

Complex hydrogen bonding patterns in bis(2-aminopyrimidinium) selenate monohydrate. Interrelation among graph-set descriptors

Marek Daszkiewicz

Received: 11 March 2011 / Accepted: 19 August 2011 / Published online: 7 September 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract The hydrogen bonding patterns can be described on the basis of the graph-set approach. Since a set of individual hydrogen bonds create the hydrogen bonding pattern, these individuals can be described independently by the graph-set descriptors. The descriptors can be added to each other resulting in a descriptor of the pattern. Algebraic equations based on the graph-set descriptors are given. Alternatively, the hydrogen bonding pattern can be considered as an ordered molecular assembly. Therefore, the proper *elementary* graph-set descriptor, $E_d^a(n)$, associated with the molecules can be utilized to obtain a descriptor of hydrogen bonding patterns. The procedure of the appropriate graph-set summation, leading to the complex pattern, is described.

Keywords Graph sum · Graph-set analysis · Hydrogen bonding pattern · Crystal structure

Introduction

A graph-set for the determination of hydrogen bonding patterns was developed many years ago [1, 2]. The graph-set descriptor, $G_d^a(n)$, designates a pattern of hydrogen bonds, $G = \{D, S, C, R\}$ where a is the number of acceptors, d is the number of donors, and n is the number of atoms in the pattern, also known as the degree of pattern. When the pattern contains only one type of hydrogen bond

it is called a motif and designated by the unitary graph-set. If multiple hydrogen bonds are involved in the pattern, then a complex (binary, ternary, etc.) graph-set describes the pattern, e.g., $C_2^2(6)$ and $C_2^2(10)$ (Fig. 1). This graph-set method was applied into the *Cambridge Structural Database* programs for visualization and characterization of non-covalent networks in molecular crystals [3, 4].

The direction of a hydrogen bond can be designated by an arrow. This concept was introduced by Grell et al. [5, 6]. Two symbols, $>$ and $<$, have subsequently been implemented into the *MERCURY* program [4]. The symbol $>$ means that the hydrogen bond is created from donor to acceptor, and the symbol $<$ has opposite meaning. For example, if the chain pattern $C_2^2(6)>a<b$ is formed by two hydrogen bonds a and b , it means that the hydrogen bond a is constructed from donor to acceptor and the second hydrogen bond b has the same acceptor as a .

However, when the crystal grows, molecules are arranged into consecutive layers and similarly the subsequent hydrogen bonds are created like a wall that is growing during bricklaying. Therefore, the pattern of hydrogen bonds can be considered as a set of individual hydrogen bonds, where the molecular units are involved in the pattern. Since each pattern is described by the graph-set descriptor, $G_d^a(n)$, the question arises whether the individual descriptors can be added to each other resulting in the descriptor of complex graph sets or not. A real example of an existing structure was chosen to answer this question.

In the crystal structure of bis(2-aminopyrimidinium) selenate monohydrate, a complex hydrogen bonding network occurs. On the basis of this particular structure, it is shown that the molecular graph-set descriptors can be added to each other resulting in complex graph-set descriptors for the hydrogen bonding network. Additionally, a mathematical formula is presented.

M. Daszkiewicz (✉)
Institute of Low Temperature and Structure Research,
Polish Academy of Sciences, Okólna str. 2, P. O. Box 1410,
50-950 Wrocław, Poland
e-mail: m.daszkiewicz@int.pan.wroc.pl

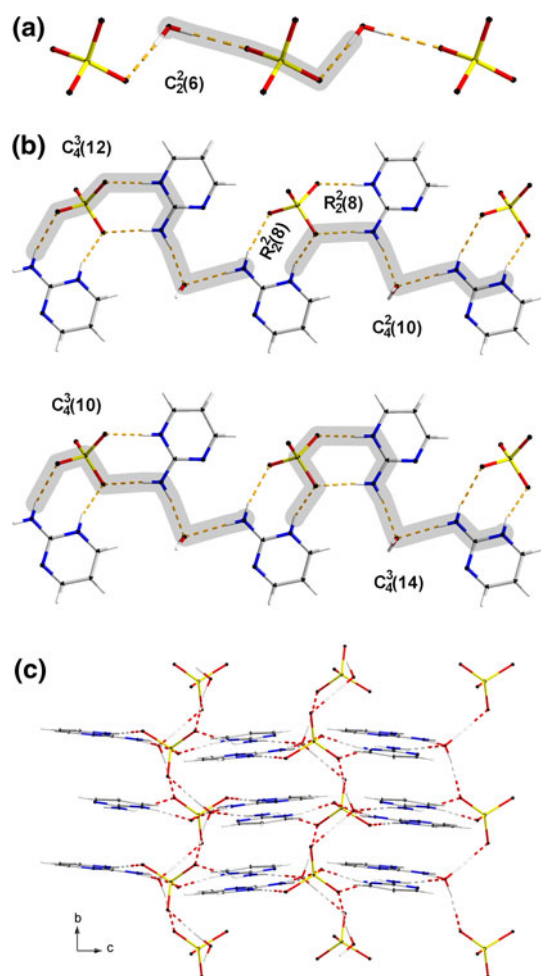


Fig. 1 The selenate anion forms **a** a chain pattern along the b axis and **b** an undulating ribbon-like pattern along the $(\bar{1}01)$ direction, and **c** a cross-section of the patterns viewed along a -axis

Experimental

Synthesis

The starting compounds, 2-aminopyrimidine [Aldrich, purum, $\geq 98\%$ (NT)] and selenic acid (Aldrich, 40 wt% in H_2O , 99.95%) were used as supplied. The dilute acid solution was added to the hot water solution of 2-aminopyrimidine in a 1:1 molar ratio. After the solution was cooled to room temperature, it remained clear without any precipitate. The solution was slowly evaporated at room temperature for several days until it resulted in the formation of good quality single crystals. The following tentative assignments can be done for IR and Raman Spectra. IR (Nujol mull, cm^{-1}): 3272 s, 3066 s ($\nu \text{N-H}\cdots\text{N}$), 3124 s, 3110 s, 3035 s ($\nu \text{C-H}$), 1683 s sh, 1652 vs ($\delta \text{N-H}$), 1626 vs ($\delta \text{H}_2\text{O}$), 1577 m sh, 1561 m, 1538 m, 1507 w ($\nu \text{C-N}$ and $\nu \text{C-C}$), 1293 m ($\delta \text{N-H}\cdots\text{N}$), 1193 w, 1182 w, 1141 w, 1121 vw ($\nu \text{C-N}$ and $\nu \text{C-C}$), 1039 w/1028 w ($\gamma \text{N-H}\cdots\text{N}$),

1008 vw, 1003 vw ($\nu \text{C-C}$), 881 vs sh, 865 vs ($\nu_3 \text{SeO}_4^{2-}$), 876 vs ($\nu_3 \text{SeO}_4^{2-}$ and ring symmetric breathing), 835 vs ($\nu_1 \text{SeO}_4^{2-}$), 698 m, 684 m, 603 w (lib H_2O), 426 vs sh, 420 vs, 407 s, 398 m, 389 m ($\nu_4 \text{SeO}_4^{2-}$), 356 w/345 w ($\nu_2 \text{SeO}_4^{2-}$), 213 w, 188 m, 174 m, 161 m, 154 m, 135 m, 115 w, 95 m, 83 w, 72 w, 68 w, 57 w (lattice vibrations). Raman (cm^{-1}): 3124 vw, 3101 vw, 3093 vw, 3032 vw ($\nu \text{C-H}$), 1626 vw ($\delta \text{H}_2\text{O}$), 1536 w ($\nu \text{C-N}$ and $\nu \text{C-C}$), 1290 vw ($\delta \text{N-H}\cdots\text{N}$), 1191 vw, 1141 vw, 1117 vw ($\nu \text{C-N}$ and $\nu \text{C-C}$), 1009 vw ($\nu \text{C-C}$), 876 vs ($\nu_3 \text{SeO}_4^{2-}$ and ring symmetric breathing), 864 w sh ($\nu_3 \text{SeO}_4^{2-}$), 834 w ($\nu_1 \text{SeO}_4^{2-}$), 422 w, 414 w, 401 vw, 390 w ($\nu_4 \text{SeO}_4^{2-}$), 354 w, 337 vw ($\nu_2 \text{SeO}_4^{2-}$), 198 vw, 156 w, 105 s (lattice vibrations).

Single crystal X-ray diffraction studies

X-ray diffraction data were collected on a KUMA Diffraction KM-4 four-circle single crystal diffractometer equipped with a CCD detector using graphite-monochromatized Mo K_α radiation ($\lambda = 0.71073 \text{ \AA}$). The raw data were treated with the CrysAlis Data Reduction Program (version 1.172.30.3) taking into account an absorption correction. The intensities of the reflection were corrected for Lorentz and polarization effects. The crystal structure was solved by direct methods [7] and refined by full-matrix least-squares method using the SHELXL-97 program [7]. Non-hydrogen atoms were refined using anisotropic displacement parameters. H atoms were placed in calculated positions and allowed to ride on the parent atom. $U_{\text{iso}}(\text{H})$ was set to $1.2U_{\text{eq}}(\text{C}, \text{N})$ or $1.5U_{\text{eq}}(\text{O})$.

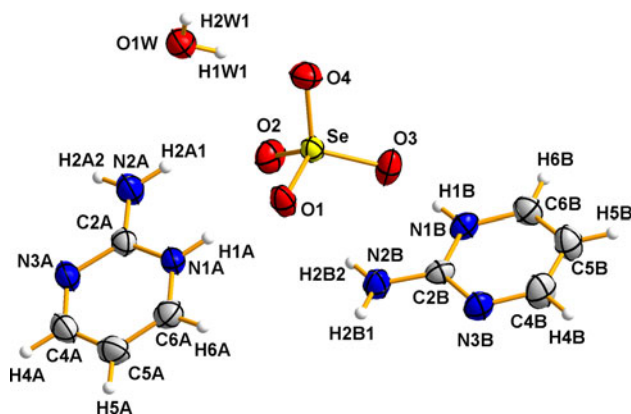
Results and discussion

Crystal structure packing of bis(2-aminopyrimidinium) selenate monohydrate

The studied compound crystallizes in the monoclinic $P2_1/n$ space group (Table 1). An asymmetric unit comprises a water molecule, two 2-aminopyrimidinium cations, Hampy^+ , which balance the negative charge of the selenate anion (Scheme 1). All these species are involved in the hydrogen bonding network (Table 2). The most obvious chain pattern of hydrogen bonds is formed by the water molecules and the selenate anions (Fig. 1a). This chain extends along the b -axis and is designated by the binary graph-set $\text{C}_2^2(6)$. On the other hand, the SeO_4^{2-} anion is also involved in forming two ring patterns with two symmetry-independent organic cations. The ring pattern is generated with the hydrogen bonds formed between the aromatic N-H, amino, and selenate anions, respectively (Fig. 1b). Both rings are described by the binary graph-set

Table 1 Crystal data and structure refinement for bis(2-aminopyrimidinium) selenate monohydrate

Empirical formula	C ₈ H ₁₄ N ₆ O ₅ Se
Formula weight	353.21
Space group	P2 ₁ /n
Unit cell dimensions	<i>a</i> = 14.190(3) Å <i>b</i> = 6.6946(13) Å <i>c</i> = 14.298(3) Å β = 91.24(3)°
Volume	1357.9(5) Å ³
Z, Calculated density	4, 1.728 g/cm ³
Absorption coefficient	2.796 mm ⁻¹
<i>F</i> (000)	712
Crystal size	0.35 × 0.35 × 0.07 mm
θ range for data collection	4.17°–29.13°
Limiting indices	−18 ≤ <i>h</i> ≤ 19 −7 ≤ <i>k</i> ≤ 9 −18 ≤ <i>l</i> ≤ 19
Reflections collected/unique	16,270/3,590 [<i>R</i> (int) = 0.0303]
Completeness to θ	98.2%, 29.13°
Absorption correction	Numerical
Max. and min. transmission	0.819 and 0.416
Refinement method	Full-matrix least-squares on <i>F</i> ²
Data/restraints/parameters	3,590/0/182
Goodness-of-fit on <i>F</i> ²	1.068
Final <i>R</i> indices [<i>I</i> > 2σ(<i>I</i>)]	<i>R</i> ₁ = 0.0257, <i>wR</i> ₂ = 0.0626
<i>R</i> indices (all data)	<i>R</i> ₁ = 0.0336, <i>wR</i> ₂ = 0.0654
Extinction coefficient	0.0071(7)
Largest diff. peak and hole	0.451 and −0.557 e/Å ³

**Scheme 1** ORTEP drawing of bis(2-aminopyrimidinium) selenate monohydrate, with the atom numbering scheme. Displacement ellipsoids for non-H atoms are drawn at the 50% probability level [8]

$R_2^2(8)$ and they have common Se and O(1) atoms from the selenate anion. The two 2-aminopyrimidinium cations and one selenate anion form the structural unit (Hampy)₂SeO₄.

The binary graph-set $C_2^2(6)$ (Fig. 1a) and both the binary graph sets $R_2^2(8)$ (Fig. 1b) intersect at the selenate anion. Additionally, the water molecule, which is a part of the $C_2^2(6)$ chain pattern, also connects the (Hampy)₂SeO₄ units together, resulting in an undulated ribbon-like pattern of hydrogen bonds along $[\bar{1}01]$ direction (Fig. 1b). Therefore, all the species of bis(2-aminopyrimidinium) selenate monohydrate are involved in the ribbon (Fig. 1b). In this complex molecular arrangement four chains of fourth level are found: $C_4^2(10)$, $C_4^3(12)$, $C_4^3(10)$, and $C_4^3(14)$ (Fig. 1b). The last two chains arises from the fact that the Se–O(1) bond simultaneously belongs to two $R_2^2(8)$ rings.

In summary, four chain patterns are incorporated into the ribbon. However, the ribbon results from the intersection of the binary $C_2^2(6)$ and $R_2^2(8)$ patterns rather than four tangled chain patterns. The ribbon propagates along the $[\bar{1}01]$ direction and is connected to the adjacent ribbon in the direction where the $C_2^2(6)$ pattern occurs. Thus, a layered structure of the hydrogen bonding network parallel to the (101) plane is created (Fig. 1c), however, the layers are not interconnected by hydrogen bonds. As a consequence of this layered structure, the crystals of the (Hampy)₂SeO₄·H₂O can be easily cleaved in the (101) plane. It appears that the layered structure results from the tetrahedral structure of the SeO₄²⁻ anion along with its ability to form hydrogen bonds in different directions.

Construction of complex hydrogen bonds

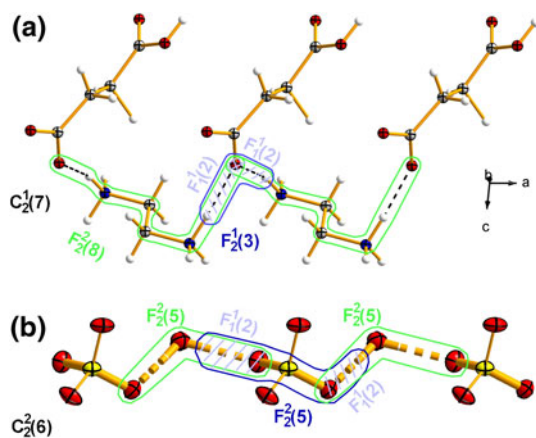
When the crystal grows, molecules are built into successive layers, and subsequently hydrogen bonds are created successively as well. According to the Etter's notation [1], the hydrogen bonding patterns can be described by $G = \{D, S, C, R\}$ graph sets. However, if each hydrogen bond is considered independently, i.e., as a building block which results in a pattern, then each hydrogen bond can be described by a *finite* graph-set, F. Thus, the F graph-set is a subgraph-set of graph-set G and therefore one can say that every chain or ring pattern is constructed by the set of F graph sets, $(F_1, \dots, F_i) \subset G$. For example, if G is represented by the discrete graph-set D, then $i = 1$ and $F = D$, because of the definition of D [2, 5].

A *unitary finite* graph-set $F_1^1(2) = F$ is always found, because in every pattern of hydrogen bonds a hydrogen atom and its acceptor exist. Thus, one can say that the symbol of the complex graph-set G can be obtained by the exact summation of the unitary graph sets F. However, this summation fails, because some atoms between adjacent *finite* patterns exist, which do not belong to any of them. So, they are not included in the complex $G_d^a(n)$ graph-set during summation.

Table 2 Hydrogen bonds (Å, °) for studied compound

	D–H...A	<i>d</i> (D–H)	<i>d</i> (H...A)	<i>d</i> (D...A)	∠(DHA)
$C_2^2(6)$	OW–HW1...O(4)	0.89	1.94	2.8283(19)	177.6
	OW–HW2...O(2) ³	0.80	1.96	2.7405(19)	168.7
$R_2^2(8)$	N(1A)–H(1A)...O(1)	0.86	1.84	2.663(2)	160.6
	N(2A)–H(2A1)...O(2)	0.86	2.19	2.977(2)	151.8
$R_2^2(8)$	N(1B)–H(1B)...O(3)	0.86	1.82	2.673(2)	175.4
	N(2B)–H(2B2)...O(1)	0.86	2.03	2.861(2)	163.0
Connect chain and rings	N(2A)–H(2A2)...OW ¹	0.86	2.11	2.964(2)	171.7
	N(2B)–H(2B1)...OW ²	0.86	2.03	2.877(2)	170.4

Symmetry codes: 1 $-x + 1, -y + 1, -z$; 2 $-x + 1/2, y - 1/2, -z + 1/2$; 3 $x, y + 1, z$

**Fig. 2** The chain pattern of **a** ethylene-1,2-diammonium bis(hydrogen succinate), $C_2^1(7)$, and **b** bis(2-aminopyrimidinium) selenate monohydrate, $C_2^2(6)$

Let us consider a *binary* $F_d^a(n)$ graph-set. In such a case, the most important feature is that it contains two hydrogen bonds at both ends. For instance, $F_2^2(8)$ and $F_2^1(3)$ graphs are parts of a chain (Fig. 2a) found in the structure of ethylene-1,2-diammonium bis(hydrogen succinate) [9]. If only an exact summation of acceptors, donors or the number of atoms associated with the symbols of two graph sets is considered, $F_{2+2}^{2+1}(8 + 3) = C_4^3(11)$, it results in an incorrect descriptor, because the *binary* graph sets $F_2^2(8)$ and $F_2^1(3)$ have a common part, i.e., one hydrogen atom and one acceptor (see Fig. 2a). However, it can be perceived that this common part also creates a graph-set, $F_1^1(2)$, and is included in the chain pattern twice. On one hand, $F_2^2(8)$ has a common part with the $F_2^1(3)$, while the $F_2^1(3)$ intersects with the next $F_2^2(8)$. Therefore, the $F_1^1(2)$ graph-set ought to be subtracted twice. As a result, a simple mathematical operation $F_2^2(8) + F_2^1(3) - 2F_1^1(2)$ gives the proper graph-set: $F_{2+2-2 \times 1}^{2+1-2 \times 1}(8 + 3 - 2 \times 2) = C_2^1(7)$.

These facts can be summarized as follows:

1. When two *binary* $F_d^a(n)$ graph sets are added to each other, then the $F_1^1(2)$ graph-set should be subtracted

2. The beginning and the end clearly define a *finite* pattern, which is described by an *open*, i.e., *finite*, graph-set. The $F_1^1(2)$ graph-set should then be subtracted $m - 1$ times from m added *binary* $F_d^a(n)$ graph sets according to the following general equation:

$$G_d^a(n) = \sum_{i=1}^m F_{d_i}^{a_i}(n_i) - (m - 1)F_1^1(2), \text{ where } G = \{F\} \quad (1)$$

3. Ring and chain patterns always create a *closed* graph-set from the topological point of view—the end is connected with the beginning. Therefore, the $F_1^1(2)$ graph-set should in this instance be subtracted m times according to the following general equation:

$$G_d^a(n) = \sum_{i=1}^m F_{d_i}^{a_i}(n_i) - mF_1^1(2), \text{ where } G = \{C, R\}. \quad (2)$$

For the bis(2-aminopyrimidinium) selenate monohydrate, several *finite* F graph sets construct chains and rings (2apy = 2-aminopyrimidinium ion):

1. $F_2^2(5)_{\text{water}} + F_2^2(5)_{\text{selenate}} - 2F_1^1(2) = C_2^2(6)$ chain (Figs. 1a, 2b),
2. $F_2^1(3)_{\text{selenate}} + F_2^2(5)_{2\text{apy}} + F_2^1(3)_{\text{water}} + F_2^2(7)_{2\text{apy}} - 4F_1^1(2) = C_4^2(10)$ chain (Fig. 1b),
3. $F_2^2(5)_{2\text{apy}} + F_2^2(5)_{\text{selenate}} + F_2^2(7)_{2\text{apy}} + F_2^1(3)_{\text{water}} - 4F_1^1(2) = C_4^3(12)$ chain (Fig. 1b),
4. $F_2^2(7)_{2\text{apy}} + F_2^2(5)_{\text{selenate}} - 2F_1^1(2) = R_2^2(8)$ both rings (Fig. 1b), etc.

Direction of hydrogen bonds in graph-set descriptors

If the direction of the hydrogen bond is relevant for one's purposes, the above algorithm for seeking complex patterns can also be used. However, the symbol m preceding the $F_1^1(2)$ descriptor should then not be used in Eqs. 1 and 2, because different *finite* patterns $F_1^1(2)$ will have different

directionality. For instance, in bis(2-aminopyrimidinium) selenate monohydrate the following chain descriptor with graph-set $C_4^2(10)$ is created and the *finite* patterns are: $F_2^1(3)>a<f$, $F_2^2(5)<f>e$, $F_2^1(3)>e<c$, $F_2^2(7)<c>a$, and unitary $F_1^1(2)<f$, $F_1^1(2)>e$, $F_1^1(2)<c$, and $F_1^1(2)>a$ (Scheme 2a). All these $F_d^a(n)$ graph sets should be written here in the order of atomic path and the directionality descriptors should also be subtracted in the order associated with the atomic path, because this is the only way the correct chain descriptor can be obtained. Thus, the use of this algorithm results in a chain descriptor $C_4^2(10)>a<f>e<c$.

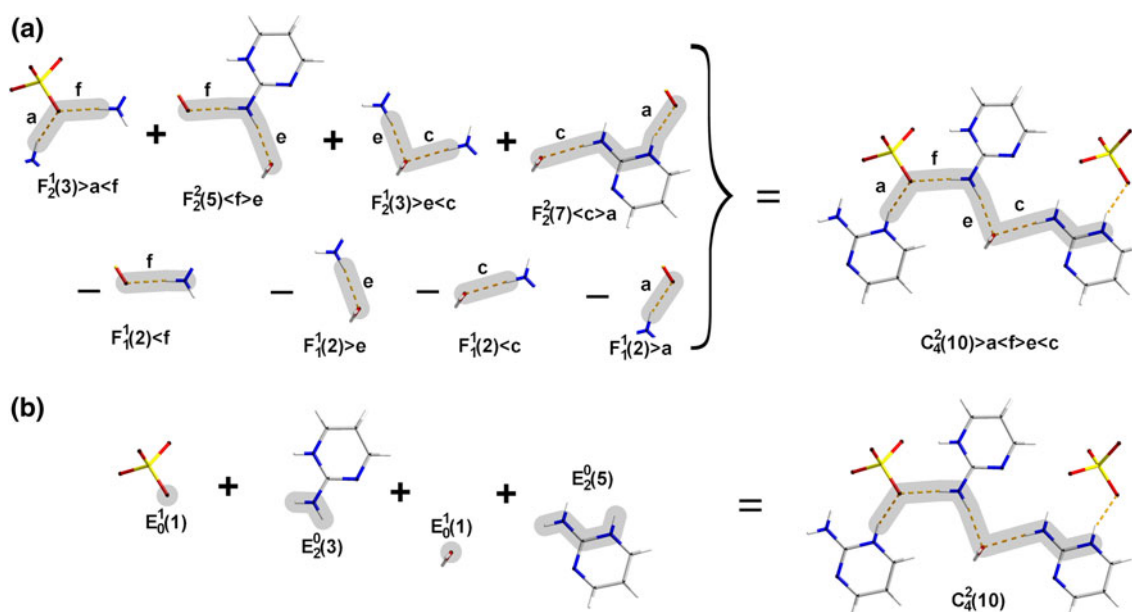
Predicting a graph-set descriptor

A complex graph-set of hydrogen bonding patterns can be constructed by individual molecular fragments taking into account presumable acceptor and donor sites. Let us now consider an isolated molecule where a simple atomic path runs through the molecule from one acceptor/donor atom to another one. For instance, a three-atomic path H–O–H and two two-atomic paths H–O are found in a water molecule. Now, all the atomic paths can be described by *elementary* graph sets. Let these be $E_d^a(n)$, distinguishing this molecular graph-set from the hydrogen bonding graph-set $F_d^a(n)$, which describes an interaction. In the $E_d^a(n)$ symbol, a and d are associated with the number of acceptors and donors located only at both ends of atomic path, respectively. So, they do not define the total number of acceptors and donors located in the whole path and thus $(a, d) = \{0, 1, 2\}$. As a consequence of this definition, the $E_d^a(n)$ is a two-node

graph-set and the following graph sets can be found in each molecule $\{E_0^1(1), E_1^0(1), E_0^2(n), E_2^0(n), E_1^1(n)\}$. For instance, three-atomic path H–O–H and the two two-atomic paths H–O are described by *elementary* graph sets $E_2^0(3)$ and two $E_1^1(2)$, respectively. Additionally, a one-atomic path can also be described, because one acceptor/donor atom itself can participate in a bifurcated hydrogen bonding interaction. Thus, the one-atomic *elementary* graph sets are $E_0^1(1)$ and $E_1^0(1)$. In the case of the SeO_4^{2-} anion, four oxygen atoms are treated as presumable acceptor sites in the hydrogen bonding network. So, in the selenate anion four one-atomic paths and six three-atomic paths are found which are described by graph sets $E_0^1(1)$ and $E_0^2(3)$, respectively.

Let us now consider the water molecule and the selenate anion together to find out what kind of hydrogen bonding patterns can be constructed by those two species. A set of *elementary* graph sets for H_2O and SeO_4^{2-} are $G_1 = \{2E_1^0(1), E_0^1(1), 2E_1^1(2), E_2^0(3)\}$ and $G_2 = \{4E_0^1(1), 6E_0^2(3)\}$, respectively. Now, a graph-set descriptor of a particular hydrogen bonding pattern can be constructed summing the *elementary* molecular graph sets G_1 and G_2 . It is worth noting here that addition of the molecular graph sets is a mathematical expression which corresponds to the creation of the hydrogen bonding interactions. For H_2O and SeO_4^{2-} , the following hydrogen bonding patterns are obtained:

1. $E_1^0(1) + E_0^1(1) = \{C_1^1(2), D_1^1(2)\}$ (Fig. 3a, b),
2. $E_1^0(1) + E_0^2(3) = \{C_1^2(4), R_1^2(4)\}$ (Fig. 3c, d),
3. $E_2^0(3) + E_0^1(1) = \{C_2^1(4), R_2^1(4)\}$ (Fig. 3e, f),



Scheme 2 Complex chain $C_4^2(10)$ results from **a** the summation of the *binary finite* graph sets $F_d^a(n)$ based on Eq. 2 (direction of hydrogen bonds is included) and **b** the summation of the *elementary* graph sets $E_d^a(n)$

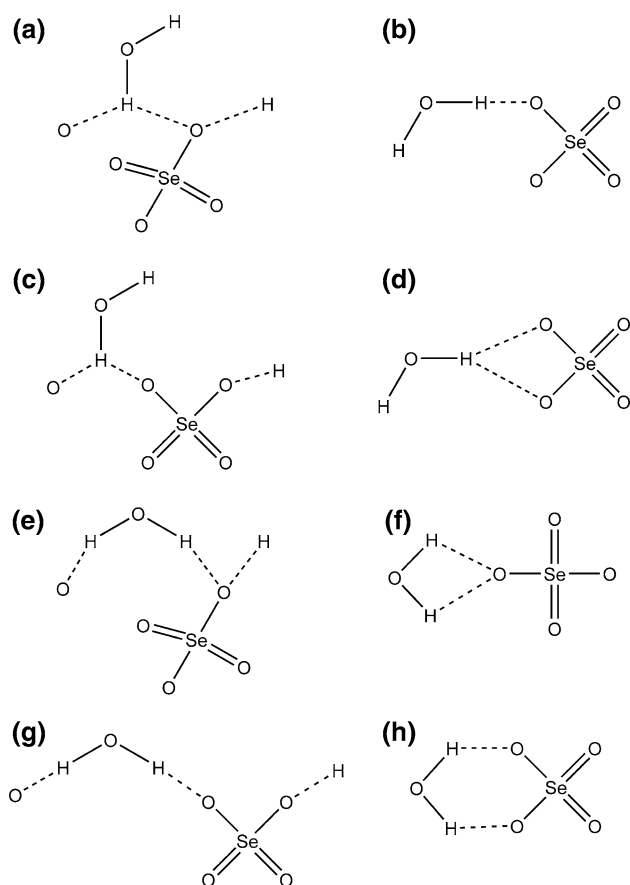
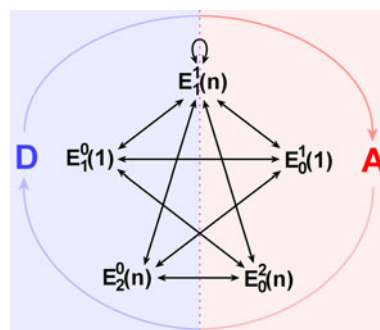


Fig. 3 Chain and ring patterns created by the water molecule and the selenate anion **a** $C_1^1(2)$, **b** $D_1^1(2)$, **c** $C_1^2(4)$, **d** $R_1^2(4)$, **e** $C_2^1(4)$, **f** $R_1^1(4)$, **g** $C_2^2(6)$, **h** $R_2^2(6)$

$$4. E_2^0(3) + E_0^2(3) = \{C_2^2(6), R_2^2(6)\} \text{ (Fig. 3g, h).}$$

Overall, a search of possible graph-set descriptors for hydrogen bonding patterns consists of the addition of each element of the G_1 set and each element of the G_2 set. However, this procedure sometimes gives an unrealistic result. For instance, the summation of $E_0^1(1)_{\text{water}} + E_0^2(3)_{\text{selenate}} = \{C_0^3(4), R_0^3(4)\}$ is unrealistic, because the resulting patterns do not involve donors. Thus, the question arises now: *When does the summation give a real result?* Since the molecules organize themselves in the order *donor atom* \rightarrow *acceptor atom*, the molecular graph sets $E_d^a(n)$ should also have the same order. Therefore, two molecules should create chain or ring patterns as follows:

- $E_1^0(1) + E_0^1(1) = C_1^1(2)$ —in this case careful consideration of the chain pattern only makes sense because a “ring” is *de facto* a discrete pattern, $D_1^1(2)$,
- $E_1^0(1) + E_0^2(n) = \{C_1^2(n+1), R_1^2(n+1)\}$,
- $E_2^0(n) + E_0^1(1) = \{C_2^1(n+1), R_2^1(n+1)\}$,
- $E_2^0(n_1) + E_0^2(n_2) = \{C_2^2(n_1+n_2), R_2^2(n_1+n_2)\}$.



Scheme 3 Allowed summations of elementary graph sets

$$5. E_1^1(n_1) + E_1^1(n_2) = \{C_2^2(n_1+n_2), R_2^2(n_1+n_2)\}.$$

In the light of the above considerations, the order of the molecular graph sets seems to be trivial. However, if it is not preserved, then the chain or the ring pattern of the hydrogen bonding pattern is not created. For instance, $E_0^1(1)_{\text{water}} + E_2^0(5)_{2\text{apy}} + E_0^1(1)_{\text{selenate}} + E_2^0(3)_{2\text{apy}}$ graph sets result in the chain $C_4^2(10)$ depicted in Fig. 1b and Scheme 2b. If they are permuted $E_0^1(1)_{\text{water}} + E_0^1(1)_{\text{selenate}} + E_2^0(5)_{2\text{apy}} + E_2^0(3)_{2\text{apy}}$, they do not create a chain graph-set because the water molecule and the selenate anion do not form a hydrogen bond here. The same logic can be applied for the two 2-aminopyridinium cations.

In summary, a set of molecular graph sets can be found in each molecule $G_i = \{E_0^1(1), E_1^0(1), E_0^2(n), E_2^0(n), E_1^1(n)\}$, where i is the number of interacting molecules. If many molecules interact with each other resulting in a chain or ring pattern, then the symbols of the molecular graph sets can be added to each other alternatively moving along the arrows from D(onor) to A(ceptor) area presented in Scheme 3. To obtain a descriptor of a chain or a ring pattern the last added *elementary* graph-set must be connected with the first one. Otherwise, a *finite* graph-set is constructed. As a consequence of the movement among the elementary graph sets shown in Scheme 3, it can be observed that the resultant descriptor of a chain or a ring pattern $\{C_d^a(n), R_d^a(n)\}$ always possesses the following feature:

$$|a-d| \leq \begin{cases} \frac{i}{2}, & \text{for even } i \\ \frac{i-1}{2}, & \text{for odd } i \end{cases}$$

Conclusions

A proper graph-set descriptor of hydrogen bonding patterns is usually obtained by simple counting of atoms. However, the pattern is a set of individual hydrogen bonds arranged in accordance with the symmetry elements. These separate hydrogen bonds are described by *finite* graph sets, $F_d^a(n)$, and they can be added to each other resulting in a graph-set

$C_d^a(n)$ or $R_d^a(n)$ associated with complex chain or ring patterns, respectively. This mathematical concept also takes into account an excessive graph-set $F_1^1(2)$ being an intersection of graph sets $F_d^a(n)$, which is subtracted as shown in Eqs. 1 and 2.

In this article, Etter's notation is used for the description of individual molecules by *elementary* graph sets, $E_d^a(n)$. It was shown, that each molecule can be described by the set of five elementary graph sets, $G_i = \{E_0^1(1), E_1^0(1), E_0^2(n), E_2^0(n), E_1^1(n)\}$. The creation of hydrogen bonds is then expressed by the summation of *elementary* graph sets resulting in a graph-set descriptor for a chain or a ring hydrogen bonding pattern. A scheme of allowed summations of graph sets has been given.

However, since this mathematical concept does not differentiate atoms from molecules, some complex patterns cannot be formed in a real crystal because of steric hindrance. For instance, it is impossible to form a $R_2^2(11)$ ring pattern with two amino groups of a 1,4-phenylenediammonium cation and two oxygen atoms of a trifluoroacetate anion, despite the fact that the summation of the *elementary* graph sets yield $E_2^0(8) + E_0^2(3) = R_2^2(11)$. However, a chain graph-set $C_2^2(11)$ can be found in the crystal structure of 1,4-phenylenediammonium trifluoroacetate [10].

Supplementary information

CCDC-707763 contains the supplementary crystallographic data for this paper. These data can be obtained free

of charge via www.ccdc.cam.ac.uk/conts/retrieving.html (or from the Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB21EZ, UK; fax: +44 1223 336033; or deposit@ccdc.cam.ac.uk).

Acknowledgment The author would like to thank Dr M. K. Marchewka for fruitful discussion on the hydrogen bonding interactions.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Etter MC, MacDonald JC, Bernstein J (1990) Acta Crystallogr B 46:256
2. Bernstein J, Davis RE, Shimoni L, Chang N-L (1995) Angew Chem Int Ed Engl 34:1555
3. Motherwell WDS, Shields GP, Allen FH (1999) Acta Crystallogr B 55:1044
4. Macrae CF, Bruno IJ, Chisholm JA, Edgington PR, McCabe P, Pidcock E, Rodriguez-Monge L, Taylor R, van de Streek J, Wood PA (2008) J Appl Crystallogr 41:466
5. Grell J, Bernstein J, Tinhofer G (1999) Acta Crystallogr B 55:1030
6. Grell J, Bernstein J, Tinhofer G (2000) Acta Crystallogr B 56:166
7. Sheldrick GM (2008) Acta Crystallogr A 64:112
8. Brandenburg K (2011) DIAMOND release 3.2 g. Crystal Impact GbR, Bonn, Germany
9. Schreuer J, Munch T (1994) Z Kristallogr 209:32
10. Anderson KM, Goeta AE, Hancock KSB, Steed JW (2006) Chem Commun 2138