## **Crystal structure of a head-to-head bis(cholic)-urea dimer**

Francisco Meijide, Juan V. Trillo, Santiago de Frutos, Cristina Gancedo, Victor H. Soto,<sup>*a*</sup> Luciano Galantini,<sup>*b*</sup> Aida Jover, and José Vázquez Tato\*

Departamento de Química Física, Facultad de Ciencias, Universidad de Santiago de Compostela, Avda. Alfonso X El Sabio s/n, 27002 Lugo, Spain. <sup>a</sup>Escuela de Química, Centro de Investigación en Electroquímica y Energía Química (CELEQ), Universidad de Costa Rica, San José, Costa Rica. <sup>b</sup>Dipartimento di Chimica, Università di Roma "Sapienza", P. le A. Moro 5, 00185 Roma, Italy.

\*Corresponding author: Prof. José Vázquez Tato Facultad de Ciencias, Universidad de Santiago de Compostela, Avda. Alfonso X el Sabio s/n, 27002 Lugo, Spain e-mail: jose.vazquez@usc.es

## Abstract

A head-to-head dimer of cholic acid linked at C3 carbon atoms by a urea bridge was synthesized and the crystal structure was studied. The crystal belongs to the monoclinic crystal system, space group I<sub>2</sub>. The packing shows a bilayer structure with alternating hydrophobic and hydrophilic layers. The horizontal and vertical planes of the two cholic residues are almost parallel to each other, the vertical ones being perpendicular to the bilayers. The hydrogen bond network is two-dimensional as it does not propagate in the direction perpendicular to the bilayers. It is noteworthy that the nitrogen atoms of the urea bridge do not participate in that network. The carbonyl urea group and only one of the ester of the two side chains form hydrogen bonds with two other dimers through the steroid hydroxy O7-H and O12-H groups. This means that each dimer forms eight hydrogen bonds but only four are different.

#### Introduction

Bile acids are natural amphiphilic compounds of great physiological importance.<sup>1</sup> In natural bile acids from mammals, the steroid nucleus have hydroxyl groups, commonly at positions 3, 7, and/or 12 oriented towards its  $\alpha$ -side, while two methyl groups (at positions 10 and 13) are oriented towards the  $\beta$ -side.<sup>2</sup> The amphipathic character of these compounds, due to the hydrophilic/hydrophobic bifacial structure, results in the fact that bile acids are among the most important biological surfactants. The hydroxyl groups and the side chain (supporting a carboxylic group) have made them very attractive building blocks for forming supramolecular structures,<sup>3-</sup> 4 novel antibiotics,<sup>5</sup> new surfactants,<sup>6-8</sup> organogelators,<sup>9</sup> etc. Bile acids derivatives are

especially useful as scaffolds for anion receptors, mainly when helpful substituents are incorporated into their structure.<sup>10-11</sup>

Small<sup>12-13</sup> has proposed that bile salts form two types of micelles in aqueous solution, primary and secondary micelles. Primary micelles are formed by the trihydroxy series of bile salts, free or conjugated with aminoacids, in water or at low salt concentrations. Secondary micelles, mainly formed by dihydroxy derivatives as deoxycholate, are probably formed by the aggregation of primary micelles through hydrogen bonds. However, the formation of hydrogen bonds by bile acids has been mainly studied in solid state,<sup>14-16</sup> and different patterns have been observed depending on the number, location and orientation of the hydroxy groups.<sup>17-18</sup> The effect of the length of the side chain has also been investigated.<sup>19</sup>

On the other hand, N,N'-disubstituted ureas can act as both hydrogen-bond donors through their two NH protons, and acceptors through the lone pairs of the C=O group. This fact has been explored for the construction of crystalline networks, and nanostructured materials.<sup>20</sup> Both groups in urea exhibit a good complementarity resulting in self-association through hydrogen-bonded networks. The crystal of N,N'-diphenylurea is a nice example.<sup>21</sup> Reasons why the interactions of urea derivatives through hydrogen bonding are important have been commented on elsewhere,<sup>20,22</sup> the formation of inclusion compounds being among them.<sup>23</sup> This property is also characteristic of bile acids<sup>24</sup> and it has been used for practical purposes as the enantioresolution of chiral compounds,<sup>25-26</sup> the formation of polymers,<sup>27-28</sup> and delayed release of drugs.<sup>29</sup>

Recently the crystal structures of some head-to-head dimers of methyl esters of cholic and deoxycholic acids have been studied. These dimers, with different bridges linking their C3 positions, lead to different conformations of the two steroid units in the dimers and also to different hydrogen bonding networks and molecular packings, affecting the inclusion capacity of the crystals.<sup>30</sup> The ability of forming hydrogen bonds by cholic acid and the characteristics of head-to-head dimers has been used for designing a crystal in which a single water molecule is encapsulated between two cholic residues in an ice-like structure.<sup>31</sup>

All this information suggests that interesting properties could arise from a dimer compound in which the characteristics of urea and cholic acid were put together in a single molecule. Thus we have selected a simple urea bridge which provides planarity and a shorter distance between the steroids, which *a priori* will restrict the spatial orientation of the steroid nucleus. Once the 3-aminoderivative of cholic acid is formed, the synthesis of the urea dimer (Figure 1) requires only one step.



**Figure 1**. Chemical structure and conventional numbering of the diester derivative of the bis(cholic)-urea dimer (C-urea-C) studied in this paper.

## Experimental

### Synthesis

The synthesis of ureas by reaction of amines with di-*tert*-butyl dicarbonate has been described by Knölker *et al.*<sup>32</sup> The reaction is performed in the presence of 4-dimethylaminopyridine as catalyst. The reaction mechanism goes through the formation of isocyanates which being *in situ* trapped by an additional equivalent of an amine provides access in high yields to ureas.<sup>33</sup> Here we have tried a similar synthesis by using diethylpyrocarbonate (DEPC) as the dicarbonate and triethylamine (TEA) as catalyst (see scheme).



## Scheme

DEPC (0.45 mL, 2.96 mmol) is added to a dispersion of methyl ester of the 3βaminocholic acid<sup>34</sup> (1.16 g, 2.75 mmol) in 25 mL of dried DMF, under argon and stirred for 10 min. The mixture is cooled to 0°C and TEA (1 mL) is added dropwise. The stirring is maintained at this temperature for 45 min and at r.t. for 24 h. Water (300 mL) is added and the precipitate is filtered, washed with water and dried in a vacuum oven. The product is purified by column chromatography with ethyl acetate:methanol (9:1) as eluent. The solvent is removed in a rotavapor, and the solid dried in a vacuum oven at 60°C. Yield: 46 %. <sup>1</sup>**H** NMR (DMSO; 500 MHz, δ/ppm): 3.86 (d, H<sub>3</sub>); 3.76-3.75 (m, H<sub>7</sub>, H<sub>12</sub>); 0.92 (d, H<sub>21</sub>); 0.87 (s, H<sub>19</sub>); 0.59(s, H<sub>18</sub>). <sup>13</sup>**C** NMR (DMSO; 125 MHz, δ/ppm): 174.3 (O=C<sub>ester</sub>); 157.7 (O=C<sub>urea</sub>); 71.9 (C<sub>12</sub>); 67.1 (C<sub>7</sub>); 51.7 (C<sub>3</sub>); 23.7 (C<sub>19</sub>); 17.7 (C<sub>21</sub>); 13.0 (C<sub>18</sub>). NMR spectra are almost identical to those published in the literature for similar derivatives and do not deserve further comments.

MALDY-TOF: 869.6  $(M^+)$ ; theor.: 869.2. Melting point: 215° (with decomposition).

## X-ray diffraction

Data for the colorless prismatic crystal of the compound were collected on a Bruker Smart-CCD-1000. Molecular graphics were from Mercury (<u>http://www.ccdc.cam.ac.uk/prods/mercury</u>). Cif files are available as electronic supporting material. CCDC 952059 contains the supplementary crystallographic data for the crystal. These data can be obtained free of charge from The Cambridge Crystallographic Data Center via <u>www.ccdc.cam.ac.uk/data\_request/cif</u>.

# **Results and discussion**

The compound was recrystallized from ethyl acetate-methanol, and the crystal belongs to the monoclinic crystal system, space group  $I_2$ . A summary of the crystal data and experimental details are listed in Table 1.

Empirical formula	$C_{51}H_{84}N_2O_9$
Formula weight	869.2
Temperature (K)	100
Wavelength (Å)	1.5418
Crystal system, space group	monoclinic I2
a (Å)	14.1521(5)
b (Å)	7.8050(3)
c (Å)	50.058(2)
α, γ (°)	90
β (°)	95.461(2)
Volume (Å <sup>3</sup> )	5504.1(4)
Z, calculated density $(g/cm^3)$	4, 1.049
Absorption coefficient (mm <sup>-1</sup> )	0.561
F (000)	1904
Crystal size (mm <sup>3</sup> )	0.30×0.15×0.02
Theta range (data collection) (°)	1.8-65.5

**Table 1**. Crystal data, data collection and refinement for the C-urea-C dimer.

	-16≤h≤16		
Index ranges	-9≤k≤8		
	0≤l≤58		
Data/restraints/parameters	9055/14/584		
Goodness-of fit on F <sup>2</sup>	1.262		

A view along the *b* crystallographic axis (Figure 2) shows that the  $\alpha$ -faces of the two steroid units of the dimer are oriented in the same direction, and place opposite to other ones (with a  $\beta$ -interdigitation between C18 and C19 methyl groups), thus forming a bilayer in which markedly hydrophobic and hydrophilic zones alternate. The structure is stabilized by hydrogen bonds (see below).



**Figure 2.** View along the *b* crystalographic axis of the molecular packing in the C-urea-C dimer. Hydrogen bonds are indicated in red and blue dot lines.

In a former paper, we have established that three torsion angles  $\varphi_i$  (i=1-3), and two common angles,  $\theta_i$  (i=1-2) were enough for describing the relative orientation in the space of the two bile acid residues of the dimer.<sup>30</sup> For defining the angles, the C3, C10, and C13 carbon atoms were used (see Figure 3 top). It must be taken into consideration that each angle C3-C10-C13 corresponds to the rigid steroid nucleus and is almost constant with an average value of 124±3° (estimated from several bile acid crystals), values being 123.3° and 124.5° for C-urea-C. The nomenclature *e* (eclipsed), *g* (gauche) and *t* (anti) for torsion angle values of 0, 90 and 180° was also adopted. Table 2 shows the values for the mentioned five angles, from which a *tet* conformer results in C-urea-C (Figure 3, top). Conformations of other dimers<sup>30</sup> are also recompiled in Table 3. The two dimers with the *tet* conformation have the belly and back sides oriented towards the same direction while the *ttt* conformer has the belly and back sides oriented in opposite directions.

**Table 2.** Torsion angles ( $\varphi_i$ ; i=1-3), common angles, ( $\theta_i$ ,i=1-2), and sequential distances between the carbon atoms C10<sup>I</sup>-C3<sup>I</sup>-C10<sup>II</sup> for the C-urea-C dimer studied in this paper. Superscripts <sup>I</sup> and <sup>II</sup> refer to both steroid units in the dimer.

Angle	Value (°)	Distance	Value (Å)
$\theta_1(C3^{II}-C3^{I}-C10^{I})$	130.3	$C10^{I}-C3^{I}$	3.052
$\theta_2(C10^{II}-C3^{II}-C3^{I})$	127.7	$C3^{I}-C3^{II}$	4.930
$\varphi_1(C10^{I}-C3^{I}-C3^{II}-C10^{II})$	8.9	$C3^{II}$ - $C10^{II}$	3.060
$\varphi_2(C13^{I}-C10^{I}-C3^{I}-C3^{II})$	175.8	$C10^{I}-C10^{II}$	8.783
$\varphi_3(C3^{I}-C3^{II}-C10^{II}-C13^{II})$	175.5		

**Table 3**. Conformation of C-urea-C and other bile acid dimesr (dimer acronyms are those of the original reference).<sup>30</sup>

Dimer	C-urea-C	C-suc-C	C-Ph-C	DC-Ph-DC
Conformer	tet	ttt	tet	i <sub>et</sub> ti <sub>et</sub>

The two substituents of the urea are in a Z-Z conformation. Because of the planar character of the urea bridge, the angles between horizontal and vertical planes of the two steroid units in the dimer are 5.04 and  $1.12^{\circ}$ , respectively. This last value is in agreement with the values of  $\varphi_2$  and  $\varphi_3$ . The vertical plane of the dimer (calculated from the C3, C10, C13, C18, and C19 carbon atoms of both cholic residues) forms an angle of 89.5° with the bilayer plane. The two steroid residues exhibit some degree of rotation with respect to the urea plane, since the angles between this plane and the vertical ones of the steroids are 28.97 and 30.08° for residues I and II, respectively.

The width of the bilayer is 11.58 Å, very close to that in the C-suc-C crystal (11.30 Å) which also shows a  $\beta$ -interdigitation.<sup>31</sup> The structure also resembles the one observed for an adamantyl derivative of cholic acid used for the enantioresolution of 2-butanol,<sup>26</sup> the width of the bilayer being 11.3 Å with  $\alpha$ -interdigitation.



**Figure 3**. (Top) Torsion and common angles used for defining the relative position of the two steroid residues in the dimer. *tet* conformer in the crystal of the C-urea-C dimer (I and II identify each cholic residue of the dimer). (Bottom) Distinctive distance values (Å) and regular trapezium formed by the  $C10^{I}-C3^{I}-C3^{II}-C10^{II}$  carbon atoms.

The four C10<sup>I</sup>-C3<sup>I</sup>-C3<sup>II</sup>-C10<sup>II</sup> carbon atoms form an almost planar and regular trapezium (Figure 3, bottom) in a close similitude with C-Ph-C.<sup>30</sup> However, now  $d(C10^{I}-C10^{II})$  is larger than  $d(C3^{I}-C3^{II})$  since  $\theta_1$  and  $\theta_2$  are higher than 90° and in C-Ph-C they are lower than 90°. Thus a direct comparison indicates that both trapezia are inverted to each other. The distances d(C10-C3) are in agreement with values for other cholic acid dimers,<sup>30</sup> the average value being  $3.019\pm0.019$  Å.

Hydrogen bonds in the crystal are only established between the hydrophilic parts of the bilayers. This means that the hydrogen bond network does not propagate in a direction perpendicular to the layers (see Figure 2). Each dimer in the crystal is forming hydrogen bonds with other two dimer molecules (Figure 4). The two terminal esters of the reference molecule are asymmetric as far as the hydrogen bonding concerns since only one of the terminal carbonyl groups is linked to the other two dimer molecules. This is to say, one of the oxygen atom of the carbonyl group of an ester is forming two hydrogen bonds while the other one is not involved in the hydrogen bond network. The situation of different hydrogen bonds at the two lateral chains of a dimer is also present in DC-Ph-DC.<sup>30</sup> The reference dimer is located above the other two molecules, arranged side by side, in a "roof" geometry (Fiugre 4 bottom). The result is that only four

different hydrogen bonds exist which involve the carbonyl oxygen atoms of urea and ester groups (each acting as a double acceptor) with the two O7 and O12 hydroxy groups (acting as donors) of the steroid nucleus of the other two dimers. Lengths of these hydrogen bonds are shown in Table 4.



Figure 4. Scheme of the hydrogen bonds in the C-urea-C crystal.

the 4. Hydrogen bond lengths in the C-dred-C erys			
D-H <sup></sup> A	$d(D^{}A)(A)$		
O12-H <sup></sup> O <sub>urea</sub>	2.760 (5)		
O7-H <sup></sup> O <sub>urea</sub>	2.705 (5)		
O12-H <sup></sup> O <sub>ester</sub>	2. 906 (3)		
O7-H <sup></sup> O <sub>ester</sub>	2.933 (2)		

Table 4. Hy	vdrogen	bond	lengths	in the	C-urea-C	crystal
	,	00110 -				

Surprisingly the two nitrogen atoms of the urea do not participate in the hydrogen bond network. These organic groups are located in a highly hydrophobic region, which viewed along the b axis is formed by four A rings and four methyl groups at C10 (two from each dimer molecule). Thus a cavity, with a cross-section of

approximate 21  $Å^2$ , is formed and could capture small hydrophobic guests. Neither ethyl acetate nor methanol, the recrystallization solvents, are included and the urea group has not any available acceptor for forming hydrogen bonds.

The hydrogen bond network is two-dimensional since it propagates perpendicularly to the *ac* plane, which is parallel to the vertical plane of the dimer (see above). This is to say, it does not propagate in the direction perpendicular to the bilayers and hydrophilic regions of different layers are not interconnected by hydrogen bonds (Figures 2 and 4 bottom). Because of the rigidity of the steroid nucleus the properties of this type of bilayer have to be very different from those of bilayers formed by lipids as, for instance, their fluidity. Applications of lipid bilayer membranes on solid supports are well kown,<sup>35-37</sup> and therefore new applications can be expected for these alternative rigid steroid bilayers.

The torsion angles C20-C22-C23-C24 (values in Table 5) show that the lateral chains of the two steroid units adopt fully extended *tttt* conformations. The most significant difference concerns with the ends of the lateral chains. Apart from the different participation of the two terminal ester groups in the hydrogen bond network (see above), another difference appears in their C22-C23-C24-O24b torsion angle values (Table 5). This fact is also illustrated by the angle between the carboxylic plane (C24O24aO24b) and the vertical plane of the steroid nucleus of each cholic residue, the values being 8.46 and 41.44° for the I and II residues, respectively. This fact is related to the orientation that the carbonyl group adopts for forming hydrogen bonds with O7H and O12H by the side chain I, an unnecessary requirement for the side chain II.

Pond sequence	φ (°)			
Bond sequence	Cholic residue I	Cholic residue II		
C13-C17-C20-C22	-170.85	-176.39		
C17-C20-C22-C23	-158.14	-172.59		
C20-C22-C23-C24	-174.07	173.96		
C22-C23-C24-O24a	-173.48	166.83		
C22-C23-C24-O24b	8.26	53.36		

**Table 5**. Relevant torsion angles of the side chains in the C-urea-C crystal.

### Conclusion

Starting from  $3\beta$ -aminocholic acid in its methyl ester form, the synthesis of an urea dimer can be performed in a single step procedure. The resolution of the crystal of the new compound evidences the formation of a bilayer structure with very well defined alternating hydrophobic and hydrophilic layers. The hydrogen bond network is two-

dimensional, being noteworthy that the nitrogen atoms of the urea bridge do not participate in the network. New applications for the rigid bilayers can be foreseen from the applications of classical lipid bilayers supported on a substrate, recently published.

### Acknowledgements

The authors thank the Ministerio de Ciencia y Tecnología, Spain, (Projects MAT2007-61721 and MAT2010-19440) for financial support. J. V. Trillo also thanks for a scholarship.

### References

- (1) Monte, M. J.; Garcia Marin, J. J.; Antelo, A.; Vázquez Tato, J. World J. Gastroenterol. 2009, 15, 804.
- (2) Mukhopadhyay, S.; Maitra, U. Curr. Sci. 2004, 87, 1666.
- (3) Tamminen, J.; Kolehmainen, E. Molecules 2001, 6, 21.
- (4) Nonappa; Maitra, U. Org. Biomol. Chem. 2008, 6, 657.
- (5) Savage, P. B.; Li, C.; Taotafa, U.; Ding, B.; Guan, Q. FEMS microbiology letters 2002, 217, 1.
- (6) Soto Tellini, V. H.; Jover, A.; Meijide, F.; Vázquez Tato, J.; Galantini, L.; Pavel, N. V. Adv. Mater. 2007, 19, 1752.
- (7) Alvarez Alcalde, M.; Jover, A.; Meijide, F.; Galantini, L.; Pavel, N. V.; Antelo, A.; Vázquez Tato, J. *Langmuir* 2008, 24, 6060.
- (8) Álvarez Alcalde, M.; Jover, A.; Meijide, F.; Galantini, L.; Pavel, N. V.; Antelo, Á.; Vázquez Tato, J. *Langmuir* 2009, 25, 9037.
- (9) Nonappa; Maitra, U. Soft Matter 2007, 3, 1428.
- (10) Davis, A. P.; Joos, J.-B. Coord. Chem. Rev. 2003, 240, 143/.
- (11) Davis, A. P. Coord. Chem. Rev. 2006, 250, 2939.
- (12) Small, D. M. Adv. Chem. Ser. 1968, 84, 31.
- (13) Small, D. M. in The Bile Acids, Chemistry, Physiology, and Metabolism; Nair, P. P., Kritchevski, D., Eds.; Plenum Press: New York. 1971, Chapter 8, 249.
- (14) Miyata, M.; Sada, K.; Yasuda, Y.; Miki, K. Mol. Cryst. Liq. Cryst. 1994, 240, 183.
- (15) Miyata, M.; Tohnai, N.; Hisaki, I. Molecules 2007, 12, 1973.
- (16) Miyata, M.; Tohnai, N.; Hisaki, I. Acc. Chem. Res. 2007, 40, 694.
- (17) Jover, A.; Meijide, F.; Soto, V. H.; Vázquez Tato, J.; Rodríguez Núñez, E.; Ton-Nu, H.-T.; Hofmann, A. F. *Steroids* **2004**, *69*, 379.
- (18) Alvarez Alcalde, M.; Jover, A.; Carrazana, J.; Meijide, F.; Soto, V. H.; Vázquez Tato, J. *Steroids* **2007**, *72*, 535.
- (19) Kato, K.; Sugahara, M.; Tohnai, N.; Sada, K.; Miyata, M. Crystal Growth Des. 2004, 4, 263.
- (20) Custelcean, R. Chem. Commun. 2008, 295.
- (21) Dannecker, W.; Kopf, J. R., H. Cryst. Struct. Commun. 1979, 8, 429.

- (22) Obrzud, M.; Rospenk, M.; Koll, A. J. Phys. Chem. B 2010, 114, 15905.
- (23) Hollingsworth, M. D.; Harris, K. D. M. Comprehensive Supramolecular Chemistry 1996, ed. D. D. MacNicol, F. Toda and R. Bishop, Pergamon Press, Oxford, vol. 6.
- (24) Miyata, M.; Sada, K. In Comprehensive supramolecular chemistry; D. D. MacNicol, T., F., Bishop, R., Eds.; Elsevier: Oxford, 1996; Vol. 6, p 147.
- (25) Bortolini, O.; Fantin, G.; Fogagnolo, M. Chirality 2005, 17, 121.
- (26) Miragaya, J.; Jover, A.; Fraga, F.; Meijide, F.; Vázquez Tato, J. Crystal Growth Des. 2010, 10, 1124.
- (27) Miyata, M.; Kitahara, Y.; Osaki, Y.; Takemoto, K. J Inclusion Phenom Macrocyclic Chem 1984, 2, 391.
- (28) Nakano, K.; Sada, K.; Miyata, M. Polymer J. 2001, 33, 172.
- (29) Froemming, K. H.; Sandmann, R. Arch. Pharmaz. 1970, 303, 371.
- (30) Meijide, F.; Trillo, J. V.; de Frutos, S.; Galantini, L.; Pavel, N. V.; Soto, V. H.; Jover, A.; Vázquez Tato, J. *Steroids* **2013**, 78.
- (31) Soto, V. H.; Alvarez, M.; Meijide, F.; Trillo, J. V.; Antelo, A.; Jover, A.; Galantini, L.; Vázquez Tato, J. *Steroids* **2012**, *77*, 1228.
- (32) Knoelker, H.-J.; Braxmeier, T.; Schlechtingen, G. Synlett 1996, 502.
- (33) Knölker, H.-J.; Braxmeier, T. Tetrahedron Lett. 1996, 37, 5861.
- (34) Anelli, P. L.; Lattuada, L.; Uggeri, F. Synth. Commun. 1998, 28, 109.
- (35) Czolkos, I.; Jesorka, A.; Orwar, O. Soft Matter 2011, 7, 4562.
- (36) Bally, M.; Bailey, K.; Sugihara, K.; Grieshaber, D.; Voeroes, J.; Staedler, B. *Small* **2010**, *6*, 2481.
- (37) Jackman, J. A.; Knoll, W.; Cho, N.-J. Materials 2012, 5, 2637.