Acta Biochimica Polonica

Vol. 45 No. 4/1998

917-928

QUARTERLY

This paper is dedicated to Professor Maciej Wiewiórowski on his 80th birthday

Review

Molecular recognition motifs in cytidinium and 2′-deoxycytidinium salts with composite anions[o]

Mirosław Gilski and Mariusz Jaskólski[✉]

Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland, and Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznań, Poland

Keywords: crystal structure, hydrogen bonding, molecular recognition, supramolecular chemistry, supramolecular synthons, crystal engineering, cytidine, 2-deoxycytidine

In the crystal structures of N3-protonated cytidinium and 2′-deoxycytidinium salts with composite XY_n anions capable of accepting hydrogen bonds through their Y atoms, the dominating motif of cytosinium...anion interactions consists of a pair of hydrogen bonds donated from the $N3^+$–H protonation site and from the exoamino N4–H41 group *cis* to N3, and accepted by two Y centers of one anion. This multipoint recognition pattern is stable and robust and thus can be classified as a supramolecular synthon. In a broader group of N3-protonated, N1-substituted cytosinium salts with composite anions it occurs with 70% frequency. The C5 side of the cytosine ring mimics the $N3^+$–H type synthon and shows a propensity to form an analogous motif in which a C5–H5...Y hydrogen bond replaces the strong $N3^+$–H...Y interaction. Since the C–H...Y bond is much weaker, the secondary motif shows higher deformability and is less frequent (44%).

Supramolecular chemistry is a dynamic and highly attractive field concerned with molecular assemblies formed from covalent units through intermolecular interactions [1]. Organic crystals are particularly attractive objects for supramolecular chemistry as they are perfect supramolecules [2] where molecular recognition is realized on a practically infi-

[o]The research of Mariusz Jaskólski was supported in part by an International Research Scholar's Award from the Howard Hughes Medical Institute.

[✉]Author for correspondence: Mariusz Jaskólski, Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Z. Noskowskiego 12/14, 61-704 Poznań, Poland; tel. (0 61) 852 8503; fax: (0 61) 852 0532; email: mariusz@krystal.amu.edu.pl

Abbreviations: Cyd, cytidine; dCyd, 2′-deoxycytidine; CSD, Cambridge Structural Database.

nite scale at an amazing level of precision [3]. Crystal engineering, cultivated by crystallographers for many years [4–9], has now gained in importance as it has the potential to link the molecular and supramolecular worlds by studying intermolecular interactions in the context of crystal packing. There is no doubt that the most powerful "tool" in the hands of "crystal engineers" is the hydrogen bond. The extraordinary role that hydrogen bonds play in the recognition and assembly of small-molecular (e.g. water) and macromolecular (e.g. DNA base pairing, protein secondary structure) systems has been recognized long ago. The knowledge accumulated throughout the years has revealed many intriguing properties of hydrogen bonds (for instance, wide range of energies, σ and π cooperative effects, decisive role of even apparently insignificant interactions) and their fundamental importance for many structures as well as processes, particularly in the biological world. Among the most fascinating chapters in hydrogen bond research is the question of existence of C–H...Y bonds. Until quite recently, such bonds were the subject of dividing controversy among structural scientists. However, through the work of Taylor & Kennard [10], Desiraju [11], Jeffrey & Saenger [12], Steiner & Saenger [13] and Steiner [14], they are now widely accepted and even believed to play decisive, scale-tipping roles where precise molecular recognition is required. The fact that C–H...Y hydrogen bonds are important, even if subtle, determinants of crystal packing (especially in the absence of stronger "classic" hydrogen bonds) is easy to demonstrate. An excellent example was provided by the structure of 1,N^6-ethenoadenosine hydrochloride [15] which turned out to be isostructural with that of adenosine hydrochloride [16] in spite of the fact that the extra 1,N^6-etheno ring had removed the most important $N1^+$–H donor of the unmodified nucleoside cation. The crystal packing in both structures is dominated by clustering of four nucleoside units around the Cl^- anion which acts as a four-fold hydrogen bond acceptor. As demonstrated by Jaskólski [15], the ethenoadenosine hydrochloride structure retained the same architecture by substituting a C–H...Cl^- bond in place of the missing $N1^+$–H...Cl^- link. More recently, Leonard *et al.* [17] have shown that C–H...O hydrogen bonds in A·T, A·U, and certain non-Watson-Crick base pairs may have a stabilizing effect in base-pairing and in secondary structure formation of nucleic acids. The existence of the third (weaker), C–H...O, hydrogen bond in A·T and A·U base pairs would make them topologically more equivalent to the triple-hydrogen-bonded G·C base pair. Recently, the significance of C–H...Y hydrogen bonds in base-pairing between adenine and a shape-analog of thymine has been demonstrated by Evans & Seddon [18] in their critique of a paper by Moran *et al.* [19] which challenged the role of hydrogen bonds, formed by nucleobases, in molecular recognition. In an elegant argument that returned the subject of hydrogen bonding in DNA into the *status quo*, Evans & Seddon [18] point out that the importance of hydrogen bonding to molecular recognition has profound consequences both for basic theories of biochemistry as well as for the foundations of supramolecular chemistry and crystal engineering. With respect to the protein world, the excellent work of Derewenda *et al.* [20] has shown beyond doubt that C–H...O hydrogen bonds are ubiquitous and important cohesive interactions in the structure of protein molecules and that they may also have implications for some enzymatic mechanisms.

This paper analyzes the packing motifs formed in the crystals of cytidine (Cyd) and 2′-deoxycytidine (dCyd) salts with composite anions including trigonal (XY_3), tetrahedral (XY_4), and octahedral (XY_6) species. It demonstrates that in such structures, most of which have been determined in our laboratory, the dominating interaction motif consists of a pair of hydrogen bonds formed by the biologically important N3 side of the nu-

cleobase. One of these bonds is donated from the $N3^+$-H protonation site and the other from the N4-H41 exoamino function (H41 is *cis* to N3, H42 is *cis* to C5, Fig. 1). The two bonds are accepted by two Y acceptor centers of a single anion. This recognition pattern is robust and forms when the structural ingredients coexist in the crystal structure. It can, therefore, be referred to as a "supramolecular

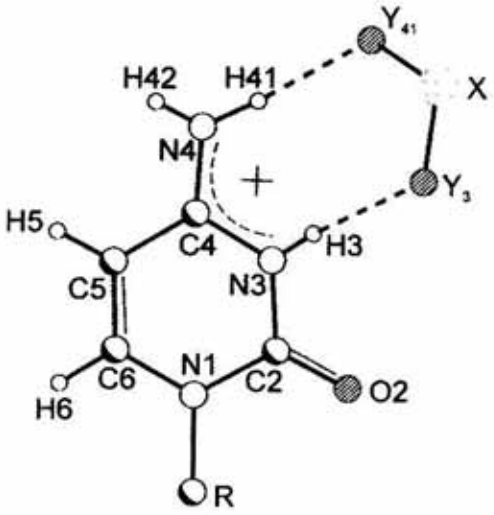

Figure 1. Pairing pattern between $N3^+$-H cytosinium cations and composite XY_n anions.

Throughout the paper, a consistent cytosine atom numbering is used with the exoamino H41 atom *cis* to N3 and H42 *cis* to C5. The interaction pattern (broken lines) represents the primary supramolecular synthon discussed in the paper. The secondary synthon is formed in an analogous way *via* the C5-H and N4-H42 donors of the cytosine unit.

synthon" following the definition of Desiraju [21] who, inspired by the seminal concept of Corey [22], used this term to describe structural units within crystalline (supramolecular) architecture involving recurring recognition patterns in intermolecular interactions. However, as recently shown by Biradha *et al.* [23], robust supramolecular synthons will also form even in the absence of the classic O/N-H donors, provided there are C-H groups which can substitute for them. The present structures provide an excellent opportunity to test this idea as the biologically inactive, C5-H, side of the nucleobase should be capable of forming a minor, or secondary motif mimicking the major $N3^+$-H synthon.

THE PRIMARY SUPRAMOLECULAR SYNTHON

The first observation of a specific interaction between the $N3^+$-H cytosinium cation and a composite anion (XY_n) capable of accepting multiple hydrogen bonds at its several (n) Y acceptor centers, was reported by Viswamitra *et al.* [24] in the structure of 2′-deoxycytidine 5′-monophosphate monohydrate. The nucleotide in that structure exists in zwitterionic form, with N3-protonated cytosine and mono-deprotonated 5′-monophosphate ester group ($-OPO_3H^-$). In the crystal lattice, two nucleotide molecules form a dimer through a pair of hydrogen bonds linking the protonated nucleobase moiety of one nucleotide with the phosphate moiety of the other. In this motif, also referred to as the primary supramolecular synthon, both hydrogen bonds are donated by the nucleobase, from the $N3^+$-H protonation site and the *cis* N4-H41 group. They are accepted by two different O atoms of the same phosphate group (Fig. 1). This multi-point recognition pattern, in which the C base is involved in two hydrogen bonds with another chemical unit, is vaguely reminiscent of its pairing trend displayed in the fullest in the C·G base pair, where three hydrogen bonds exist. A closer similarity (with respect to the nature, but not the number of hydrogen bonds) could be found with a C^+·C base pair formed between a protonated and a neutral C base, in which the protonated component is a donor of two hydrogen bonds, from the $N3^+$-H and N4-H41 sites [25, 26]. In the isostructural dihydrogenphosphate salts of cytidine (Cyd-$H^+ \cdot H_2PO_4^-$ [27]) and 2′-deoxycytidine (dCyd-$H^+ \cdot H_2PO_4^-$ [28]) the cytosinium... phosphate primary supramolecular synthon is also found in spite of the fact that the crystal architecture involves an intriguing pattern of

..$HOPO_2OH$..O5′-sugar-O3′..HO-PO_2OH..

O5′-sugar-interactions between the phosphate and sugar units [27, 28]. We have later found that the primary synthon prevails in all crystal structures where the (2′-deoxy)-cytidinium cation coexists with a composite anion, as in $CydH^+ \cdot NO_3^-$ [29]. The hydrogen bonds within this supramolecular synthon are close to linear (Table 1) and in the case of O acceptors are characterized by average N3...O and N4...O distances of 2.79 Å and 2.84 Å, respectively (Table 1).

The most elegant realization of this supramolecular synthon has been found in the structure of 2′-deoxycytidinium sulfate methanol solvate ($(dCydH^+)_2 \cdot SO_4^{2-} \cdot MeOH$) where the two independent 2′-deoxycytidinium cations form an ion cluster with one common SO_4^{2-} anion [30]. Within this cluster, which is almost perfectly symmetric, each of the cations forms the $N3^+$-H type synthon with the anion. One of the sulfate O atoms serves as a common joint, accepting hydrogen bonds from both N4-H41 donors (Fig. 2). The molecule of methanol solvating this crystal is disordered and involved in complicated interactions with the sugar moieties, forming columns composed of C-OH groups. Later, we have crystallized the same salt, 2′-deoxycytidinium sulfate, in a different (hydrated) form ($(dCydH^+)_2 \cdot SO_4^{2-} \cdot H_2O$) (M. Jaskólski, unpublished). The packing of this crystal is entirely different with the molecule of water of hydration directly intertwined into the pattern of nucleobase...sulfate interactions. In spite of this, the dominance of the $N3^+$-H cytosinium...anion synthon is clearly visible (Fig. 3). One of the cations (B) forms the synthon in its classic form while the other (A) uses the water molecule as a relay in a chain of cooperative N3-H3...OW-H...OSO_3 hydrogen bonds. In this way, the intervening water molecule not only did not disrupt the primary supramolecular synthon formed by

Table 1. Hydrogen bond interactions forming the ($N3^+$-H3, N4-H41)...Y_2X motif in selected cytidinium and 2′-deoxycytidinium salts

Salt		N3-H3 Å	H3...Y_3 Å	N3...Y_3 Å	N3-H3 ...Y_3 (°)	N4-H41 Å	H41 ...Y_{41} Å	N4...Y_{41} Å	N4-H41 ...Y_{41} (°)	Reference
5′-dCMP·H_2O		1.00	1.88	2.874	170	0.91	1.79	2.691	173	[24]
$dCydH^+ \cdot H_2PO_4^-$		0.84	1.84	2.668	168	0.76	2.21	2.939	162	[28]
$CydH^+ \cdot H_2PO_4^-$		0.91	1.85	2.700	155	0.95	2.02	2.945	163	[27]
$CydH^+ \cdot NO_3^-$		0.79	1.93	2.725	174	0.82	2.14	2.956	170	[29]
$(CydH^+)_2 \cdot SO_4^{2-}$	A	0.93	2.08	2.991	164	0.83	1.98	2.802	168	[31]
	B	1.01				0.95				
$(dCydH^+)_2 \cdot SO_4^{2-} \cdot$ MeOH	A	0.96	1.77	2.721	176	0.92	1.87	2.758	162	[30]
	B	1.04	1.83	2.830	160	0.95	1.91	2.838	165	
$(dCydH^+)_2 \cdot SO_4^{2-} \cdot H_2O$	A	0.86	1.98[b]	2.834[d]	172[f]	0.96	1.89	2.786	155	unpubl.
		0.90[a]	1.87[c]	2.764[e]	173[g]					
	B	0.74	2.15	2.843	156	0.91	1.90	2.806	176	
$(dCydH^+)_2 \cdot SiF_6^{2-}$		0.83	1.89	2.718	175	0.95	1.98	2.911	170	unpubl.
$(CydH^+)_2 \cdot SiF_6^{2-}$	A	0.99	1.90	2.751	143	0.80	2.09	2.882	170	unpubl.
	B	1.03	1.83	2.833	164	0.82	2.00	2.802	169	

[a] OW-HW; [b] H3...OW; [c] HW...OSO_3; [d] N3...OW; [e] OW...OSO_3; [f] N3-H3...OW; [g] OW-HW...OSO_3

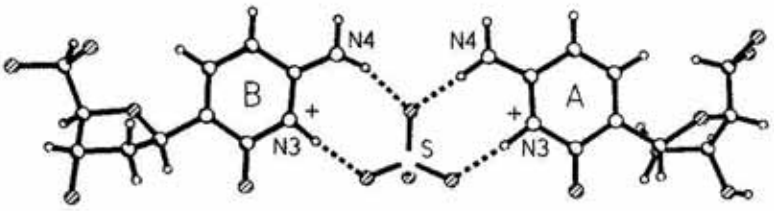

Figure 2. $(dCydH^+)_2 \cdot SO_4^{2-} \cdot MeOH$ – ion cluster involving one sulfate anion and the two independent nucleoside cations.

The two primary ($N3^+$–H type) supramolecular synthons (broken lines) are formed in a nearly symmetric fashion. The disordered nature of the ribose side chains is indicated by two O5′ atoms attached to each C5′ atom [30].

cation A but was incorporated to reinforce it (Table 1).

The structure of cytidinium sulfate ($(CydH^+)_2 \cdot SO_4^{2-}$) [31] is somewhat different. Here we have again two independent cytidinium cations clustered around one common anion. One of the cations (A) forms the classic synthon with the SO_4^{2-} anion. The other cation (B) donates the two hydrogen bonds ($N3^+$–H and N4–H41) to this same anion but they are accepted by one, rather than by two, O atoms leading to significant distortion of the characteristic motif.

THE SECONDARY SUPRAMOLECULAR SYNTHON

Topologically, the H42–N4–C4–C5–H5 fragment of the cytosinium cation is very similar to the H41–N4–C4–$N3^+$–H3 fragment. Geometrically it is thus quite possible that also on the biologically inactive, C5–H, side of the nucleobase a similar multi-point recognition pattern could be formed. However, this secondary synthon would involve a weak C5–H5...Y hydrogen bond instead of the strong $N3^+$–H3...Y bond donated by the

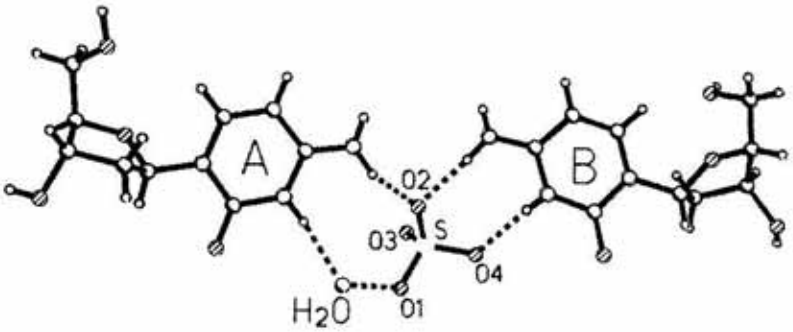

Figure 3. $(dCydH^+)_2 \cdot SO_4^{2-} \cdot H_2O$ – two primary cation...anion supramolecular synthons formed by one sulfate anion.

One of these motifs (A) is formed *via* an auxiliary water molecule.

In the structure of cytidinium hexafluorosilicate ($(CydH^+)_2 \cdot SiF_6^{2-}$) (M. Gilski, unpublished), the SiF_6^{2-} anion again interacts with two independent cytidinium cations (Fig. 4). In both cases the classic $N3^+$–H supramolecular synthons are formed (Fig. 4, Table 1). The structure of 2′-deoxycytidinium hexafluorosilicate ($(dCydH^+)_2 \cdot SiF_6^{2-}$) (M. Gilski, unpublished) is more symmetric. It contains only one independent 2′-deoxycytidinium cation which forms the primary $N3^+$–H type synthon with the two-fold symmetric anion.

protonation site (Table 1) and might not be competitive enough (in energetic terms) in the presence of the primary $N3^+$–H type motif. Formation of the secondary synthon in typical $N3^+$–H salts of cytidine and 2′-deoxycytidine would be automatically disadvantaged because the primary synthon would dictate the crystal packing principle which might be incompatible with the secondary synthon. Therefore, simultaneous presence of both supramolecular synthons would be a strong indication of their robustness and would actu-

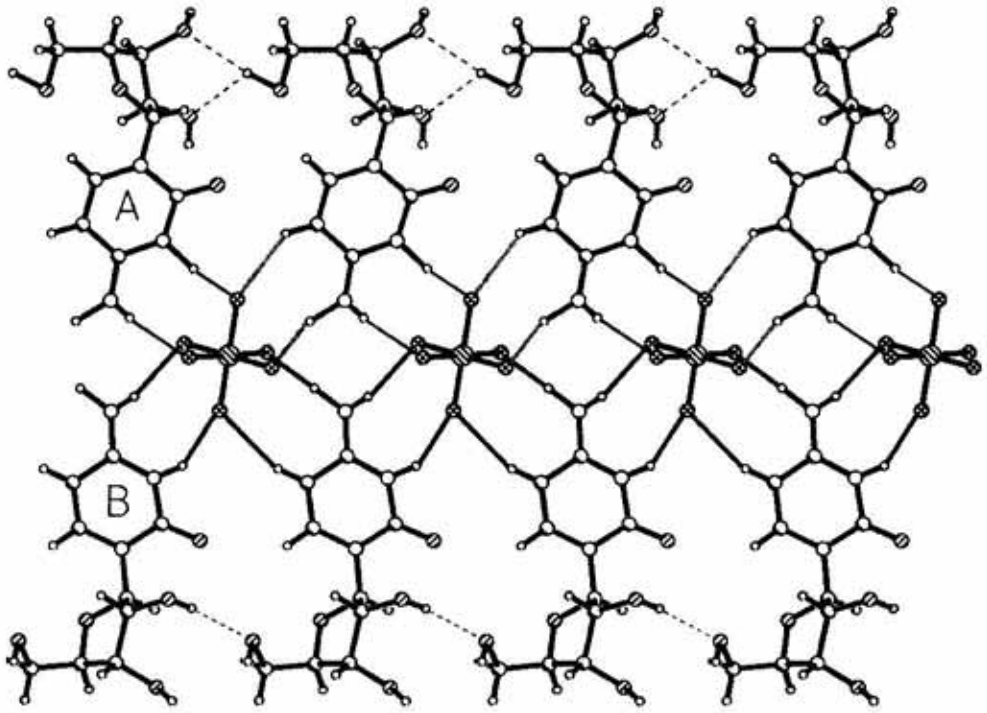

Figure 4. $(CydH^+)_2 \cdot SiF_6^{2-}$ — the primary ($N3^+$–H type, light/dark blue) and secondary (C5–H type, pink/orange) supramolecular synthons formed by the two independent cytidinium cations (A/B).

ally speak in favor of the interactions involving C–H...Y hydrogen bonds.

Out of the several C–H groups in the cytidinium or 2′-deoxycytidinium cation, the C5–H5 group is most likely to act as a hydrogen bond donor. In their MO calculations (using the CNDO/2 method) of net atomic charges in $CydH^+ \cdot NO_3^-$ and $CydH^+ \cdot Cl^-$, Wiewiórowski *et al.* [32] have shown that the charge accumulated in the C5–H5 group (C5 −0.17, H5 +0.08) makes it a very good hydrogen bond donor. For comparison, the strong N–H proton donors are characterized by the following (average) net atomic charges: N3/H3 −0.19/+0.18, N4/H −0.21/+0.18, and the average hydroxyl group by O/H −0.26/+0.16. Among the C–H hydrogen atoms, the highest positive charge is accumulated at H5, the next being H6 (+0.06). However, the C6–H6 group is not likely to be a hydrogen bond donor as the C6 atom has a very high positive net charge (+0.24). The aliphatic C–H groups have all positive net charges at the carbon atoms and net atomic charges at the hydrogen atoms (max +0.04) much lower than H5. It is, therefore, quite obvious from those results, that the highest potential to be a donor in a C–H...Y hydrogen bond resides in the C5–H5 group.

We have analyzed all the structures discussed in the previous section to see how often the secondary supramolecular synthon was formed (Table 2). It is quite surprising that the rate of occurrence of this secondary synthon is quite high. In several salts it accompanies the primary synthon and is realized in all possible cases, as in $(CydH^+)_2 \cdot SiF_6^{2-}$ (Fig. 4, Table 2).

In 5′-dCMP·H_2O, the C5–H synthon cannot be formed because, in the *anti* orientation of the glycosidic bond, the C5–H side of the nucleobase is shielded by the phosphate group of the nucleotide and is, therefore, inaccessible to intermolecular acceptors. An intermolecular N4–H42...O(phosphate) hydrogen bond is, however, formed (Table 2). In the two isomorphous (2′-deoxy)cytidinium dihydrogenphosphates, in $CydH^+ \cdot NO_3^-$, and in the two

Table 2. Hydrogen bond interactions forming the (C5–H5, N4–H42)...Y_2X motif in selected cytidinium and 2'-deoxycytidinium salts

Salt		C5–H5 Å	H5...Y_5 Å	C5...Y_5 Å	C5–H5 ...Y_5 (°)	N4–H42 Å	H42 ...Y_{42} Å	N4 ...Y_{42} Å	N4-H42 ...Y_{42} (°)	Reference
5'-dCMP·H_2O		1.00				1.05	1.94	2.980	173	[24]
$dCydH^+ \cdot H_2PO_4^-$		1.07	2.30	3.290	153	0.97	1.73	2.699	177	[28]
$CydH^+ \cdot H_2PO_4^-$		1.08	2.35	3.339	152	0.91	1.84	2.740	170	[27]
$CydH^+ \cdot NO_3^-$		1.01	2.28	3.279	168	0.87	2.16	3.033	172	[29]
$(CydH^+)_2 \cdot SO_4^{2-}$	A	0.96	2.45	3.249	140	0.99	2.42	3.376	145	[31]
	B	0.96				0.77				
$(dCydH^+)_2 \cdot SO_4^{2-} \cdot$ MeOH	A	0.96				0.98				[30]
	B	0.96				0.99				
$(dCydH^+)_2 \cdot SO_4^{2-} \cdot$ H_2O	A	0.96				0.85				unpubl.
	B	0.96	2.46	3.321	150	0.83	2.09	2.838	151	
$(dCydH^+)_2 \cdot SiF_6^{2-}$		0.96	2.22	3.109	154	0.72	2.23	2.914	159	unpubl.
$(CydH^+)_2 \cdot SiF_6^{2-}$	A	0.93	2.39	3.268	158	1.02	1.84	2.853	176	unpubl.
	B	0.93	2.41	3.194	142	1.06	1.86	2.864	156	
$dCyd_2H^+ \cdot H_2PO_4^-$	A	0.96	2.63	3.268	124	0.85	1.98	2.828	174	[26]
	B	0.96	2.90	3.616	132	0.87	2.00	2.869	176	
	C	0.96	2.83	3.441	122	0.89	1.92	2.805	175	
	D	0.96	2.27	3.121	148	0.80	2.54[a]	3.232[a]	145[a]	
							2.43[a]	3.172[a]	156[a]	

[a] Bifurcated hydrogen bond

hexafluorosilicates, the C5–H supramolecular synthons are formed and are characterized by quite similar geometrical parameters (Table 2). In cytidinium sulfate the secondary synthon is only formed by that cytidinium cation which also forms the primary synthon (A). Cation B forms neither the primary nor the secondary synthon. An interesting situation arises in $(dCyd\ H^+)_2 \cdot SO_4^{2-} \cdot H_2O$ where the secondary supramolecular synthon is not formed by cation A. It is the same cation which on its $N3^+$–H side formed the primary synthon in an unusual way, with an extra H_2O link (Fig. 5). One might, therefore, say that the secondary synthon of cation A was "sacrificed" in order to "rescue" the stronger interactions of the primary synthon and to accommodate the water of hydration. None of the two 2'-deoxycytidinium cations in $(dCyd \cdot H^+)_2 \cdot SO_4^{2-} \cdot$ MeOH forms the secondary supramolecular synthon. However, while cation B does not form any interactions with the anion through the H42–N4–C5–H5 region (N4–H42 is hydrogen bonded to the carbonyl O2B atom in another copy of this cation), the N4–H42/C5–H5 donors of cation A are accepted by a single O atom of the SO_4^{2-} anion.

It could be expected that when the primary supramolecular synthon cannot be formed because of chemical reasons, the ions would enjoy more freedom and would be more likely to be involved in the formation of the secondary synthon. This is, however, not always the case. In the structures of N^3-methylcytidinium nitrate [33] and 3,N^4-ethenocytidinium dihydrogenphosphate [34] the primary synthon cannot be formed because the N^3 site of the cytosine system is blocked by a –C sub-

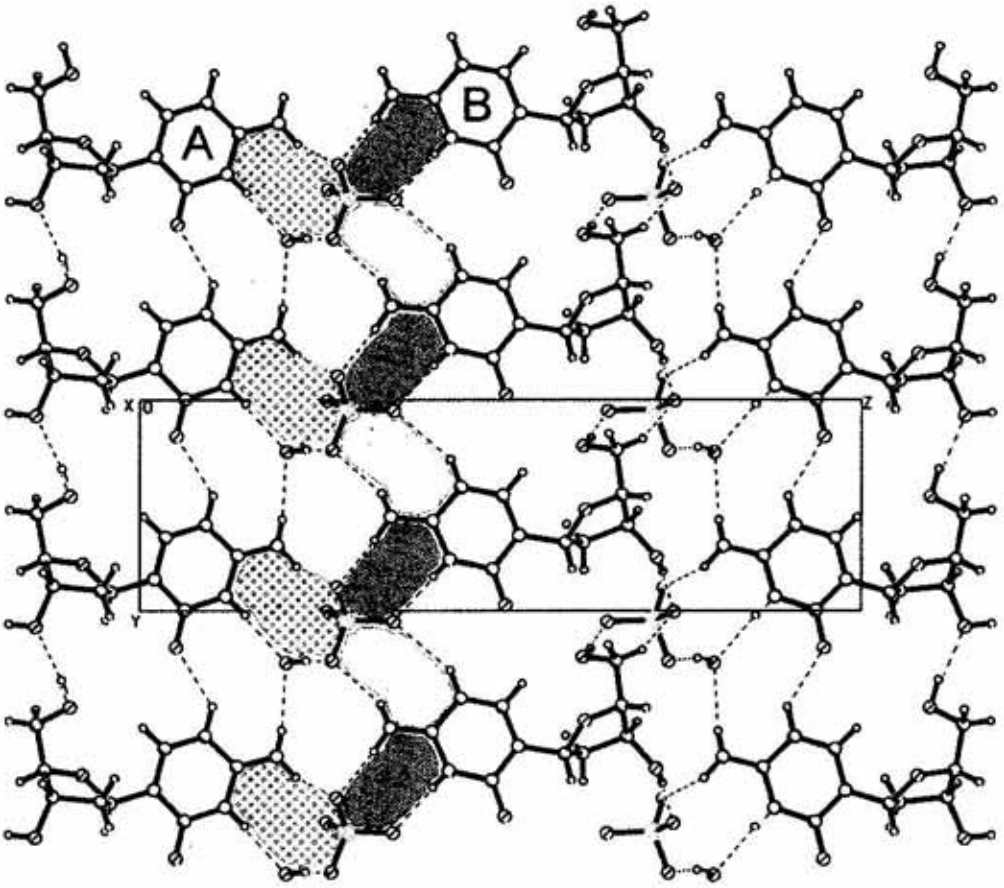

Figure 5. Crystal packing of $(dCydH^+)_2 \cdot SO_4^{2-} \cdot H_2O$.

The primary synthons (also depicted in Fig. 3) are indicated using pink shading: solid color for cation B, stippled color for cation A which forms the primary synthon through an intervening water molecule. The secondary supramolecular synthon formed by cation B is shaded green. Cation A does not form a secondary synthon.

stituent. In spite of this, the C5–H synthon is also absent. On the other hand, in the structure of 2′-deoxycytidine hemidihydrogenphosphate the four independent dCyd units have proper orientation with respect to the phosphate anions for the formation of the secondary supramolecular synthon (Fig. 6, Table 2), even though the $H_2PO_4^-$ ions themselves are involved in a very peculiar system of hydrogen bonds (tight infinite columns) which seems to dominate the crystal packing [25]. It has to be admitted, however, that the C5–H5...O interactions in these motifs are rather weak, as illustrated by the long H5...O distances and bent C5–H5...O angles listed for units A, B, C in Table 2. Interestingly, 2′-deoxycytidine D, which forms a much shorter and more linear C5–H5...O bond, has the N4–H42...O component compromised (bifurcated and long hydrogen bond, Table 2). The absence of the primary synthons in this structure is the consequence of the $C^+ \cdot C$ base pairing which involves both the N3 and N4–H41 sites of all the dCyd units.

DISCUSSION

The data of Table 1 clearly indicate that in the listed salts of cytidine and 2′-deoxycytidine with composite anions, the ($N3^+$–H, N4–H41)...Y_2X multi-point recognition pattern is a constant and dominating motif of association between the cytosinium cation and the anion. Its robustness and geometrical constancy justify, therefore, its classification as a

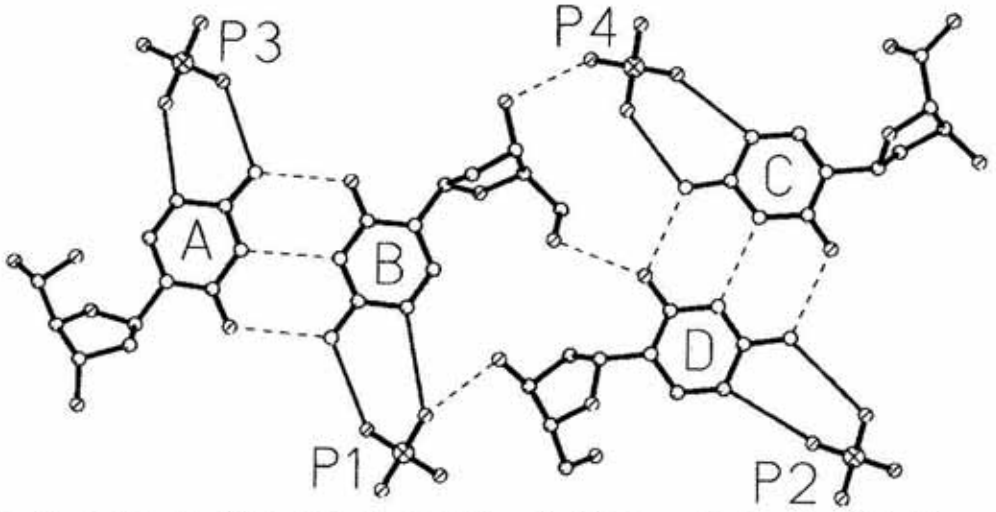

Figure 6. A layer in the crystal packing of 2′-deoxycytidine hemidihydrogenphosphate ($dCyd_2H^+ \cdot H_2PO_4^-$) [25, 26].

The four independent nucleoside units (A,B,C,D) form two protonated $C^+ \cdot C$ base pairs using their N3 and N4–H41 functions and are, therefore, incapable of forming the primary supramolecular synthon with the (two-fold symmetric) dihydrogenphosphate anions (P1, P2, P3, P4). Instead, the full potential of nucleobase...anion interactions is directed towards the C5–H side of cytosine and all four possible secondary synthons are formed (green lines).

supramolecular synthon, as defined by Desiraju [21]. In Table 1 there is only one case when this synthon is formed in a non-classic way, using a water molecule as a relay in the $N3^+$–H...O(anion) interaction ($(dCydH^+)_2 \cdot SO_4^{2-} \cdot H_2O$) and one case when it is absent ($(CydH^+)_2 \cdot SO_4^{2-}$, cation B). In the latter case, there is still a direct interaction between one anion and the two proton-donor sites ($N3^+$–H, N4–H41) of the cytosinium cation but these interactions involve a single acceptor atom of the anion. To confirm the robustness of this primary supramolecular synthon, we have analyzed the data in the Cambridge Structural Database (CSD) [35] searching for all crystal structures (excluding metal complexes) in which an N3-protonated N1-substituted cytosinium cation coexists with a composite anion. There are 33 such cases in the 5.14 release of the CSD, and in 23 of them (70%) the primary supramolecular synthon is observed.

The secondary supramolecular synthon, involving the C5–H side of the cytosine system, is formed less frequently. In the structures listed in Table 1, it is formed in 8 out of 13 possible cases. The main reason for its reduced prevalence in N3-protonated cytidinium and 2′-deoxycytidinium salts seems to be the fact that the much stronger primary synthon formed through the $N3^+$–H protonation site dominates the crystal packing. In cytidinium and 2′-deoxycytidinium salts incapable of forming the primary synthon (chemical modification at the N3 site), the secondary synthon may have increased chances of occurrence, as in the structure of 2′-deoxycytidine hemidihydrogenphosphate where the recognition patterns, even if weak, seem to be present in all four possible situations. There are, however, also cases (3-methylcytidinium nitrate, 3,N^4-ethenocytidinium dihydrogenphosphate) where the secondary synthon is not formed even in the absence of the $N3^+$–H donor. The weak nature of the C5–H5...Y hydrogen bond, which explains the reduced competitiveness of the secondary supramolecular synthon as a determinant of crystal packing, is illustrated in Table 2 where the H5...O distances range from 2.27 to 2.90 Å (1.77 to 2.15 Å in the $N3^+$–H3...O interactions, Table 1). The N4–H42...Y links of the

secondary synthon are not much longer than the N4–H41...Y interactions. However, both hydrogen bonds forming the secondary synthon are noticeably less linear than the $N3^+$–H3...Y/N4–H41...Y interactions (Table 2, 1) illustrating again the relative stability of these two motifs which is directly correlated with their robustness. In a statistical survey of the Cambridge Structural Database similar to that described above, the secondary supramolecular synthon was found to be present in 17 out of 39 possible cases. Interestingly, in this statistical sample, the secondary synthon is always absent when the $N3^+$–H group does not form the primary synthon. When interpreting the reduced frequency of the secondary supramolecular synthon (44%), it has to be noted that among those 22 cases when it is not formed, there are situations, like in cytidine 2′,3′-cyclic-phosphate [36], where the anion is under stereochemical strain that puts strong limitation on its simultaneous participation in two supramolecular synthons. It is then, naturally, the $N3^+$–H type synthon that is favored. The availability of the anion for cytosine...anion recognition may be restricted in a somewhat similar way for intermolecular reasons, as in the structure of 2′-deoxycytidine hemidihydrogenphosphate where the $H_2PO_4^-$ anions form infinite helical columns with a pair of very strong O...H...O hydrogen bonds connecting the consecutive segments [25]. In this particular structure, however, the phosphate...cytosine recognition takes place through the secondary, C5–H, synthon because of the involvement of the N3 side of the nucleobases in $C^+ \cdot C$ pairing (see above). In the other dihydrogenphosphate salt of a nucleoside incapable of forming the primary synthon (3,N^4-ethenocytidinium dihydrogenphosphate [34]), the $H_2PO_4^-$ anions form similar, though less tight, hydrogen-bonded columns, but the C5–H donor is not recognized by the anion. In this context it is interesting to note that when the $N3^+$–H side of the nucleobase becomes available, as in cytidinium and 2′-deoxycytidinium dihydrogenphosphate [27, 28], the strong tendency for self-association of the $H_2PO_4^-$ anion no longer dominates but is superseded by the primary supramolecular synthon. This provides an elegant demonstration of the importance of this multi-point recognition between cytidinium and 2′-deoxycytidinium cations and composite anions.

We wish to thank Professor Maciej Wiewiórowski for his stimulating interest throughout all the phases of this project and Professor Gautam Desiraju and Dr. Ashwini Nangia (University of Hyderabad, India) for helpful discussions.

REFERENCES

1. Lehn, J.-M. (1988) Supramolecular chemistry – scope and perspectives: Molecules, supermolecules, and molecular devices (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **27**, 89–112.
2. *Perspectives in Supramolecular Chemistry: The Crystal as a Supramolecular Entity* [1996], vol. 2 (Desiraju, G.R., ed.) Wiley, Chichester.
3. Dunitz, J.D. (1991) Phase transitions in molecular crystals from a chemical viewpoint. *Pure Appl. Chem.* **63**, 177–185.
4. Schmidt, G.M.J. (1971) Photodimerization in the solid state. *Pure Appl. Chem.* **27**, 647–678.
5. Schmidt, G.M.J. (1976) *Solid State Photochemistry* (Ginsburg, D., ed.) Verlag Chemie, New York.
6. Paul, I.C. & Curtin, D.Y. (1973) Thermally induced organic reactions in the solid state. *Acc. Chem. Res.* **6**, 217–225.
7. Thomas, J.M. (1974) Review lecture: Topography and topology in solid-state chemistry. *Philos. Trans. Roy. Soc. London, Ser. A.* **277**, 251–286.
8. Addadi, L., Ariel, S., Lahav, M., Leiserowitz, L., Popovitz-Biro, R. & Tang, C.P. (1980)

Chemical Physics of Solids and their Surfaces (Roberts, M.W. & Thomas, J.M., eds.) The Royal Society of Chemistry, London.

9. Desiraju, G.R. (1984) Carrying out organic chemistry within crystalline solids. *Endeavour* **8**, 201-206.

10. Taylor, R. & Kennard, O. (1982) Crystallographic evidence for the existence of C-H...O, C-H...N, and C-H...Cl hydrogen bonds. *J. Am. Chem. Soc.* **104**, 5063-5070.

11. Desiraju, G.R. (1991) The C-H...O hydrogen bond in crystals: What is it? *Acc. Chem. Res.* **24**, 290-296.

12. Jeffrey, G.A. & Saenger, W. (1991) *Hydrogen Bonding in Macromolecular Structures.* Springer, Berlin.

13. Steiner, T. & Saenger, W. (1993) Role of C-H...O hydrogen bonds in the coordination of water molecules. Analysis of neutron diffraction data. *J. Am. Chem. Soc.* **115**, 4540-4547.

14. Steiner, T. (1996) C-H...O Hydrogen bonding in crystals. *Cryst. Rev.* **6**, 1-57.

15. Jaskólski, M. (1984) Crystal and molecular structure of 1,N^6-ethenoadenosine hydrochloride. *J. Cryst. Spectr. Res.* **14**, 45-57.

16. Shikata, K., Ueki, T. & Mitsui, T. (1973) The crystal and molecular structure of adenosine hydrochloride. *Acta Cryst.* B**29**, 31-38.

17. Leonard, G.A., McAuley-Hecht, K., Brown, T. & Hunter, W.N. (1995) Do C-H...O hydrogen bonds contribute to the stability of nucleic acid base pairs? *Acta Cryst.* D**51**, 136-139.

18. Evans, T.A. & Seddon, K.R. (1997) Hydrogen bonding in DNA – A return to the *status quo*. *Chem. Commun.* 2023-2024.

19. Moran, S., Ren, R.X.-F. & Kool, E.T. (1997) A thymidine phosphate shape analog lacking Watson-Crick pairing ability is replicated with high sequence selectivity. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10506-10511.

20. Derewenda, Z.S., Lee, L. & Derewenda, U. (1995) The occurrence of C-H...O hydrogen bonds in proteins. *J. Mol. Biol.* **252**, 248-262.

21. Desiraju, G.R. (1995) Supramolecular synthons in crystal engineering – A new organic synthesis. *Angew. Chem. Int. Ed. Engl.* **34**, 2311-2327.

22. Corey, E.J. (1967) General methods for the construction of complex molecules. *Pure Appl. Chem.* **14**, 19-37.

23. Biradha, K., Nangia, A., Desiraju, G.R., Carrell, C.J. & Carrell, H.L. (1997) C-H...O hydrogen bonded multi-point recognition in molecular assemblies of dibenzylidene ketones and 1,3,5-trinitrobenzenes. *J. Mater. Chem.* **7**, 1111-1122.

24. Viswamitra, M.A., Reddy, B.S., Lin, G.H.-Y. & Sundaralingam, M. (1971) Stereochemistry of nucleic acids and their constituents. XVII. Crystal and molecular structure of deoxycytidine 5'-phosphate monohydrate. Possible puckering for furanoside ring in B-deoxyribonucleic acid. *J. Am. Chem. Soc.* **93**, 4565-4573.

25. Jaskólski, M., Gdaniec, M., Gilski, M., Alejska, M. & Bratek-Wiewiórowska, M.D. (1994) Crystal structure of 2'-deoxycytidine hemidihydrogenphosphate reveals $C^+ \cdot C$ base pairs and tight, hydrogen-bonded $(H_2PO_4^-)_\infty$ columns. *J. Biomol. Struct. Dyn.* **11**, 1287-1306.

26. Dauter, Z. & Jaskólski, M. (1996) 2'-Deoxycytidine hemidihydrogenphosphate: Comparison of crystal structure refinements based on diffractometer and image plate data. *Pol. J. Chem.* **70**, 1018-1029.

27. Jaskólski, M. (1989) Structure of cytidinium dihydrogenphosphate. *Acta Cryst.* C**45**, 85-89.

28. Jaskólski, M. (1991) Structure of 2'-deoxycytidinium dihydrogenphosphate. *Acta Cryst.* C**47**, 153-156.

29. Guy, J.J., Nassimbeni, L.R., Sheldrick, G.M. & Taylor, R. (1976) Cytidinium nitrate. *Acta Cryst.* B**32**, 2909-2911.

30. Jaskólski, M. & Wiewiórowski, M. (1994). Isomorphism, solid-state transformations, and disorder in the crystal structures of cytidinium and 2′-deoxycytidinium salts; in: *Correlations, Transformations, and Interactions of Organic Crystal Structures* (Jones, D.W. & Katrusiak, A., eds.) Oxford University Press.

31. Gilski, M. & Jaskólski, M. (1993) Crystal and molecular structure of cytidinium sulfate. *Pol. J. Chem.* **67**, 1841-1850.

32. Wiewiórowski, M., Bratek-Wiewiórowska, M.D., Alejska, M., Perkowska, A., Jaskólski, M. & Rychlewska, U. (1986) The influence of counter anions on the conformations and molecular interactions of protonated nucleosides. *Chemica Scripta* **26**, 229-240.

33. Jaskólski, M. & Wiewiórowski, M. (1987) Structure of 3-methylcytidinium nitrate. *Acta Cryst.* C**43**, 89-92.

34. Gilski, M., Bartoszak, E. & Jaskólski, M. (1991) Structure of 3,N^4-ethenocytidinium dihydrogenphosphate. *Acta Cryst.* C**47**, 1460-1463.

35. Allen, F.H. & Kennard, O. (1993) 3D Search and research using the Cambridge Structural Database. *Chemical Design Automation News* **8**, 31-37.

36. Reddy, B.S. & Saenger, W. (1978) Molecular and crystal structure of the free acid of cytidine 2′,3′-cyclophosphate. *Acta Cryst.* B**34**, 1520-1524.