnes and H. Ratajczak, *J. Mol. Struct.* 1995, **354**, 109-118. D-D. Wu and T.W.C. Mak, *J. Chem. Crystallogr.* 1994, **24**, 689-694.
9. (a) Z. Dega-Szafran, I. Kowalczyk and M. Szafran, *Bull. Polish Acad. Sci. Chem.* 1995, **43**, 303-312; (b) Z. Dega-Szafran, A. Komasa, M. Grundwald-Wyspiańska, M. Szafran, G. Buczak and A. Katrusiak, *J. Mol. Struct.* 1997, **404**, 13-23; (c) M.M. Ilczyszyn, T. Lis and H. Ratajczak, *J. Mol. Struct.* 1995, **372**, 9-27; (d) J. Baran, A.J. Barnes, B. Engelen, M. Panthofer, A. Pietraszko and H. Ratajczak, *J. Mol. Struct.* 1995, **372**, 29-40.
10. (a) S. Haussühl, *Z. Kristallogr.* 1989, **188**, 311-320; (b) S. Sato, *J. Phys. Soc. Japan*, 1968, **25**, 185-201; (c) G. Schaack, *Ferroelectrics*, 1990, **104**, 147-158.
11. A.D. Headley, R. E. Corona and E.T. Cheung, *J. Phys. Org. Chem.* 1997, **10**, 898-900
12. (a) A.H. Schmidt, *Synthesis*, 1980, 961-994; (b) M. Panigrahi, S. Dash. S. Patel and B.K. Mishra, *Tetrahedron*, 2012, **68**, 781-805.
13. L.I. Markova, V.L. Malinovskii, L.D. Patsenker and R. Häner, *Org. Biomol. Chem.* 2012, **10**, 8944-8947.
14. (a) Y. Wang, G.D. Stucky and J. M. Williams, *J. Chem. Soc. Perkin Trans. 2*, 1974 35-38; (b) D. Semmingsen, F.J. Hollander and T.F. Koetzle, *J. Chem. Phys.* 1977, **66**, 4405-4412; (c) F.J. Hollander, D. Semmingsen and T.F. Koetzle, *J. Chem. Phys.* 1977, **67**, 4825-4831; (d) D. Semmingsen, *Acta. Chem. Scand. A*, 1973, **27**, 3961-3972; (e) A. Katrusiak and R.J. Nelmes, *J. Phys. C. Solid State Phys.* 1986, **19**, L765-L772.
15. (a) S. Cohen, J.R. Lacher and J.D. Park, *J. Am. Chem. Soc.* 1959, **81**, 3480; (b) D. J. MacDonald, *J. Org. Chem.* 1968, **33**, 4559-4560; (c) D.T. Ireland and H.F. Walton, *J. Phys. Chem.* 1967, **71**, 751-754.
16. R. West and D.L. Powell, *J. Am. Chem. Soc.* 1963, **85**, 2577-2579.
17. (a) G. Gilli, V. Bertolasi, P. Gilli and V. Ferretti, *Acta Crystallogr. B*, 2001, **57**, 859-865; (b) G. Gilli and P. Gilli, *The Nature of the Hydrogen Bond, Outline of a Comprehensive Hydrogen Bond Theory*, Oxford University Press, 2009, p. 251-253; (c) V. Bertolasi, P. Gilli, V. Ferretti and G. Gilli, *Acta Crystallogr. B*, 2001, **57**, 591-598; (d) B.B. Koleva, T.

- Kolev, R.W. Seidel, M. Spitteller, H. Mayer-Figge and W.S. Sheldrick, *J. Phys. Chem. A*, 2009, **113**, 3088-3095; (e) S. Mathew, G. Paul, K. Shivasankar, A. Choudhury and C.N.R. Rao, *J. Mol. Struct.* 2002, **641**, 263-279; (f) B.B. Koleva, T. Kolev, R.W. Seidel, H. Mayer-Figge, M. Spitteller and W.S. Sheldrick, *J. Phys. Chem. A*, 2008, **112**, 2899-2905.; (g) T. Kolev, R. Wortmann, M. Spitteller, W.S. Sheldrick and M. Heller, *Acta Crystallogr. E*, 2004, **60**, o956-o957; (h) A. Bulut, O.Z. Yeşilel, N. Dege, H. Icbudak, H. Olmez and O. Büyükgüngör, *Acta Crystallogr. C*, 2003, **59**, o727-o729; (i) I. Uçar, A. Bulut, O.Z. Yeşilel and O. Büyükgüngör, *Acta Crystallogr. C*, 2004, **60**, o585-o588 ; (j) B.B. Koleva, T. Kolev, R.W. Seidel, T. Tsanev, H. Mayer-Figge, M. Spitteller and W.S. Sheldrick, *Spectrochim. Acta A*, 2008, **71**, 695-702; (k) I.L. Karle, D. Ranganathan and V. Haridas, *J. Am. Chem. Soc.* 1996, **118**, 7128-7133.
18. (a) T. Kolev, D. Yancheva, M. Spitteller, W.S. Sheldrick and H. Mayer-Figge, *Acta Crystallogr E*, 2006, **62**, o463-o465; (b) O. Angelova, R. Petrova, V. Radomirska and T. Kolev, *Acta Crystallogr C*, 1996, **52**, 2218-2220; (c) O. Angelova, V. Velikova, T. Kolev and V. Radomirska, *Acta Crystallogr C*, 1996, **52**, 3252-3256; (d) T. Kolev, H. Mayer-Figge, R.W. Seidel, W.S. Sheldrick, M. Spitteller and B.B. Koleva, *J. Mol. Struct.* 2009, **919**, 246-254; (e) T. Kolev, R. Stahl, H. Preut, V. Koniczek, P. Bleckmann, V. Radomirska, *Zeit. Kristallogr. New Cryst. Struct.* 1998, **213**, 167-169; (f) T. Kolev, R. Stahl, H. Preut, P. Bleckmann, V. Radomirska, *Zeit. Kristallogr. New Cryst. Struct.* 1998, **213**, 169-170; (g) T. Kolev, Z. Glavcheva, R. Stahl, H. Preut, P. Bleckmann, V. Radomirska, *Zeit. Kristallogr. New Cryst. Struct.* 1999, **214**, 193-194; (h) T. Kolev, M. Spitteller, W.S. Sheldrick, H. Mayer-Figge, *Acta Crystallogr C*, 2006, **62**, o299-o300; (i) T. Kolev, M. Spitteller, W.S. Sheldrick, H. Mayer-Figge, *Acta Crystallogr C*, 2005, **61**, o4292-o4294.
19. J.M. Berg, J.L. Tymoczko and L. Stryer, *Biochemistry*, W.H. Freeman and Company, New York, Basingstone, 2007.
20. (a) H.A. Saroff, *J. Chem. Educ.* 1994, **71**, 637-643; (b) *Nomenclature and Symbolism for amino acids and peptides*, *Pure Appl. Chem.* 1984, **56**, 595-624.
21. (a) Y. Iitaka, *Acta Crystallogr.* 1960, **13**, 35-45, 1960, **14**, 1-10; (b) K. Ishikawa, M. Tanaka, T. Suzuki, A. Sekine, T. Kawasaki, K. Soai, M. Shiro, M. Lahav and T. Asahi, *Chem. Commun.* 2012, **48**, 6031-6033, and references cited therein.
22. (a) J.A. Kanters, A. Schouten, J. Kroon and E. Grech, *Acta Crystallogr. C*, 1991, **47**, 807-810; (b) P. Barczyński, Z. Dega-Szafran, A. Katrusiak and M. Szafran, *J. Mol. Struct.*

- 2012, **1013**, 95-101; (c) U. Korkmaz and I. Uçar, A. Bulut and O. Büyükgüngör, *Struct. Chem.* 2011, **22**, 1249-1259.
23. (a) A. Mostad and S. Natarajan, *Acta Chem. Scand.* 1989, **43**, 1004-1006; (b) B. Dittrich and M.A. Spackman, *Acta Crystallogr. A*, 2007, **63**, 426-436.
24. M.C. Etter, J.C. MacDonald and J. Bernstein, *Acta Crystallogr. B*. 1990, **46**, 256-262.
25. (a) G.M. Barrow, *J. Am. Chem. Soc.* 1958, **80**, 86-88; (b) A.D. Headley, B. Patel and E.T. Cheung, *Tetrahedron Lett.* 1996, **37**, 6673-6676.
26. (a) Z. Dega-Szafran, A. Katrusiak and M. Szafran, *J. Mol. Struct.* 2012, **1030**, 184-190; (b) E.J. MacLean, P.S. Wheatley, G. Ferguson and C. Glidewell, *Acta Crystallogr. C*, 1999, **55**, 1892-1896.
27. (a) A. Leiper and E.R. Lippincott, *J. Am. Chem. Soc.* 1957, **79**, 5098-5101; (b) A. Barth, *Prog. Biophys. Mol. Biol.* 2000, **74**, 141-173; (c) M.T. Rosado, M. Leonor, T.S. Duarte and R. Fausto, *Vib. Spect.* 1998, **16**, 35-54.
28. A.E. Frisch, R.D. Dennington, T.A. Keith, A.B. Neilsen and A.J. Holder, *Gauss View*, revision 3.0.9, Gaussian, Inc., Pittsburgh PA. 2003.
29. (a) F.G. Baglin and C.B. Rose, *Spectrochim. Acta A*, 1970, **26**, 2293-2304; (b) S.L. Georgopoulos, R. Diniz, M.I. Yoshida, N.L. Speziali, H.F. Dos Santos, G. M. A. Junqueira, and L.F.C. de Oliveira, *J. Mol. Struct.* 2006, **794**, 63-70.
30. U. Korkmaz and A. Bulut, *J. Mol. Struct.* 2013, **1050**, 61-68.
31. (a) P. Barczyński, Z. Dega-Szafran, A. Katrusiak and M. Szafran, *J. Mol. Struct.* 2011, **998**, 240-245; (b) T. Kolev, R. W. Seidel, B. B. Koleva, H. Mayer-Figge, M. Spitteller and W. S. Sheldrick, *J. Mol. Struct.* 2009, **931**, 45-49; (c) B. Koleva, R. W. Seidel, T. Kolev, S. Zareva, H. Mayer-Figge, I. M. Opiel, and W. S. Sheldrick, *J. Mol. Struct.* 2009, **921**, 163-171. (d) T. Głowiak and A.W. Szemik, *J. Cryst. Spectrosc. Res.* 1986, **16**, 79-89; (e) J. Nowicka-Scheibe, E. Grech, W. Sawka-Dobrowolska, G. Bator and L. Sobczyk, *Polish J. Chem.* 2007, **81**, 643-652; (f) T.M. Kolev, B.A. Stamboliyska, D.Y. Yancheva and V. Enchev, *J. Mol. Struct.* 2004, **691**, 241-248; (g) P. Barczyński, Z. Dega-Szafran, A. Katrusiak and M. Szafran, *J. Mol. Struct.* 2012, **1018**, 28-34; (h) P. Barczyński, Z. Dega-Szafran, A. Katrusiak and M. Szafran, *J. Mol. Struct.* 2012, **1013**, 95-101; (i) Z. Dega-Szafran, A. Katrusiak, A. Komasa and M. Szafran, *Supramolecular. Chem.* 2013, **25**, 432-440.
32. (a) *KUMA KM4* CCD Software, version 161, **1999**, Kuma Diffraction (Wroclaw, Poland); (b) *CrysAlis I62*, Kuma Diffraction, **1999** (Wroclaw, Poland).
33. G. M. Sheldrick, *Acta Crystallogr. A*, 2008, **64**, 112-122.

34. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez and J.A. Pople, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford CT, 2004.
35. A.D. Becke, *J. Chem. Phys.* 1993, **98**, 5648-5652; 1997, **107**, 8554-8580.
36. C. Lee, W. Yang and G.R. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.
37. W.J. Hehre, L. Random, P.V.R. Schleyer and J.A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.

**Table 1.** Selected experimental and calculated by the B3LYP/6-311G(d,p) bond lengths (Å), bond and torsion angles (°) for complexes of squaric acid with glycine (**1**, **1a**), sarcosine (**2**, **2a**), dimethylglycine (**3**, **3a**) and betaine (**4**, **4a**).

Parameters	<b>1</b>		<b>1a</b>		<b>2</b>		<b>2a</b>		<b>3</b>		<b>3a</b>		<b>4</b>		<b>4a</b>	
	A <sup>a</sup>	B <sup>b</sup>	A <sup>a</sup>	B <sup>b</sup>	A <sup>a</sup>	B <sup>b</sup>	A <sup>a</sup>	B <sup>b</sup>	A <sup>a</sup>	B <sup>b</sup>	A <sup>a</sup>	B <sup>b</sup>	A <sup>a</sup>	B <sup>b</sup>	A <sup>a</sup>	B <sup>b</sup>
<i>Bond lengths</i>																
O(1)-C(1)	1.242(2)		1.215		1.250(1)	1.208	1.236(2)	1.235(2)	1.219	1.212	1.207(4)	1.216(4)	1.210	1.217		
O(2)-C(2)	1.251(2)		1.201		1.215(1)	1.219	1.218(2)	1.223(2)	1.201	1.220	1.248(4)	1.240(4)	1.227	1.227		
O(3)-C(3)	1.265(2)		1.305		1.247(1)	1.296	1.287(2)	1.258(2)	1.309	1.302	1.314(4)	1.304(4)	1.303	1.302		
O(4)-C(4)	1.249(2)		1.322		1.308(1)	1.323	1.300(2)	1.318(2)	1.317	1.305	1.240(4)	1.250(4)	1.285	1.285		
O(5)-C(5)	1.295(2)	1.312(2)	1.215	1.217	1.314(1)	1.247	1.266(2)	1.305(2)	1.229	1.242	1.298(4)	1.305(4)	1.262	1.262		
O(6)-C(5)	1.211(2)	1.197(2)	1.320	1.324	1.201(1)	1.251	1.245(2)	1.213(2)	1.267	1.255	1.207(4)	1.185(4)	1.233	1.233		
<i>Bond angles</i>																
C(1)-C(2)-C(3)	89.9(1)		86.67		87.7(1)	88.5(1)	87.7(1)	88.5(1)	86.56	87.91	88.3(3)	88.9(3)	88.33	88.37		
C(2)-C(3)-C(4)	91.2(1)		92.86		89.4(1)	89.7(1)	91.4(1)	89.6(1)	92.50	92.10	93.3(3)	93.2(3)	92.74	92.73		
C(3)-C(4)-C(1)	89.7(1)		93.02		93.4(1)	94.2(1)	91.9(1)	94.2(1)	92.74	92.97	89.9(3)	89.1(3)	91.39	91.38		
C(4)-C(1)-C(2)	89.3(1)		87.45		89.4(1)	87.6(1)	89.0(1)	87.6(1)	88.16	87.01	87.9(3)	88.7(3)	87.54	87.55		
<i>Torsion angles</i>																
C(1)-C(2)-C(3)-C(4)	-0.4(1)		0.16		-1.3(1)	-0.15	1.0(1)	0.8(1)	0.42	-0.31	-0.1(3)	-1.8(3)	0.19	0.19		
C(2)-C(3)-C(4)-C(1)	0.4(1)		-0.17		1.3(1)	0.16	-1.1(1)	-0.9(1)	-0.43	0.32	0.1(3)	1.8(3)	-0.19	0.19		
C(3)-C(4)-C(1)-C(2)	-0.4(1)		0.16		-1.3(1)	-0.15	1.0(1)	0.9(1)	0.42	-0.30	-0.1(3)	-1.7(3)	0.19	-0.19		
C(4)-C(1)-C(2)-C(3)	0.4(1)		-0.15		1.3(1)	0.14	-1.0(1)	-0.8(1)	-0.39	0.28	0.1(3)	1.7(3)	-0.18	0.18		
O(5)-C(5)-C(6)-N(1)	169.1(1)	170.4(1)	-176.42	-160.34	178.0(1)	143.51	161.8(1)	-166.4(1)	-176.36	-174.04	-168.3(3)	167.2(3)	170.91	-170.90		
O(6)-C(5)-C(6)-N(1)	-11.8(2)	9.2(2)	4.05	20.24	-0.5(2)	34.03	-21.4(2)	15.4(2)	4.18	4.54	-10.7(6)	-13.9(6)	-8.03	8.04		

<sup>a</sup> Part A - the positions denoted by the unprimed atomic labels

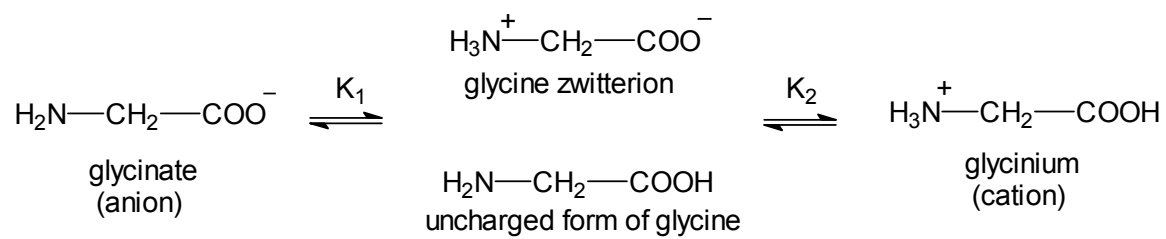
<sup>b</sup> Part B - the positions denoted by the primed atomic labels

**Table 2.** Experimental and calculated by the B3LYP/6-311G(d,p) approach geometry of the hydrogen bonds for complexes of squaric acid with glycine (**1**, **1a**), sarcosine (**2**, **2a**), dimethylglycine (**3**, **3a**) and betaine (**4**, **4a**), and calculated energies (HF, a.u.) and dipole moments ( $\mu$ , D)

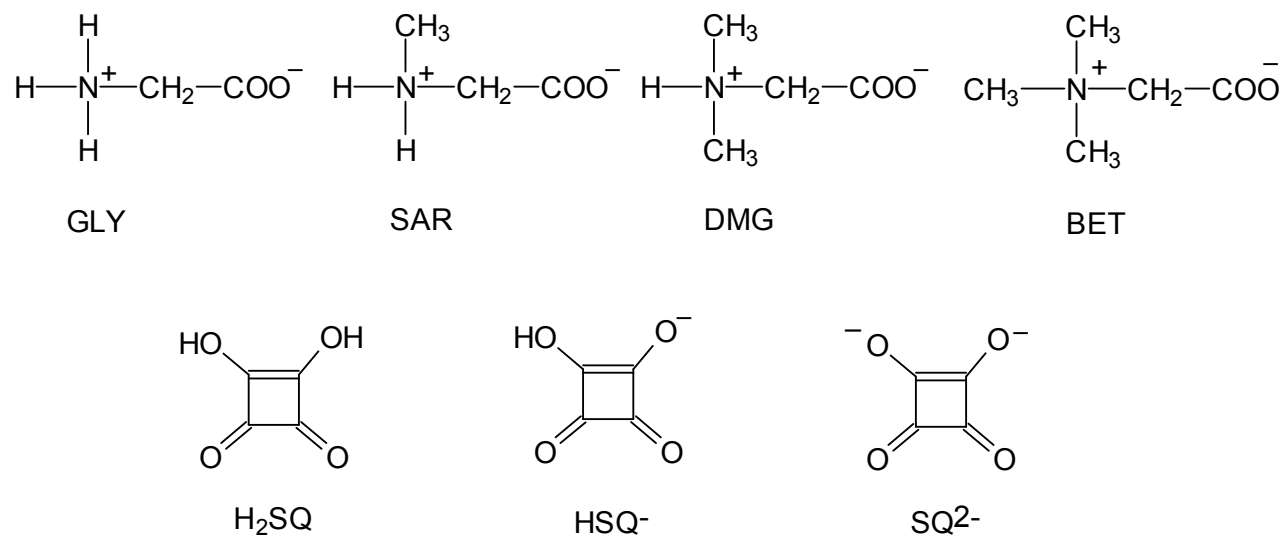
	D-H...A	D-H	H...D	D...A	$\angle$ DHA	
<i>X-Ray</i>						Symmetry code
<b>1</b>	O(5)-H(3)...O(3)	1.00(2)	1.46(2)	2.444(2)	167(2)	-
	O(5')-H(4)...O(4)	0.80(2)	1.77(2)	2.571(2)	175(3)	-
	N(1')-H(2'N)...O(2)	0.93(2)	1.90(2)	2.815(2)	139(1)	1.5-x, 0.5+y, 1.5-z
	N(1')-H(3'N)...O(5)	0.85(2)	2.11(2)	2.805(2)	133(1)	2-x, y, 1.5-z
	N(1)-H(2N)...O(1)	0.87(2)	2.08(2)	2.916(2)	157(2)	2-x, 2-y, 2-z
	N(1)-H(2N)...O(6')	0.89(2)	2.72(2)	2.828(2)	88(2)	2-x, 2-y, 2-z
<b>2</b>	O(5)-H(3)...O(3)	0.92(2)	1.70(2)	2.612(1)	171(2)	-
	O(1W)-H(1W)...O(3)	0.85(2)	2.00(1)	2.849(1)	173(5)	-x, y, 1.5-z
	O(4)-H(4)...O(1)	0.96(2)	1.60(2)	2.528(1)	161(2)	0.5+x, 0.5-y, 2-z
	N(1)-H(2N)...O(1)	0.91(2)	2.09(2)	2.924(1)	152(1)	-0.5+x, 0.5-y, -0.5+z
	N(1)-H(1N)...O(1W)	0.97(1)	1.90(2)	2.867(1)	173(2)	-0.5+x, 0.5+y, z
<b>3</b>	O(3)-H(3)...O(5)	0.82	1.65	2.444(2)	163	-
	O(4)-H(4)...O(6)	0.88(2)	1.66(2)	2.527(2)	170(2)	-
	O(4')-H(4')...O(1)	0.94(2)	1.68(2)	2.624(2)	176(2)	-
	O(5')-H(3')...O(3')	0.93(2)	1.59(2)	2.514(2)	174(2)	-
	N(1)-H(1)...O(1')	0.90(2)	2.16(2)	2.870(2)	136(1)	1-x, 1-y, 1-z
	N(1')-H(1')...O(2)	0.92(2)	1.94(2)	2.803(2)	157(2)	-x, 1-y, -z
	N(1')-H(1')...O(6')	0.92(2)	2.55(2)	3.035(2)	114(1)	-x, -0.5+y, -0.5-z
<b>4</b>	O(5)-H(4)...O(4)	0.82	1.76	2.559(4)	164	-
	O(3)-H(3)...O(2')	0.97(4)	1.55(4)	2.503(4)	167(4)	-
	O(3')-H(3')...O(2)	0.90(4)	1.61(4)	2.501(4)	166(3)	-
	O(5')-H(4')...O(4')	0.82	1.73	2.546(4)	178	-
<i>B3LYP/6-311G(d,p)</i>						Energy/ $\mu$ ,
<b>1a</b>	O(3)-H(3)...O(5)	0.997	1.648	2.644	176.57	-1023.673350 / 8.574
	O(4)-H(4)...O(5')	0.995	1.658	2.648	172.63	
	O(6)-H(1N)...N(1)	0.990	1.862	2.588	127.64	
	O(6')-H(1N')...N(1')	0.991	1.865	2.574	125.89	
<b>2a</b>	O(3)-H(3)...O(5)	1.023	1.518	2.537	173.67	-854.906921 / 7.001
	O(W)-H(1W)...O(6)	0.973	1.868	2.815	163.74	
	N(1)-H(2N)...O(2)	1.030	1.783	2.722	149.71	
<b>3a</b>	O(3)-H(3)...O(5)	0.987	1.693	2.671	170.46	-1635.534053 / 25.829
	O(4)-H(4)...O(6)	0.993	1.673	2.658	171.35	
	O(4')-H(4')...O(1)	0.999	1.646	2.643	175.80	
	O(3')-H(3')...O(5')	1.013	1.561	2.570	172.69	
	N(1)-H(N1)...O(6)	1.064	1.669	2.495	130.49	
	N(1')-H(N1')...O(6')	1.094	1.556	2.460	135.60	
<b>4a</b>	O(4)-H(4)...O(5)	1.052	1.427	2.473	171.72	-1714.166126 / 0.003
	O(3)-H(3)...O(2')	1.000	1.600	2.591	169.85	
	O(3')-H(3')...O(2)	1.000	1.601	2.591	169.87	
	O(4')-H(4')...O(5')	1.052	1.427	2.473	171.71	

**Table 3.** Crystal data and structure refinement for complexes of squaric acid with glycine (**1**), sarcosine (**2**), dimethylglycine (**3**) and betaine (**4**).

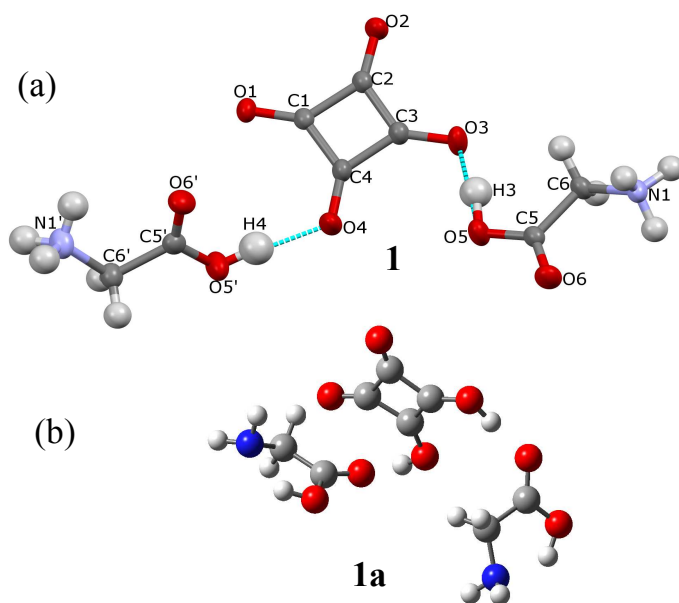
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
Empirical formula	C <sub>8</sub> H <sub>12</sub> N <sub>2</sub> O <sub>8</sub>	C <sub>14</sub> H <sub>18</sub> N <sub>2</sub> O <sub>12</sub> ·H <sub>2</sub> O	C <sub>16</sub> H <sub>22</sub> N <sub>2</sub> O <sub>12</sub>	C <sub>9</sub> H <sub>13</sub> NO <sub>6</sub>
Formula weight	264.20	424.32	434.36	213.20
Temperature (K)	296(2)	296(2)	296(2)	296(2) K
Wavelength (Å)	1.71073	1.54184	0.71073	0.71073 Å
Crystal system	Monoclinic	Monoclinic	Monoclinic	Orthorhombic
Space group	<i>C2/c</i>	<i>C2/c</i>	<i>P2<sub>1</sub>/c</i>	<i>Pca2<sub>1</sub></i>
Unit cell dimensions				
<i>a</i> (Å)	16.8392(3)	16.6068(2)	17.5110(5)	22.2718(14)
<i>b</i> (Å)	8.3163(2)	7.21436(8)	6.5385(2)	6.0451(3)
<i>c</i> (Å)	15.8267(3)	15.44410(17)	17.3475(4)	16.2765(8)
β (°)	100.063(2)	97.0249(11)	108.249(3)	90.00
Volume (Å <sup>3</sup> )	2182.27(8)	1836.43(4)	1886.32(9)	2191.4(2)
Z	8	4	4	8
Calculated density (g/cm <sup>3</sup> )	1.608	1.535	1.529	1.402
Absorption coefficient (mm <sup>-1</sup> )	0.146	1.213	0.133	0.119
F(000)	1104	888	912	976
Crystal size (mm)	0.18 x 0.22 x 0.36	0.16 x 0.26 x 0.33	0.22 x 0.28 x 0.38	0.19 x 0.23 x 0.38
θ range for data collection (°)	3.13 – 27.59	5.37 – 75.31	2.47 – 29.13	2.22 – 29.14
Max/min. indices <i>h, k, l</i>	-20/21, -10/10, -19/19	-20/20, -8/7, -18/19	-23/22, -7/8, -23/22	-29/20, -5/8, -22/18
Reflections collected / unique	10909 / 2371	6657 / 1862	11666 / 4520	8662 / 4068
R <sub>int</sub>	0.0156	0.0098	0.0206	0.0353
θ <sub>Max</sub> (°) / Completeness (%)	93.5	98.0	89.3	88.7
Refinement method	Full-matrix least-squares on F <sup>2</sup>	Full-matrix least-squares on F <sup>2</sup>	Full-matrix least-squares on F <sup>2</sup>	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	2371 / 0 / 199	1862 / 0 / 147	4520 / 0 / 341	4068 / 1 / 297
Goodness-of-fit on F <sup>2</sup>	0.697	0.807	0.968	0.766
Final R1 / wR2 indices [I>2σ <sub>I</sub> ]	0.0382 / 0.1098	0.0331 / 0.1013	0.0386 / 0.0970	0.0400 / 0.0891
R1 / wR2 indices (all data)	0.0420 / 0.1159	0.0337 / 0.1020	0.0610 / 0.1042	0.0942 / 0.1069
Largest diff. peak and hole (e.Å <sup>-3</sup> )	0.296 and -0.278	0.234 and -0.141	0.297 and -0.328	0.156 and -0.185

Scheme 1. Binding of protons to glycine.<sup>20</sup>

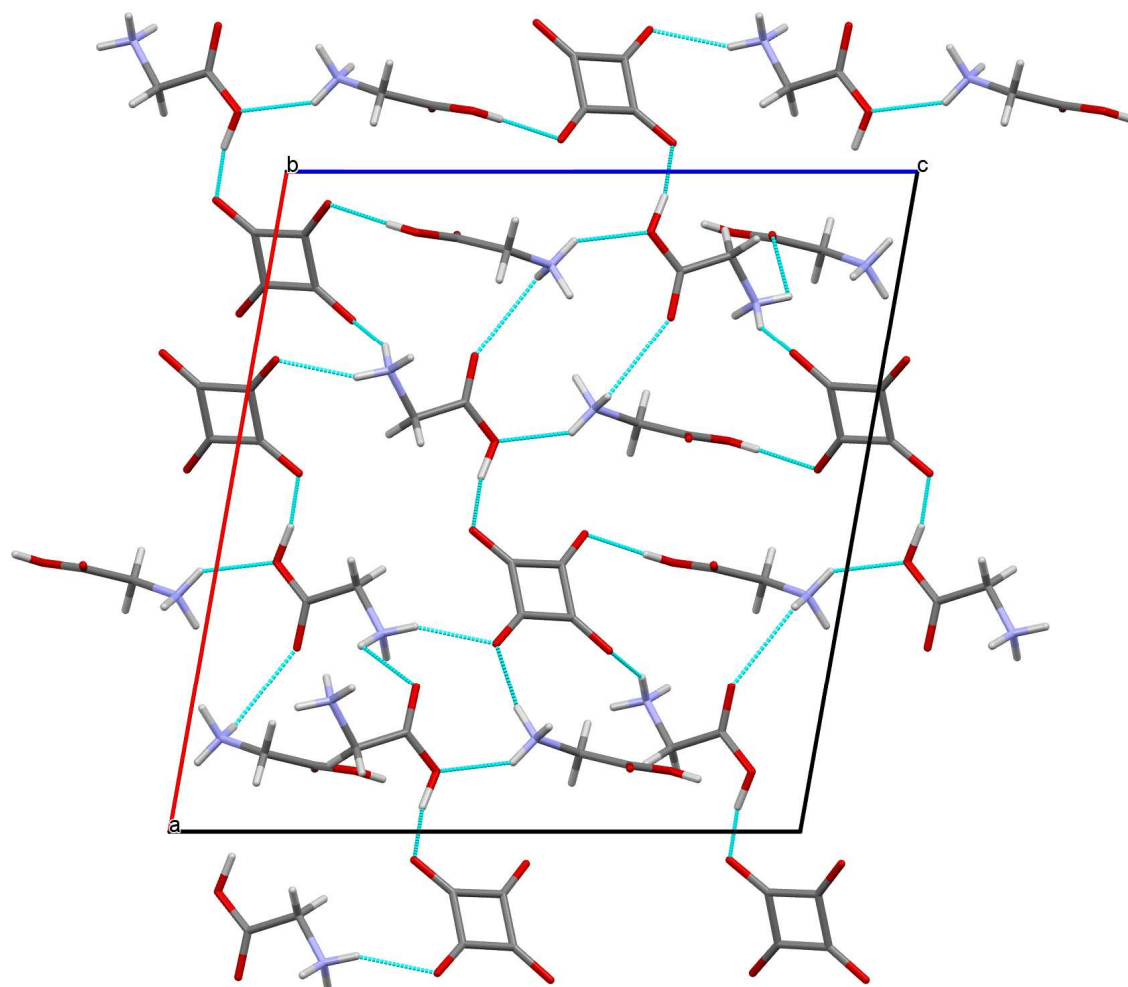




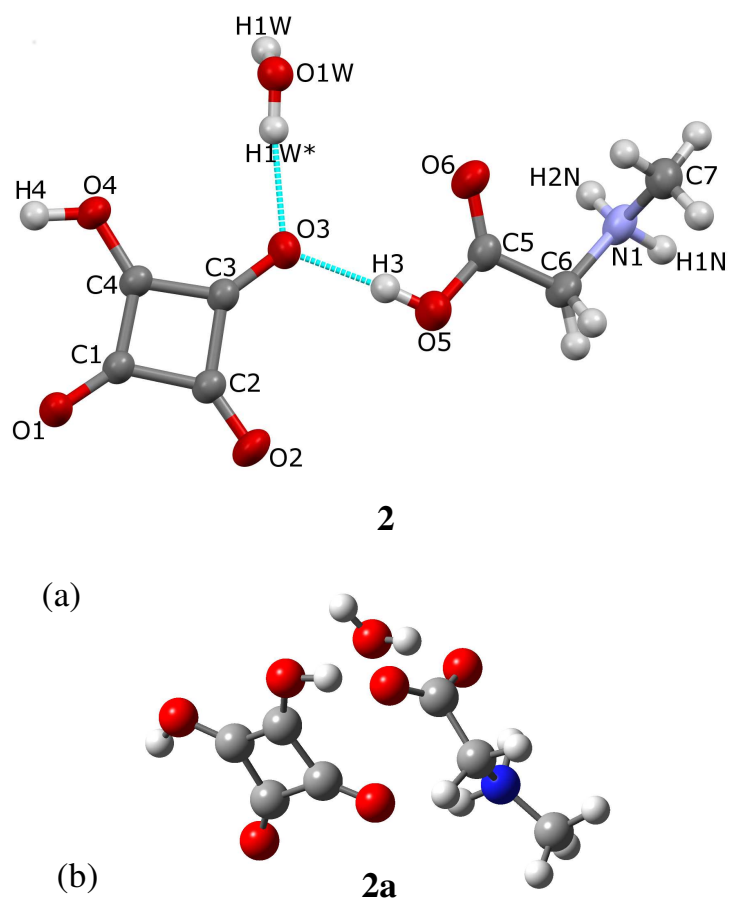
**Fig 1.** Molecular structures of glycine (GLY), sarcosine (SAR), dimethylglycine (DMG), betaine (BET), squaric acid (H<sub>2</sub>SQ), hydrogen squarate anion (HSQ<sup>-</sup>) and squarate dianion (SQ<sup>2-</sup>).



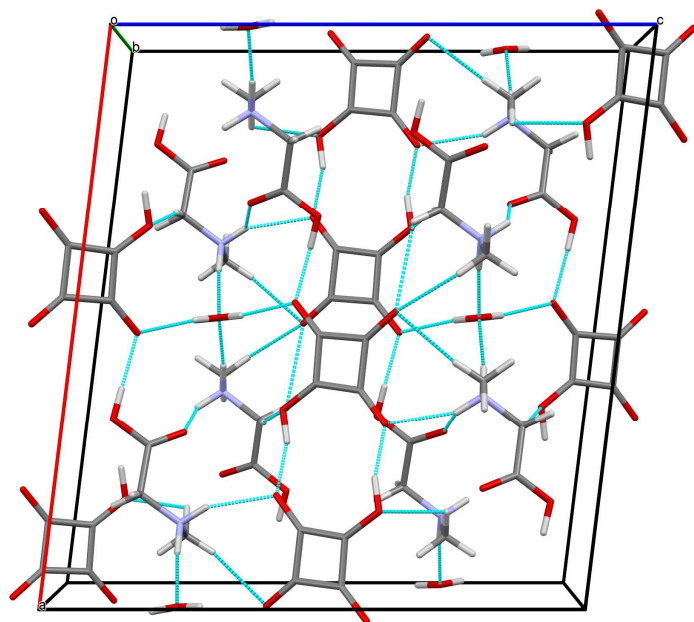
**Fig. 2.** (a) Two glycinium cations and squarate dianion in the crystal structure (**1**) and the atomic labeling scheme. Thermal parameters are shown at the 50% probability level and H-bonds are indicated as the dotted lines; (b) the optimized structure of **1a**, by the B3LYP/6-311G(d,p) approach showing two uncharged glycine molecules and squaric acid.



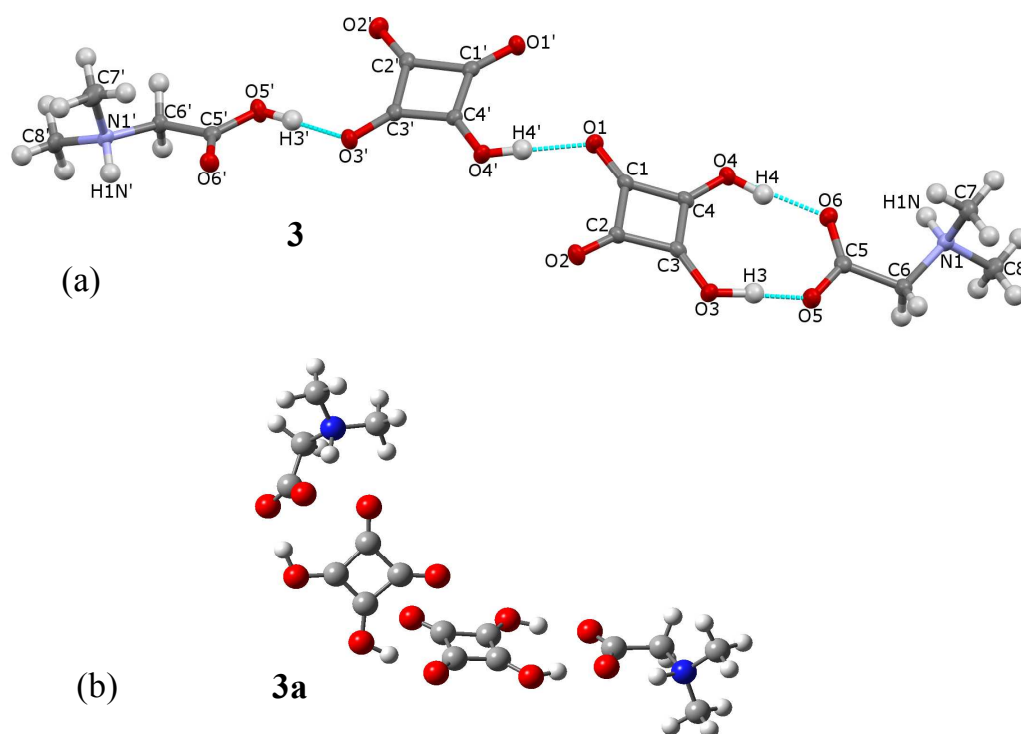
**Fig. 3.** Autostereogram of crystal structure **1** viewed down the direction [010]. The hydrogen bonds are indicated as the dotted lines.



**Figure 4.** (a) The hydrogen squarate anion, sarcosinium cation and water molecule in the crystal structure **2** and the atomic labeling scheme. All atoms are at general position, except O(1W) located on the 2-fold axis, hence only H(1W) of the water molecule is symmetry-independent. H1W\* is at equivalent position  $(-x, y, 1.5-z)$ . Thermal parameters are shown at the 50% probability level and H-bonds are indicated as the dotted lines; (b) the optimized structure of **2a**, by the B3LYP/6-311G(d,p) approach showing sarcosine, squaric acid and water molecules.



**Fig. 5.** Two hydrogen-bonded cyclamers  $R_{\frac{3}{2}}^2(16)$  and  $R_{\frac{2}{2}}^2(10)$  in the structure of **2**. The O-H $\cdots$ O and N-H $\cdots$ O bonds are indicated as dotted lines.



**Fig. 6.** (a) The crystal structure of the dimer of dimethylglycine-squaric acid and dimethylglycinium-hydrogen squarate, **3**, and its atomic-labeling scheme. All atoms are at general positions and the asymmetric part of cell contains hydrogen squarate anion, squaric acid, N-dimethylglycinium cation and N-dimethylglycine as zwitterion. Thermal parameters are shown at the 50% probability level and H-bonds are indicated as the dotted lines; (b) the optimized structure of **3a**, by the B3LYP/6-311G(d,p) approach showing two squaric acids and two dimethylglycinium zwitterions linked into linear dimer.

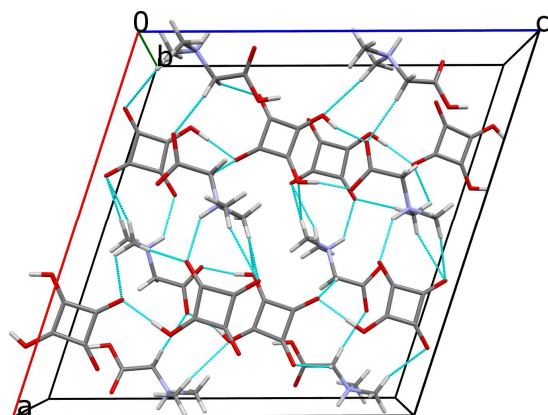
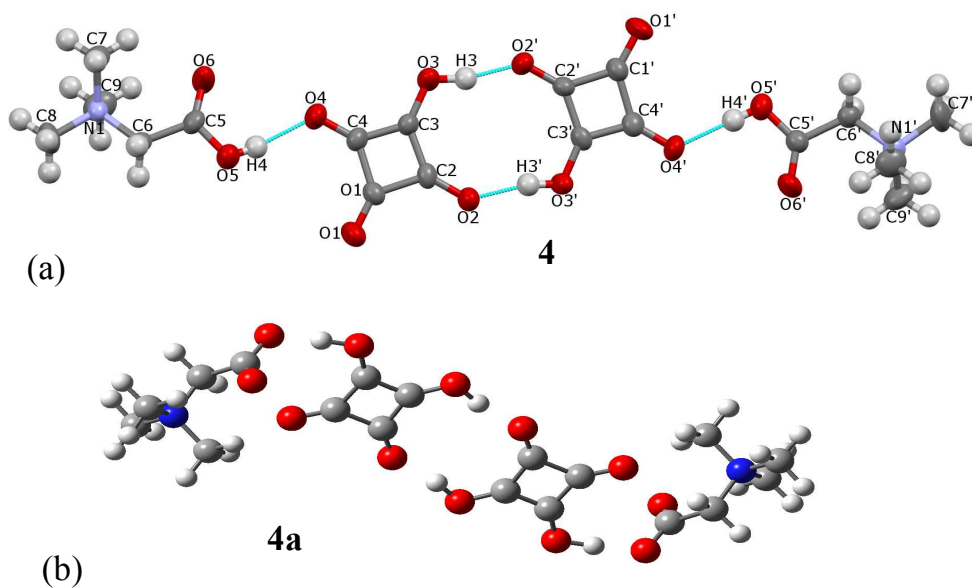
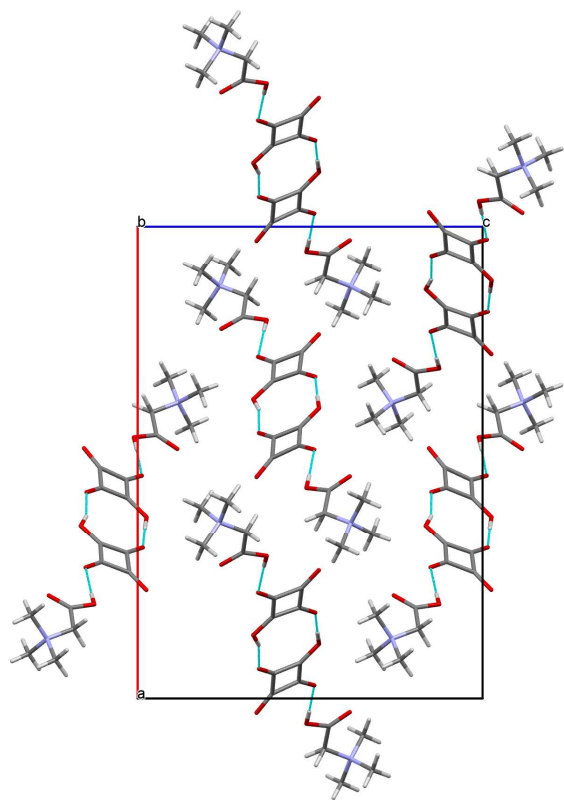


Fig. 7. Two hydrogen-bonded cyclamers  $R_2^2(30)$  in the structure of **3**. The O-H...O and N-H...O bonds are indicated as dotted lines.

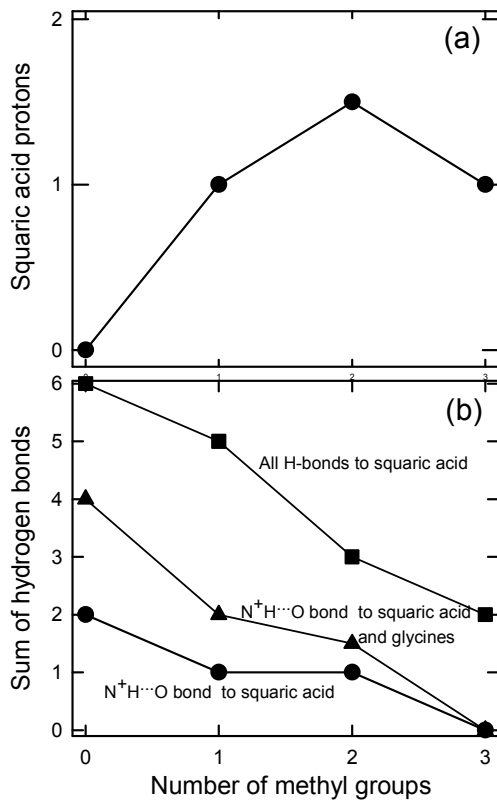


**Fig. 8.** (a) The crystal structure of the H-bonded cyclic dimer of betainium hydrogen squarate, **4**, and its labeling scheme. All atoms are at general positions and the asymmetric part of the unit cell contains two squarate anions and two betainium cations. Thermal parameters are shown at the 50% probability level and H-bonds are indicated as the dotted lines; (b) the optimized structure of **4a**, by the B3LYP/6-311G(d,p) approach showing two squaric acids and two betaines linked into a centrosymmetrical dimer.

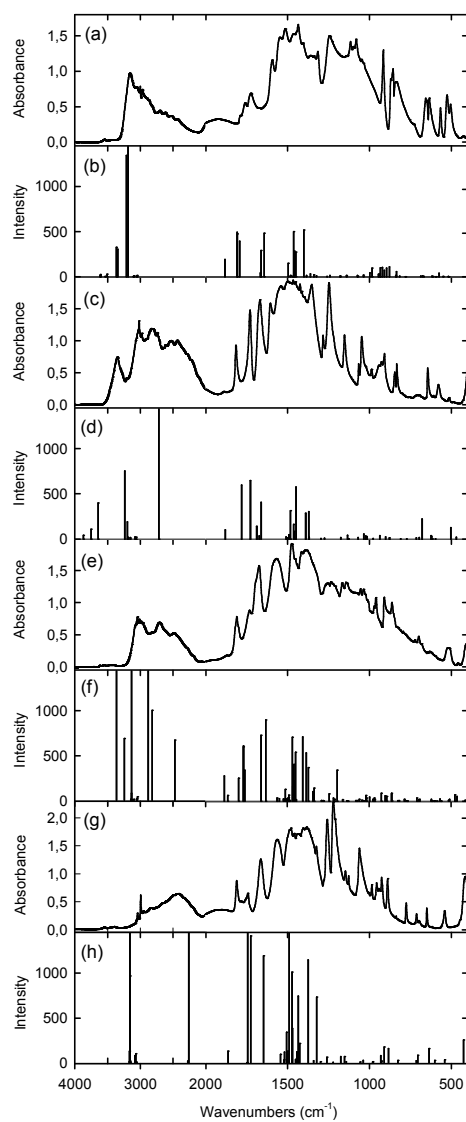




**Fig. 9.** Autostereogram of unit-cell viewed down [010] in **4**. The hydrogen bonds are indicated as dotted lines.



**Fig. 10.** (a) The numbers of protons in squaric acid moiety as a function of the number of methyl group in the glycine derivatives, two protons in  $\text{H}_2\text{SQ}$ , one in  $\text{HSQ}^-$  and zero in  $\text{SQ}^{2-}$ ; (b) The numbers of H bonds ( $\text{N}-\text{H}\cdots\text{O}$ ,  $\text{O}-\text{H}\cdots\text{O}$ ) to squaric acid and glycine derivatives as a function of the number of methyl group in the glycine derivatives; ● -  $\text{N}^+-\text{H}\cdots\text{O}$  bonds to squaric acid moiety, Δ -  $\text{N}^+-\text{H}\cdots\text{O}$  bonds to squaric acid moiety and glycine derivative; ■ - all H-bonds to squaric acid moiety.

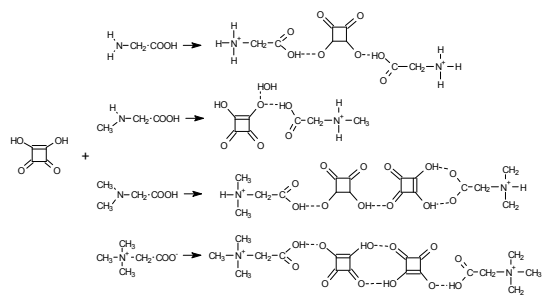


**Fig. 11.** Experimental solid-state spectra and calculated frequencies by the B3LYP/6-311G(d,p) approach (a) and (b) for **1** and **1a**; (c) and (d) for **2** and **2a**; (e) and (f) for **3** and **3a**; (g) and (h) for **4** and **4a**, respectively.

## GRAPHICAL ABSTRACT

**NH $\cdots$ O and OH $\cdots$ O interactions of glycine derivatives with squaric acid**

Michalina Aniola, Zofia Dega-Szafran, Andrzej Katrusiak, Mirosław Szafran



Hydrogen-bonded complexes of glycine, sarcosine, dimethylglycine and betaine with squaric acid have been synthesized and characterized by X-ray diffraction, FTIR, NMR and DFT methods.