ORIGINAL PAPER

A comparative study of the hydrogen-bonding patterns and prototropism in solid 2-thiocytosine (potential antileukemic agent) and cytosine, as studied by ^{1}H-^{14}N NQDR and QTAIM/ DFT

Jolanta N. Latosińska · Janez Seliger · Veselko Žagar ·
Dorota V. Burchardt

Abstract A potential antileukemic and anticancer agent, 2-thiocytosine (2-TC), has been studied experimentally in the solid state by ^{1}H-^{14}N NMR-NQR double resonance (NQDR) and theoretically by the quantum theory of atoms in molecules (QTAIM)/density functional theory (DFT). Eighteen resonance frequencies on ^{14}N were detected at 180 K and assigned to particular nitrogen sites ($-NH_2$, $-N=$, and $-NH-$) in 2-thiocytosine. Factors such as the nonequivalence of molecules (connected to the duplication of sites) and possible prototropic tautomerism (capable of modifying the type of site due to proton transfer) were taken into account during frequency assignment. The result of replacing oxygen with sulfur, which leads to changes in the intermolecular interaction pattern and molecular aggregation, is discussed. This study demonstrates the advantages of combining NQDR and DFT to extract detailed information on the H-bonding properties of crystals with complex H-bonding networks. Solid-state properties were found to have a profound impact on the stabilities and reactivities of both compounds.

J. N. Latosińska (✉)
Faculty of Physics, Adam Mickiewicz University,
Umultowska 85,
61–614 Poznań, Poland
e-mail: Jolanta.Latosinska@amu.edu.pl

J. Seliger · V. Žagar
"Jozef Stefan" Institute,
Jamova 39,
1000 Ljubljana, Slovenia

J. Seliger
Faculty of Mathematics and Physics, University of Ljubljana,
Jadranska 19,
1000 Ljubljana, Slovenia

D. V. Burchardt
Department of Paediatric Dentistry,
Karol Marcinkowski University of Medical Sciences,
Bukowska 70,
60–812 Poznań, Poland

Introduction

The compound called 2-thiocytosine (2-TC), an analog of the canonical natural nucleic acid base cytosine where the oxygen has been replaced with sulfur, is present in the tRNA of several organisms, including *Escherichia coli* [1]; however, it has not been detected in the DNA or RNA of natural mammals. The replacement of cytosine by 2-thiocytosine may cause significant changes in DNA structure due to the perturbation of the base-pairing process, or it can produce point mutations. Any alteration of or modification to the base-pairing scheme of DNA due to the existence of a different tautomeric form also may result in the perturbation of the replication process and spontaneous mutations (i.e., reduced stability of DNA). The presence of an abnormal tautomer of a thiosubstituted base can be even more deleterious to the stability of DNA. On the other hand, these unique abilities to modify DNA may be relevant in the design of DNA-binding drugs with high antitumor efficacies. Actually, 2-TC possesses important biological properties; it has shown to have a significant biological effect on the mitosis of human lymphocytes and

is thus a potential antileukemic [2] and anticancer agent [3–9]. The antileukemic activity of fluorinated 2-TC has been proven only recently [10]. Numerous substituted 2-TC derivatives, including 1-(β-D-arabinofuranosyl)-2-thiocytosine and its analogs and complexes with trimethylplatinum have been synthesized, and their enzymatic reactivities and antitumor activities have been studied [11–13]. The cytotoxic activities of some of them were found to be even higher than that of cisplatin, and they were even active against cisplatin-resistant cell lines [13]. It was revealed that the nucleosides of 2-TC exhibit moderate inhibitory activity against Epstein–Barr viruses (EBV) in cell cultures [14, 15], and halogenated nucleosides of 2-TC exhibit considerable inhibitory activity against herpes simplex virus (HSV), and are potent inhibitors of varicella zoster virus (VZV) replication [16]. It was recently found that sulfur metabolic pathways are essential for survival and expression of virulence in many pathogenic bacteria, including *Mycobacterium tuberculosis* [17], so 2-TC (which contains sulfur) may represent a valuable lead for antibacterial and antituberculosis drug development.

Experimental [18–20] and theoretical [21–25] studies have shown that cytosine occurs in six tautomeric forms (the two most stable are the amino-oxo and amino-hydroxy forms), whose relative stabilities depend on the environment. It is well known that in the crystalline state, cytosine adopts its amino-oxo tautomeric form, with four planar molecules arranged in sheets exhibiting a network of H-bonds involving –NH$_2$, –NH, –N=and C=O groups [24, 25]. 2-TC, just like cytosine, exists in six tautomeric forms (Fig. 1) as a result of thiol–thione and amino–imino equilibria. It was recently shown that its environment has an important impact on the tautomeric equilibrium of 2-TC (just as it does for cytosine); this equilibrium is significantly different in the gas phase (the predominant tautomer is the amino-thiol form), in solution (the predominant tautomer is the 1H-thione-amino form), and in nitrogen and argon matrices at low temperatures (exclusively amino-thiol) [26–29].

The preference for different tautomeric forms is a clear indication of the importance of intermolecular interactions, in particular H-bonding, in determining the structure of the condensed phase. The dominant tautomer in the crystalline state of 2-TC is—according to X-ray studies—the amino-thione form [30], but the channel required for proton migration (i.e., hydrogen bonds) is available in the crystalline structure. The arrangement of molecules in the crystals of 2-TC is essentially the same as that found for cytosine, but it is more complicated due to the existence of two nonequivalent molecules in the elementary cell. The presence of the "ordered" and "disordered" H-bonds in crystalline cytosine recently detected using the low-temperature FTIR [31] suggests that the possibility of
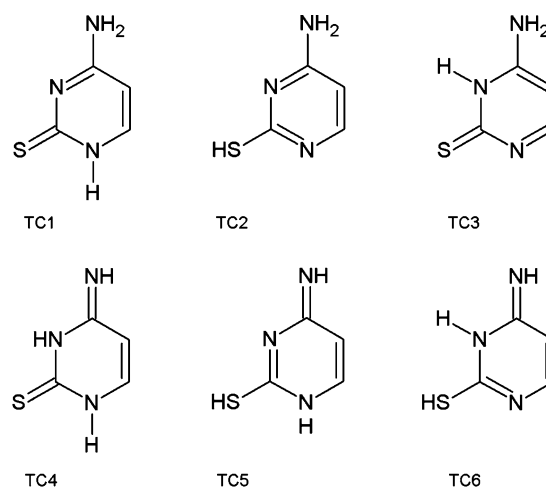
TC1>>TC3>>TC2>TC4>>TC6>TC5

TAUTOMER	ΔE [kJ/mol]	μ [D]
TC1	-	3.09
TC3	62.8	13.05
TC2	92.1	4.56
TC4	119.2	9.40
TC6	148.1	6.74
TC5	155.3	9.08

Fig. 1 Tautomeric structures of 2-thiocytosine

formal proton migrations in cytosine requires further investigation. In our previous papers, we showed that ^{14}N NQDR and QTAIM/DFT combined studies are extremely helpful for interpreting the complicated NQR spectra and for studies of intermolecular interactions in the crystalline state [25, 32–34].

[14]N NDQR is the method of choice because [17]O NQR cannot be applied to 2-TC, although it is a great aid for identifying tautomers, as the [17]O parameters are very sensitive to oxygen hybridization and hydrogen bonds. It is worth noting that the electron density distribution at the –N=, –NH or –NH$_2$ sites is not only more accurately but it is also fully experimentally determined by NQDR. The same cannot be said of any structural method, such as standard X-ray diffraction (where spherically averaged pseudo-atom electronic distributions are used), because of the sensitivity of NQDR to short-range interactions, in contrast to the long-range periodic order seen by the X-ray-based techniques. Although the electron density distribution in the unit cell of a single crystal can be calculated from the structural factors determined in X-ray density quality studies, this electron density is loaded with error arising from experimental errors in structural factors and Fourier truncation errors (only a finite number of reflections can be collected), and phase information is lost during the measurement (the measured intensities are proportional to

the squared structural factors). Thus, density quality studies actually require further modeling of the electron density using independent atoms [35] or a multipole approach [36], and the result is not purely experimental.

Additionally, the study of weak interactions has recently evoked enormous interest, as they commonly occur in important biological systems. The fact that the less abundant tautomer is often the most reactive one (Hammond rule [37]) further reinforces the relevance of the use of combined ^{14}N NDQR and QTAIM/DFT methods to investigate intermolecular interactions in solid-state 2-TC, and to compare these interactions with those occurring in cytosine.

Experimental

A high-purity polycrystalline sample of 2-thiocytosine (97%) was purchased from Sigma–Aldrich and used without any additional purification.

[1]H-^{14}N-NQDR

NQR spectroscopy provides information on the quadrupolar interaction energy of the nuclear charge distribution with the electric field gradient (EFG) evaluated at the nuclear site.

When there is no magnetic field, a ^{14}N nucleus ($I=1$) has three generally nondegenerate nuclear quadrupole energy levels. The resonance (NQR) frequencies, usually termed $\nu_+ > \nu_- \geq \nu_0$, are expressed as [38]:

$$\nu_+\left(^{14}N\right) = \frac{e^2Qq}{4h}(3+\eta)$$
$$\nu_-\left(^{14}N\right) = \frac{e^2Qq}{4h}(3-\eta) \quad (1)$$
$$\nu_0\left(^{14}N\right) = \nu_+\left(^{14}N\right) - \nu_-\left(^{14}N\right) = \frac{e^2Qq}{2h}\eta$$

where the $e^2Qq/h = e^2Qq_{zz}/h$ is the nuclear quadrupole coupling constant, $\eta = \left|\frac{q_{yy}-q_{xx}}{q_{zz}}\right|$ is the asymmetry parameter, $q_{ii} = \frac{\partial^2 V(r)}{\partial x_i^2}$ ($i=x$, y and z) are the principal components of the second-rank symmetric EFG tensor, and $V(r)$ is the external electrostatic potential, which satisfies $|q_{xx}| \leq |q_{yy}| \leq |q_{zz}|$. The quadrupole coupling constant e^2Qq/h and the asymmetry parameter η for a ^{14}N nucleus can be derived from these frequencies using the formulae:

$$e^2Qq/h = \frac{2}{3}\left[\nu_+\left(^{14}N\right) + \nu_-\left(^{14}N\right)\right]$$
$$\eta = 2\frac{\left[\nu_+\left(^{14}N\right) - \nu_-\left(^{14}N\right)\right]}{e^2Qq/h} \quad (2)$$

Different double resonance techniques based on magnetic field cycling were used to detect the ^{14}N NQR frequencies.

The proton spin system was polarized in $B_0=0.75$ T for 30 s. Then the sample was pneumatically transferred within 0.1 s into another magnet, where it was left for 0.3 s. In this other magnet, the magnetic field can be varied continuously between zero and 0.1 T. After staying in this other magnet for <0.1 s, the sample was pneumatically transferred back into the first magnet, and the proton NMR signal was measured immediately after the sample had been stopped in the first magnet.

As the first method we used the [1]H–[14]N cross-relaxation spectroscopy [39, 40]. In this method, the sample is left to relax in a low magnetic field for 0.3 s, and the low magnetic field is changed between magnetic field cycles in steps of approximately 0.5 mT corresponding to the step in the proton Larmor frequency ν_L of 20 kHz. When the proton Larmor frequency ν_L matches the ^{14}N NQR frequency ν_Q, the proton spin-lattice relaxation time shortens, which results in a decrease in the proton NMR signal after the cycle. In some cases, especially at higher proton Larmor frequencies, a step of 40 kHz can be used. On the other hand, at around $\nu_L=\nu_Q$, the step is reduced to 10 kHz to improve the resolution.

In the second step, we used the solid-effect technique [41]. In this method, the low magnetic field is fixed and the sample is irradiated in the low magnetic field for 0.5 s with a strong rf magnetic field at variable frequencies. When the frequency ν of the rf magnetic field is equal to $\nu_Q\pm\nu_L$, simultaneous spin flips take place in both the ^{1}H and ^{14}N spin systems, so the proton magnetization drops to a lower value. This experiment is repeated at a few low magnetic field values in order to clarify the spectrum and get rid of the signal artefacts caused by direct proton absorption of the rf power at multiples of the proton Larmor frequency and the level crossing signals produced by the higher harmonics of the rf magnetic field.

As a final technique combining the three ^{14}N NQR frequencies assigned to a given nitrogen site, we used the two-frequency irradiation technique [42]. Here, the proton Larmor frequency ν_L is set in resonance with the lowest ^{14}N NQR frequency ν_0 and the sample is irradiated with two rf magnetic fields at the frequencies $\nu_1=\nu$ and $\nu_2=\nu + \nu_0$. When $\nu_1 = \nu_-$ and $\nu_2 = \nu_- + \nu_0 = \nu_+$, the proton relaxation rate in the low magnetic field increases, which causes the proton NMR signal at the end of the magnetic field cycle to drop to a low value.

The application of all three double-resonance techniques is necessary to unambiguously determine the triplets (ν_+, ν_- and ν_0) of the ^{14}N NQR frequencies arising from various nonequivalent nitrogen positions in the crystal.

In the NQR spectrum of the compound studied, eighteen resonance lines were detected at 180 K, corresponding to six nonequivalent nitrogen positions in the crystal. The accuracy of NQR frequency determination was 10 kHz.

DFT/QTAIM calculations

Quantum chemical calculations were carried out within the GAUSSIAN03[TM] code [43] run on the CRAY supercomputer at the Poznań Supercomputer and Network Centre (PCSS). All calculations were performed within the density functional theory (DFT) with an exchange-correlation hybrid functional, B3LYP (three-parameter exchange functional of Becke B3 [44] combined with the Lee–Yang–Parr correlation functional LYP [45]), using an extended basis set with polarization and diffuse functions, 6-311++G**, which has been shown to produce highly accurate electron densities and electric field gradient (EFG) tensors. Satisfactory reproduction of the EFG tensors at ^{14}N quadrupolar nuclei by DFT in the solid requires the use of the methodology first applied by us in 1997 [46], which was then improved and widely discussed in our previous papers, including our most recent [25, 32–34], but is only briefly summarized here. The calculations were carried out under the assumption of the presence of monomers and clusters. Formation of molecular clusters was divided into three steps. The first was the formation of six molecular clusters composed of ten molecules which assumed their X-ray atomic positions; the second was a partial optimization achieved using the Berny algorithm, during which only the positions of the hydrogen atoms were allowed to relax while those of all other atoms remained frozen. The third step involved the assumption that the proton positions for the central molecule were the most representative (i.e., they were insignificantly disturbed when PBCs: periodic boundary conditions—were ignored). The clusters constructed in this way—built from equivalent molecules of a certain tautomeric form—were used for further calculations of the NQR parameters and for topological analysis of the electron density within Bader's quantum theory of atoms in molecules (QTAIM). The results of the EFG and QTAIM calculations are only discussed below for the central molecules in these clusters. Such a methodology ensures quite adequate results and is a very good compromise from the point of view of time/disk space requirements and the accuracy with which the electron density distribution can be determined. Moreover, according to the definition of the EFG tensor components, they decrease very quickly with increasing distance; e.g., towards the zz axis as x^2/r^5. In view of the above, it can be assumed that the long-range effect on the reproduction of EFG tensor components is smaller than the effect that follows from taking into account the $\pi \cdots \pi$ stacking interactions (the latter does not exceed 5%). It should be noted that a valid reproduction of the EFG tensor and consequently the abovementioned NQR parameters ensures that a highly reliable electron density distribution is obtained. Thus, the NQR parameters—quadrupole coupling constants, asymmetry parameters,

and frequencies at all nitrogen atoms—were theoretically calculated at the B3LYP/6-311++G** level assuming that different molecular aggregations were formed by the intermolecular interactions depending on the tautomeric forms.

Further theoretical analysis of the intermolecular interactions was performed within QTAIM [47]. In this approach, the electron density $\rho(r)$ of a molecule treated as a scalar field can be examined by analyzing its gradient vector field. The values of $\rho(r)$ at the start and end points of a gradient path that follows the largest increase in $\rho(r)$ take the extreme points (maxima, saddle points, or minima in the electron density) and, depending on the nature of the extremum, they are termed nuclear attractor, bond, ring, and cage critical points, and denoted NACP, BCP, RCP and CCP, respectively. The kind of extremum can easily be determined with the help of a Hessian matrix consisting of nine second–order derivatives of $\rho(r)$. In its diagonalized form, the Hessian provides the three eigenvalues λ_1, λ_2, λ_3 (satisfying $\lambda_1 \leq \lambda_2 \leq \lambda_3$), which characterize the curvature of $\rho(r)$ along the three main curvature axes at the point r. Each critical point in $\rho(r)$ can be classified by a rank equal to the number of nonzero eigenvalues λ_i and a signature calculated as the algebraic sum of the signs of the eigenvalues λ_i of the Hessian. Maxima in the negative Laplacian of the electron density are indicative of local charge concentrations called valence shell charge concentrations (VSCC), and are classified as either bonded (located along bonds) or nonbonded (associated with lone pairs or unshared single electrons) maxima. In addition to the electron density and its Laplacian, the ellipticity of the bond, $\varepsilon = -\lambda_1/\lambda_2| - 1$, the total electron energy density at BCP (H_{BCP}) as well as its components—the local kinetic energy density (G_{BCP}) and the local potential energy density (V_{BCP}), as well as the hydrogen bonding energy according to Espinosa [48] were calculated.

Results and discussion

The [1]H-^{14}N NQDR spectrum of 2-TC as obtained by the solid-effect technique at 180 K is presented in Fig. 2. The experimental NQR parameters are tabulated in Table 1.

The assignment of the frequencies to particular nitrogen sites is not trivial. Admittedly, all resonance signals were detected and assigned for cytosine [25, 49–51], but this task was made easier by having one molecule in the elementary cell. Although the presence of different kinds of ^{14}N sites (−NH$_2$, −N=, and −NH–) in 2-TC facilitates this task, as the NQR parameters (the coupling constant and asymmetry parameter) are significantly different for such sites, the assignment of eighteen NQR frequencies to particular nitrogen sites in the 2-TC molecule is still difficult
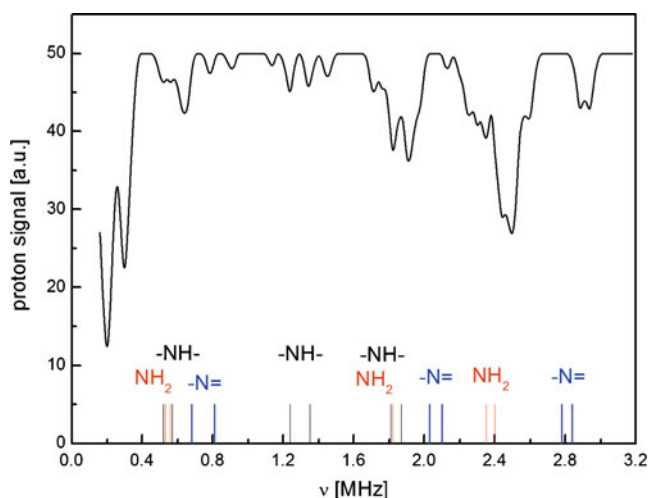
Fig. 2 The experimental ^{1}H-^{14}N NQDR spectrum of 2-thiocytosine obtained at T=180 K by the solid effect technique at the proton Larmor frequency ν_L=100 kHz. Using this technique, three dips are generally observed around the NQR frequency ν_Q at the frequencies ν_Q and $\nu_Q \pm \nu_L$. The dip at the frequency $\nu_Q + \nu_L$ is usually the most pronounced. The ^{14}N NQR frequencies, as confirmed by the two-frequency irradiation techniques and assigned to various nitrogen positions in the molecule on the basis of DFT calculations, are shown on the frequency scale

considering the nonequivalence of molecules (connected with the duplication of sites) and the possibility of prototropic tautomerism (capable of modifying the type of site due to proton transfer). The number of lines that can be grouped into six sets of three resonance lines (Fig. 2) with the help of the third equation of set (1) suggests that there are two tautomeric forms or two crystallographically nonequivalent molecules in the elementary cell. The latter seems to be consistent with the X-ray data [30], according to which 2-TC crystallizes in the monoclinic C2 space group with a=19.5980(17), b=4.0943(3), c=13.5270(7) Å,

and $\beta = 97.54°$. However, the presence of the "ordered" and "disordered" H-bonds in crystalline cytosine, which is structurally related, suggests that the possibility of formal proton migration in 2-TC should be considered.

The stability pattern of 2-TC tautomers in the gas phase (TC2 >> TC1 > TC3 >>TC4 > TC6 >> TC5), obtained at the B3LYP/6-311++G** level of theory, differs in the relative stabilitiy values but is generally in good agreement with the results obtained by Podolyan [9] at the MP2 and MP4 levels. To gain insight into the proton transfer processes in 2-TC in the solid, we modeled large clusters: decamers.

The stability pattern of 2-TC tautomers in the solid phase (TC1>TC3 >> TC2>TC4 >>TC6>TC5) differs from that in the gas phase, but only by the position of the tautomer (TC2) that is most stable in the gas phase. The DFT calculations predict that the TC2, TC4, TC5 and TC6 tautomers of 2-TC are much less stable than TC1 or TC3 in the solid state (see Fig. 1). The dipole moment of 2-thiocytosine (in TC1 form) is higher than that of the cluster (built from TC1 forms) (8.67 and 8.70 D versus 3.09 D).

Moreover, in the TC3 form, which is predicted to be energetically less stable than TC1, the monomer of 2-TC has much a higher dipole moment (10.2 D) and the cluster is also very polar (13.05 D). This suggests that in the solid (just as in polar media), intermolecular interactions stabilize the thione form. The dipole moment in general has a major impact on the packing properties in the solid state. Differences in the polarizations of adjacent molecules are closely connected with the specific arrangement. It can be expected that the polar nature of the molecules acts as a guiding template for a highly ordered solid-state structure. The predicted higher stability of TC1 than TC3 agrees well with the general observation that the requirement for a small number of transferred protons is an obstacle to prototropy

Table 1 The experimental NQR parameters for cytosine and 2-thiocytosine

Site	Cytosine					Reference	2-Thiocytosine					Molecule **
	e^2Qqh^{-1}	η	ν_+	ν_-	ν_0		e^2Qqh^{-1}	η	ν_+	ν_-	ν_0	
–NH$_2$	2.916	0.394	2.49	1.941	0.548	[49] (RT)	2.780	0.381	2.35	1.82	0.53	A
	2.94	0.374	2.480	1.925	0.550	[50] (77 K)	2.813	0.398	2.40	1.82	0.56	B
	2.943	0.387	2.492	1.922	0.570	[51] (77 K)						
–NH–	2.180	0.699	2.016	1.247	0.769	[49] (RT)						
	2.17	0.757	2.035	1.215*	0.820	[50] (77 K)	2.147	0.484	1.87	1.35	0.52	A
	2.165	0.740	2.028	1.219	0.801	[51] (77 K)	2.033	0.561	1.81	1.24	0.57	B
–N=	2.865	0.786	2.712	1.586	1.126	[49] (RT)						
	–***	–	–	–	–	[50] (77 K)	3.247	0.499	2.84	2.03	0.81	A
	2.862	0.806	2.723	1.570	1.153	[51] (77 K)	3.253	0.418	2.78	2.10	0.68	B
O	7.20	0.70	3.549*	2.002*	1.547*	[60]	–	–	–	–	–	–

* Not detected but calculated from the other data (frequencies or $e^2 Qq/h$ and η)

** Assignment given on the basis of ^{14}N NQDR and DFT results

*** Not detected

in infinitely long chains in solids [52]. Thus, the DFT method applied to solids produces reliable conclusions. To gain insight into the influence of proton transfer processes in the solid state on the NQR frequencies, e^2Qq/h and η at ^{14}N atoms in 2-TC, we performed the calculations under the assumption of the presence of six clusters built of ten molecules each at the B3LYP/6-311++G** level; see Table 2 and Fig. 3. A satisfactory reproduction of the e^2Qq/h, η and NQR frequencies at ^{14}N quadrupolar nuclei by DFT in solids requires the clustering methodology. The evident discrepancy between the experimental values and the results of DFT calculations performed using the monomer assumption (see Table 3 and Fig. 4) confirms the validity of our method. The quality of reproduction of the ^{14}N NQR parameters by DFT, assuming one of the most stable tautomers TC1 or TC3, is reasonable at first sight (the correlation coefficients are 0.992 and 0.977, and the standard deviations of the curve fit are 0.102 MHz and 0.339 MHz); see Fig. 4. However, the standard error of the curve fit for TC3 is three times as large as that for TC1. While the experimental NQR frequencies at the same type sites in both nonequivalent molecules differ by at most 0.11 MHz, the NQR frequencies calculated by DFT assuming TC1 differ by 0–0.08 MHz, while those calculated assuming TC3 differ by 0.03–0.55 MHz. The assumption of one of the four remaining tautomers TC2, TC4, TC5

and TC6 leads to unreliable values of e^2Qq/h (overestimated for TC2 and TC4, but underestimated for TC5 and TC6) and η (underestimated for TC2 and TC4 and overestimated for TC5 and TC6). This clearly shows that eighteen resonance lines detected by ^{14}N NQDR are the result of nonequivalence of molecules in the elementary cell, and, similarly to cytosine [25], only one tautomeric form predominates in the NQR spectra of 2-TC. It is worth noting that the stability pattern of 2-TC tautomers differs from that of cytosine [25], but in both only one tautomer predominates in the solid state. Thus, the presence of the "disordered" H-bonds in crystalline cytosine, as described by Rozenberg [31], results from the scattering of the proton position, but not from formal proton migrations. This conclusion is supported by the scatter in the experimental ^{14}N NQR data for cytosine (even at 77 K, which excludes the temperature factor) reported in the literature [49–51].
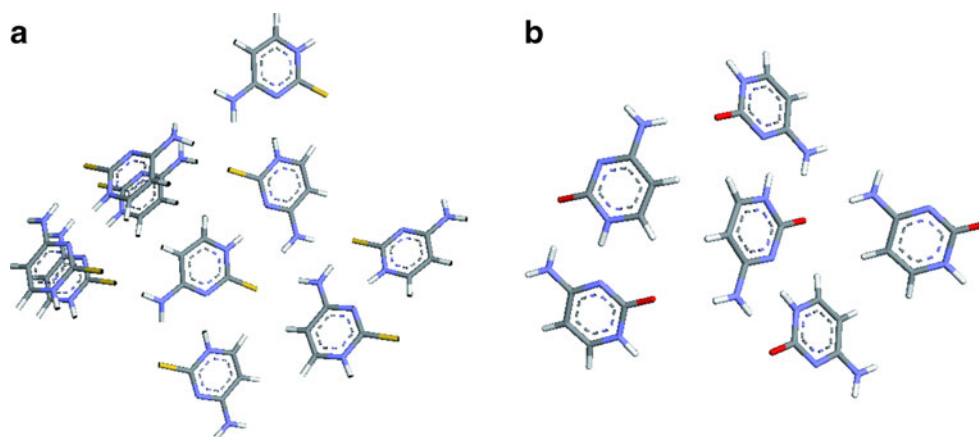
Pattern of intermolecular bonds

In the crystalline structure of 2-TC, the basic unit is a near-planar dimer made of two crystallographically nonequivalent molecules of the amino-thione tautomeric form, connected by two hydrogen bonds N(1′)–H(1′)…N(3) and N(4)–H(4)…S(2′) with lengths of 3.022 and of 3.345 Å, respectively [30]; see Figs. 3a, 5. The structural unit formed

Table 2 NQR parameters calculated at the B3LYP/6-311++G(d,p) level of theory for 2-thiocytosine (clusters of ten molecules)

Tautomer	Formal site	Molecule A					Molecule B				
		e^2Qq/h	η	ν_+	ν_-	ν_0	e^2Qq/h	η	ν_+	ν_-	ν_0
TC1	–NH$_2$	3.215	0.385	2.721	2.102	0.619	3.244	0.382	2.743	2.123	0.620
	–NH–	2.463	0.530	2.174	1.521	0.653	2.471	0.461	2.138	1.568	0.570
	–N=	3.684	0.518	3.240	2.286	0.954	3.649	0.541	3.230	2.243	0.987
	S=	37.301	0.521		20.338		37.386	0.540		20.510	
TC3	–NH$_2$	3.279	0.466	2.841	2.077	0.764	3.535	0.413	3.016	2.286	0.730
	–NH–	2.687	0.315	2.227	1.804	0.423	2.341	0.586	2.099	1.413	0.686
	–N=	3.533	0.589	3.170	2.130	1.040	4.045	0.354	3.392	2.676	0.716
	S=	37.598	0.532		20.573		37.553	0.572		20.824	
TC2	–NH$_2$	4.276	0.139	3.356	3.058	0.297	4.065	0.098	3.148	2.949	0.199
	–NH–	5.342	0.122	4.169	3.844	0.326	3.525	0.168	2.792	2.496	0.296
	–N=	4.836	0.177	3.841	3.413	0.428	3.758	0.225	3.030	2.607	0.423
TC4	–NH$_2$	4.019	0.101	3.116	2.913	0.203	4.407	0.111	3.428	3.183	0.245
	–NH–	3.709	0.242	3.006	2.557	0.449	4.929	0.186	3.926	3.468	0.458
	–N=	3.416	0.175	2.711	2.413	0.299	5.299	0.110	4.120	3.829	0.291
TC6	–NH$_2$	3.129	0.745	2.930	1.764	1.166	2.624	0.719	2.440	1.496	0.943
	–NH–	3.673	0.745	3.439	2.071	1.368	2.695	0.415	2.301	1.742	0.559
	–N=	2.873	0.344	2.402	1.908	0.494	4.073	0.545	3.610	2.500	1.110
TC5	–NH$_2$	2.517	0.974	2.501	1.275	1.226	3.455	0.477	3.003	2.179	0.824
	–NH–	2.775	0.583	2.486	1.677	0.809	3.733	0.164	2.953	2.647	0.306
	–N=	3.644	0.513	3.200	2.266	0.935	2.694	0.418	2.302	1.739	0.563

Fig. 3 Molecular aggregations formed by the intermolecular bonds (1D layer): **a** 2-thiocytosine (12 molecules) and **b** cytosine (six molecules)

by the H-bonds, as shown in Fig. 5, can be termed a "supramolecular synthon" [53]. These supramolecular synthons (dimers) are linked to the neighboring supramolecular synthons by intermolecular interactions of the same pattern; i.e., weaker H-bonds: N(1)–H(1)…N(3′) and N(4′)–H(4′)…S(2) of 3.114 and 3.408Å. Neighboring ribbons are linked together by much longer (and thus weaker) N–H…S bonds with lengths of 3.466 and 3.551Å. It is known that the following geometric criteria often hinder H-bonding: (1) a distance between the proton and the acceptor that is smaller than the sum of the van der Waals radii of these atoms; (2) a donor–proton–acceptor angle that is >90°; (3) elongation of the proton–donor bond. Therefore, the so-called topological criteria formulated by Koch and Popelier [54] on the basis of Bader's QTAIM theory are widely applied: (1) the presence of a BCP for the proton–acceptor contact as confirmation of the existence of a hydrogen-bonding interaction; (2) electron density at the proton–acceptor BCP that is within the range 0.002–0.040 au; (3) a corresponding Laplacian that is within the range 0.024–0.139 au; (4) "mutual penetration" of the hydrogen and acceptor atoms; (5) a loss of charge on the hydrogen atom; (6) energetic destabilization of the hydrogen atom; (7)

decreases in dipolar polarization and the volume of the hydrogen atom. The application of QTAIM not only allows us to verify the existence of these H-bond interactions, identify their nature and characterize their strengths, but it also permits the detection and distinction of many weaker interactions in the crystalline structure of 2-TC, which is not possible in standard X-ray studies [30]. As both N and S show relatively high electronegativities and have free electron pairs, under the right circumstances they can form multicenter bonds where one atom is simultaneously bonded to a few other atoms. Indeed, the crystalline structure of 2-TC contains four varieties of intermolecular interactions, such as the H-bonds N–H…N, N–H…S, C–H…S and C–H…C and three N…N, N…C and N…S interactions involving non-H atoms (see Fig. 5). Their molecular topology was described in terms of BCPs and RCPs, as shown in Table 4. Very small values of $\rho(r)$, small and positive values of the Laplacian, relatively high values of ε, near-zero values of H_{BCP}, and values of $|V_{BCP}|/G_{BCP} \leq 1$ indicate that these interactions are purely closed shell in the dichotomous classification based on the sign of the Laplacian [55, 56]; closed shell—excluding N(4)–H(4)…S(2′), which is of the transit type (it has partial covalent and

Table 3 NQR parameters calculated at the B3LYP/6-311++G(d,p) level of theory for single molecules of 2-thiocytosine (most stable tautomers)

Compound	Tautomer	Formal site	Molecule A					Molecule B				
			e^2Qq/h	η	ν_+	ν_-	ν_0	e^2Qq/h	η	ν_+	ν_-	ν_0
2-Thiocytosine	TC1	–NH$_2$	4.024	0.293	3.31	2.72	0.59	4.060	0.247	3.300	2.790	0.510
		–NH–	2.492	0.804	2.37	1.37	1.00	2.643	0.494	2.310	1.660	0.650
		–N=	4.219	0.343	3.53	2.80	0.72	4.118	0.355	3.450	2.720	0.730
		S=	36.993	0.982	24.442	37.047	0.979	24.441				
	TC3	–NH$_2$	5.008	0.070	3.844	3.668	0.175	5.045	0.065	3.866	3.702	0.164
		–NH–	3.479	0.300	2.870	2.348	0.522	3.491	0.297	2.877	2.359	0.518
		–N=	4.674	0.194	3.732	3.279	0.453	4.713	0.183	3.750	3.319	0.431
		S=	36.729	0.983		24.280		36.890	0.982		24.374	

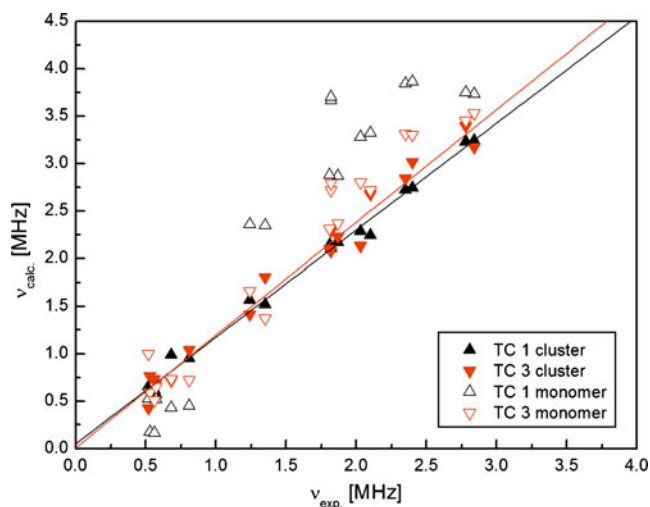The nonequivalent molecules A and B are depicted in Fig. 4

Fig. 4 The correlation between the experimental and calculated NQR frequencies (monomers and clusters of TC1 and TC3; *solid line* linear fit for cluster)

partial ionic character)—in terms of the classification based on the nondimensional $|V_{BCP}|/G_{BCP} < 1$ ratio [55, 57]; or closed shell—excluding N(4)–H(4)…S(2′), which is open shell—in terms of the classification based on the atomic valence shell and both the local (BCP) and integral properties [55, 58]. It is worth noting that for N(4)–H(4) …S(2′) bonds, the H…S distance exceeds 2.2 Å, so their classification as weak H-bonds coincides with the classification proposed by Jeffrey [59], who considered that hydrogen bonds with $d(H…Y) > 2.2$ Å are weak and mainly electrostatic, and hydrogen bonds with $d(H…Y) < 1.5$ Å are strong and covalent.
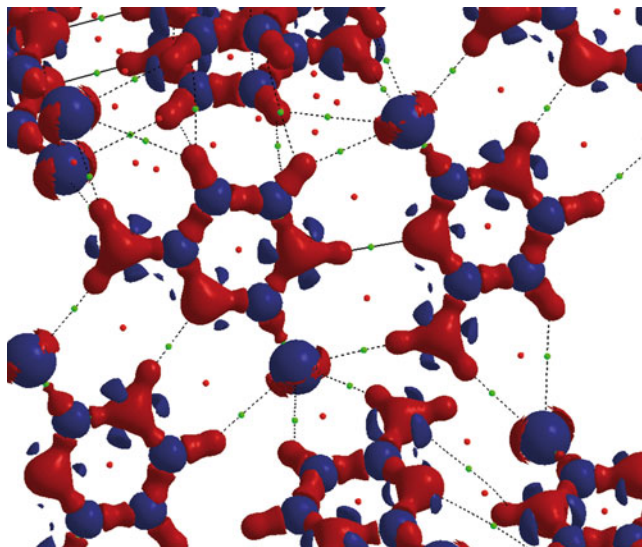

Fig. 5 The 3*d* distribution of the electron density Laplacian calculated by DFT for 2-thiocytosine (isocontour ±0.35 a.u.); the regions of negative Laplacian are shown in *red*, the regions of positive Laplacian in *blue*, *small circles* correspond to critical points (*red* RCP, *green* BCP)

In terms of the total energy density, the N(1′)–H(1′)…N(3) interaction is found to be the strongest (−23.18 kJ mol⁻¹), followed by N(4)–H(4)…S(2′) (−17.13 kJ mol⁻¹); see Table 4. The third H-bond linking adjacent molecules into a dimer, C(6)–H(6)…S(2″)(length: 3.532 Å, strength: −9.43 kJ mol⁻¹), is much weaker and is therefore not revealed by X-ray studies [30]. The hydrogen bonds linking adjacent supramolecular synthons are N(1)–H(1)…N(3′) (3.114 Å, −16.10 kJ mol⁻¹) and N(4′)–H(4′)…S(2) (3.408 Å, −13.95 kJ mol⁻¹). The H-bonds linking neighboring ribbons are two N(4)–H(4)…S (2′) (3.466 Å, −3.17 kJ mol⁻¹ and 3.551 Å, −8.01 kJ mol⁻¹), C(5)–H(5)…S(2′) (4.037 Å, −2.85 kJ mol⁻¹), and two C(5)–H (5)…C(6′) and C(5′)–H(5′)…C(6) (3.688, −3.99 kJ mol⁻¹ and 3.943 Å, −1.91 kJ mol⁻¹), respectively. The reasons for the differences in the strengths of the N–H…S bonds (each linking –NH₂ and S) are their different lengths and the lattice vibrations comprising the motions of the NH₂ groups, which produce averaging of the electron density and electric field gradient, which was not considered.

Additionally, a few π–π interactions (stacking) involving non-H atoms, purely van der Waals in nature, were detected between layered dimers: N(1)…N(3″), N(1)…C(5″) and N (3)…S(2″), with lengths of 4.094, 3.411 and 3.516 Å and approximate estimated energies of −1.14, −4.06 and −4.01 kJ mol⁻¹. It is worth noting that N(4) from the NH₂ group opposite to N(1) and N(3) does not participate in these stacking π–π interactions. The hierarchy of structures in crystals of 2-TC (molecules, supramolecular synthons (dimers), ribbons, stacks) is reflected in progressively weaker bonds.

A more accurate analysis of the Laplacian contour maps and isosurfaces for 2-TC reveals additional details (see Figs. 5 and 6), which are important from the point of view of further comparison with cytosine. It is worth noting that three VSCCs that are oriented towards their bonding partners and consistent with sp^2 hybridization are found at the sulfur atom. Nonetheless, the presence of bond paths that pass through regions of low charge density and BCP clearly indicates the existence of five intermolecular bonds in addition to the C=S bond; see Figs. 5 and 6. It should be noted that three of these intermolecular bonds are in fact very weak, which may explain why they were difficult to discern. One of the VSCCs at the sulfur atom points towards the carbon atom, and one VSCC at the carbon points towards the sulfur atom. This suggests that the C=S bond is not strongly polarized. The other two VSCCs point towards the proton atoms from NH₂ groups bonded to sulfur via the N–H…S hydrogen bonds. In the isosurface representation of the Laplacian (see Fig. 5), the maxima corresponding to these VSCCs are pronounced and quite symmetric, irrespective of whether molecule A or B is analyzed. The search for other minima in the Laplacian around the –NH₂ identifies an additional VSCC, oriented

Table 4 Topological parameters of ρ for 2-thiocytosine and cytosine: the electron density at BCP and RCP $\rho_{BCP}(r)$, its Laplacian $\Delta_{BCP}(\rho)$, the potential electron energy density V_{BCP}, the kinetic electron energy density G_{BCP}, the total electron energy density H_{BCP} and the estimated hydrogen-bonding energy according to Espinosa (E_E) calculated at the B3LYP/6-311++G(d,p) level of theory

Compound	Critical point type	R(X–H...Y) in Å	R(Y...H) in Å	$\rho_{BCP}(r)$ in a.u.	$\Delta_{BCP}(\rho)$ in a.u.	ε	G_{BCP} in a.u.	V_{BCP} in a.u.	H_{BCP} in a.u.	$|V_{BCP}|/G_{BCP}$	$G_{BCP}/\rho(r)$ in a.u.	$H_{BCP}/\rho(r)$ in a.u.	E_E in kJ/mol
2-Thiocytosine	N(1')–H(1')...N(3)	3.022	2.004	0.0272	0.0761	0.0622	0.0183	−0.0177	0.0006	0.97	0.67	0.02	−23.18
	N(4)–H(4)...S(2')	3.345	2.330	0.0235	0.0495	0.0890	0.0127	−0.0130	−0.0003	1.02	0.54	−0.01	−17.13
	C(6)–H(6)...S(2'')	3.532	2.581	0.0148	0.0384	0.0767	0.0084	−0.0072	0.0012	0.86	0.57	0.08	−9.43
	N(1)–H(1)...N(3')	3.114	2.108	0.0213	0.0629	0.0640	0.0140	−0.0123	0.0017	0.88	0.66	0.08	−16.1
	N(4')–H(4')...S(2)	3.408	2.382	0.0209	0.0444	0.0622	0.0109	−0.0106	0.0003	0.97	0.52	0.01	−13.95
	C(5)–H(5)...S(2')	4.037	3.165	0.0050	0.0137	0.3965	0.0028	−0.0022	0.0006	0.79	0.56	0.12	−2.85
	C(5)–H(5)...C(6')	3.688	2.390	0.0063	0.0204	0.8972	0.0040	−0.0030	0.0010	0.75	0.63	0.16	−3.99
	C(5')–H(5')...C(6)	3.943	3.102	0.0065	0.0082	0.8613	0.0016	−0.0012	0.0004	0.75	0.25	0.06	−1.91
	N(4)–H(4)...S(2')	3.466	2.458	0.0166	0.0400	0.0482	0.0090	−0.0081	0.0009	0.90	0.54	0.05	−3.17
	N(4')–H(4')...S(2'')	3.551	2.564	0.0128	0.0340	0.5120	0.0072	−0.0061	0.0011	0.85	0.56	0.09	−8.01
	N...S	3.517	–	0.0063	0.0192	4.5466	0.0039	−0.0031	0.0008	0.79	0.62	0.13	−4.06
	N...N	4.094	–	0.0018	0.0062	0.2816	0.0012	−0.0009	0.0003	0.75	0.67	0.17	−1.14
	N...C	3.411	–	0.0059	0.0161	0.6576	0.0035	−0.0031	0.0004	0.89	0.59	0.07	−4.01
Cytosine	N(1')–H(1')...N(3)	2.84	1.89	0.034	0.102	0.084	0.026	−0.027	−0.0010	1.04	0.76	−0.03	−35.08
	N(4)–H(4)...O(2')	2.98	2.13	0.016	0.061	0.023	0.013	−0.011	0.0020	0.85	0.81	0.13	−14.08
	C(5)–H(5)...O(2')	3.58	2.88	0.005	0.015	0.079	0.003	−0.003	0.0000	1.00	0.60	0.00	−3.30
	N(4)–H(4)...O(2'')	3.03	2.15	0.016	0.067	0.086	0.014	−0.011	0.0030	0.79	0.88	0.19	−14.13
	N...O	3.668	–	0.004	0.014	0.580	0.006	−0.003	0.0030	0.01	1.50	0.75	−3.35
	N...N	3.814	–	0.003	0.008	0.621	0.002	−0.001	0.0010	0.50	0.67	0.33	−1.85
	N...C	3.392	–	0.005	0.015	0.600	0.003	−0.003	0.0000	1.00	0.60	0.00	−3.50

Fig. 6 Laplacian contours in 2-thiocytosine. *Left*: molecule A; *right*: molecule B (negative regions in *red*, positive regions in *blue*)
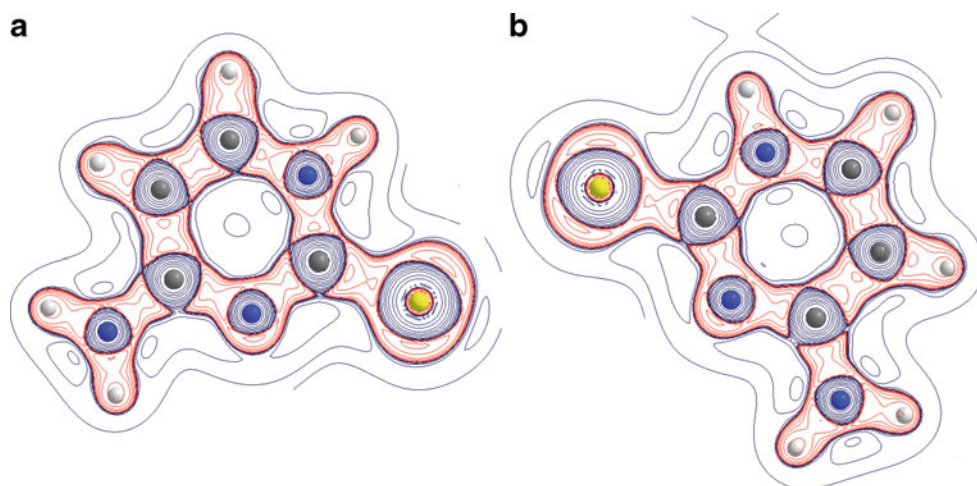
towards the carbon atom. The search for minima in the Laplacian around each of the nitrogen atoms (–N= and –NH–) of the ring gives three VSCCs that are oriented towards their bonding partners and consistent with sp^2 hybridization. The isosurface representation of the Laplacian around –N= reveals the expected VSCC of the H-bond, which is oriented in the plane of the lone pair with a pronounced maximum; by contrast, the donor nitrogen atom –NH– exhibits a more symmetrical distribution of the Laplacian and a VSCC that is directed towards H with an even more pronounced (higher) maximum. The VSCCs of carbon C(2) that are directed towards these nitrogen atoms are shifted to the nitrogen basins in both molecules. Nevertheless, the N–C and N–H bonds show features that are typical of covalent interactions. The differences between the isosurface representations of the Laplacian at the nitrogen sites in both nonequivalent molecules are small, but they suggest that the –NH– site has a more symmetrical electron density distribution in molecule A than in B, while the –N= and –NH$_2$ sites have more symmetrical electron density distributions in molecule B than in A, which is in a good agreement with experimentally obtained values.

2-Thiocytosine versus cytosine

Although the crystal structures of cytosine and 2-TC are generally similar, cytosine crystallizes with only one molecule in the asymmetric unit [24], which gives an extra advantage when studying its intermolecular interactions. A comparison of the geometry of single molecules of 2-TC and cytosine in the crystalline state reveals significant differences in the neighborhood of the C(2) atom, where oxygen is replaced with sulfur. The C(2)–S(2) bond in 2-TC is much longer than the C(2)–O(2) bond in cytosine (1.702 in molecule A and 1.701 in molecule B versus 1.24 Å). The absolute values of the potential electron energy density and the kinetic electron energy density are much smaller in 2-TC than in cytosine (−0.3983 for molecule A

and − 0.4012 for molecule B versus −1.1657 a.u. and 0.1556 for molecule A and 0.1557 for molecule B versus 0.5237 a.u., respectively). The number of electrons shared by adjacent atoms, as measured by the delocalization index, at the BCP in 2-TC is higher than that in cytosine (1.372 for molecule A and 1.360 for molecule B versus 1.132), and the asymmetry ε describing the curvature of $\rho(r)$ and thus indicating multiple bonding or electronic depletion at the BCP is much lower for C=O than for C=S (0.0891 versus 0.1244 for molecule A and 0.1109 for molecule B). Therefore, we can conclude that C=S has more double-bond character than C=O.

However, the NBO analysis clearly shows that the Wiberg and the atom–atom overlap-weighted NAO bond orders for C=O are noticeably higher than for C=S (1.6073 versus 1.4941 for molecule A and 1.4975 for molecule B, and 1.226 versus 1.1420 for molecule A and 1.1451 for molecule B, respectively, for the Wiberg and NAO bond orders).

Detailed analysis of NBO orbitals suggests that C=O has more π and σ character than C=S, as revealed by the lower occupancies of σ and π for C=S than for C=O (σ: 1.9753 for molecule A and 1.9754 for molecule B versus 1.9926; π: 1.9685 for molecule A and 1.9692 for molecule B versus 1.9861). These NBO results are in good agreement with the sum of the Hessian eigenvalues $\lambda_1+\lambda_2$ [−1.9143 for C=O and −0.4467 (for molecule A) and −0.4460 (for molecule B) for C=S] and with λ_3 (1.4404 versus 0.0981 for molecule A and 0.1027 for molecule B).

It is worth noting that $|\lambda_1|/\lambda_3$ is much lower for C=O than C=S (0.6928 versus 2.411 for molecule A and 2.287 for molecule B), which suggests that the C=O bond has more ionic character but the C=S bond is more covalent. The larger integrated net atomic charges in C=O than in C=S (−1.2221 at carbon and 1.7231 at oxygen in cytosine versus −0.1944 and 0.6875 for molecule A and −0.2153 and 0.6864 for molecule B at the carbon and sulfur, respectively, in 2-TC) agree with stronger polarization of

the C=O bond. The dipole moment of the bond, an indicator of polarizability, is much smaller for the C=S bond than the C=O bond (0.118 versus 1.128 D). The position of the BCP, which is located near the middle of the bond for C=O but is shifted to the sulfur atom for C=S, also complies with the higher polarizability of C=O.

Thus, the greater ellipticity of C=S than C=O (confirmed by the delocalization index value), which in fact varies along the bond path, does not seem to be the result of the higher bond order but rather the more symmetric BCP location, in view of the much smaller charge transfer in C=S than in C=O.

The adjacent bonds C(2)–N(3) and N(1)–C(2) in 2-TC are shorter than in cytosine (1.368 and 1.342 Å, versus 1.374 and 1.364 Å). The electron densities at the BCP and RCP in cytosine are lower than in 2-TC (BCP: 0.3297 e for C(2)–N(3) and 0.3095 e for N(1)–C(2) versus 0.3411 e for C(2)–N(3) and 0.3149 e for N(1)–C(2); RCP: 0.02392 versus 0.02413 e). λ_3 for C(2)–N(3) is higher than that for the N(1)–C(2) bond in cytosine (0.3757 versus 0.3677), in contrast to the situation in 2-TC, in which λ_3 is higher for N(1)–C(2) than for the C(2)–N(3) bond (0.3937 versus 0.3619). At the same time, the hybridization of the C(2)–N(3) and N(1)–C(2) bonds in cytosine and 2-TC differs insignificantly (by less than 5%). Moreover, the integrated net atomic charges at N(1), N(3), C(4), N(4) and H(5) increase, while those at C(2), C(5), C(6), H(1), H(4) and H(6) decrease upon substituting the sulfur atom with oxygen. Thus, the change in the curvature along the C(2)–N(3) and N(1)–C(2) bond paths reflects the charge polarization of the C–N bond due to the substitution of sulfur for oxygen, rather than a change in hybridization. The replacement of oxygen with sulfur means that there is no strong inductive effect of oxygen. The resonance donation effect (lone pairs on the oxygen adjacent to the π system), which increases the electron density on the ring and subsequently at the N(1) and N(3) sites *ortho* and decreases this density at C(2) adjacent to oxygen, exceeds the inductive electron-withdrawing effect (through σ bonds) of oxygen. This is because sulfur is less efficient than oxygen at donating its nonbonded (lone) electron pairs in a resonance interaction with the heterocyclic ring. These changes can be interpreted in terms of the flow of bonding electrons to the neighboring C–N bonds (especially the N(3)–C(2) bond) from the external double bond upon substituting the oxygen atom for sulfur. Again, these results are in good agreement with the results of NBO analysis, as the occupancy of N(3)–C(2) exceeds that of C(2)–N(1) in cytosine (1.9951 versus 1.98246), while for 2-TC the situation is reversed, as the occupancy of C(2)–N(3) is lower than that of C(2)–N(1) (1.98149 versus 1.98833).

Topological analysis of the electron density distribution using its Laplacian indicates that three VSCCs are found at the oxygen atom. One of the VSCCs at the oxygen atom points towards the carbon atom, but there is practically no VSCC at the carbon pointing towards the oxygen atom, in contrast to the situation for the C=S in 2-TC, where there is a pronounced VSCC at the carbon pointing towards the sulfur atom. This justifies the description of the C=O bond as very strongly polarized. The presence of bond paths that pierce through the regions of low charge density clearly indicates that there are four intermolecular bonds (two C–H…O and two N–H…O) in addition to the C=O bond.

It should be noted that two of these intermolecular bonds of C–H…O type are in fact very weak, and the bonds of both types occur in close proximity, which explains why it is difficult to find VSCCs, as they are not very pronounced in the isosurface representation of the Laplacian. It is worth noting that the bond paths correspond to preferable interactions. A search for other minima in the Laplacian around each of the nitrogen atoms (–N=, –NH– and –NH$_2$) in cytosine reveals that (just as in 2-TC) they form VSCCs that are oriented towards their bonding partners. The electron densities at almost every BCP (excluding N(3)–C(4) and C(4)–NH$_2$) in cytosine are higher than those in 2-TC.

The charge concentrations around the nitrogen –NH– in cytosine and in 2-TC are much greater in the direction of the hydrogen atom than in the direction of the aromatic ring.

The reverse situation is found for –N= and –NH$_2$; the most distinct concentration is oriented towards the aromatic ring and the less pronounced VSCCs into the hydrogen bond.

This suggests that the π electrons flow from C=O to not only neighboring bonds but also the whole heterocyclic ring. The differences between the isosurface representations of the Laplacian in cytosine and 2-TC are well pronounced at oxygen and sulfur, but the nitrogen –NH– exhibits a more symmetrical distribution in 2-TC than cytosine, as reflected by an increase in the NQR asymmetry parameter. Detailed analysis of the changes in ρ_{BCP} suggests that the most pronounced changes in electron density are those at the N–H bonds of –NH$_2$ and at H-bonds created using the lone pair of –N=. Nonetheless, the NQR parameters at –NH$_2$ differ only slightly, while the change caused by replacing oxygen with sulfur is evident at –N=. Taking into account the orientations of the components of the EFG tensor at each nitrogen site and symmetry considerations, we can conclude that the *zz* axis of the EFG tensor is parallel to the bond path with the BCP at which there is a large change in electron density only at –N=. This explains why the highest decrease in the quadrupole coupling constant upon replacing sulfur with oxygen is observed at the –N= site.

The structural differences caused by exchanging sulfur for oxygen are not limited to single molecules; this substitution also strongly influences the intermolecular

interaction pattern. One of the critical differences between oxygen and sulfur is sulfur's lower electronegativity, which is no less important than its larger van der Waals radius (180 pm instead of 152 pm). Sulfur is able to make less polar and longer bonds with hydrogen than oxygen, so it is expected to participate in weaker intermolecular interactions. In the crystalline structure of cytosine, a molecule in the amino-oxo tautomeric form is linked to adjacent molecules by two strong hydrogen bonds N(1′)–H(1′)...N(3) and two N(4)–H(4)...O(2′) bonds with lengths of 2.84, 2.98, and 3.03 Å, respectively [24]. These hydrogen-bonding interactions in cytosine are shorter and are therefore expected to be stronger than those found earlier in 2-TC, in agreement with the difference in melting points (593 versus 558 K).

The parameters that characterize the weak interactions in the crystalline structure of cytosine (shown in Fig. 7) are given in Table 4. Comparing the Laplacian isosurface and contours for cytosine (see Figs. 7 and 8) to those for 2-TC reveals the differences in VSCC localization. Four H-bonds—two type NH...O and two type CH...O—were found in cytosine, similar to the situation for 2-TC. The N(4)–H(4)...O(2′) bonds in cytosine have strengths of −14.08 and −14.13 kJ mol^{-1}, while those for C(5)-H(5)...O(2′) are −2.94 and −3.30 kJ mol^{-1}.

Although the two strongest hydrogen bonds involving oxygen from cytosine and sulfur from 2-TC are limited to a single octant in 3D space, the bonds with the oxygen from cytosine are weaker and hence the asymmetry parameter of the EFG tensor at ^{17}O is much greater (0.700 [62]) than that predicted by DFT at sulfur.

Although –NH$_2$ groups participate in three interactions with sulfur (oxygen in cytosine)—two hydrogen bonds with strengths of −17.12 and −10.62 kJ mol^{-1} (−13.69 and
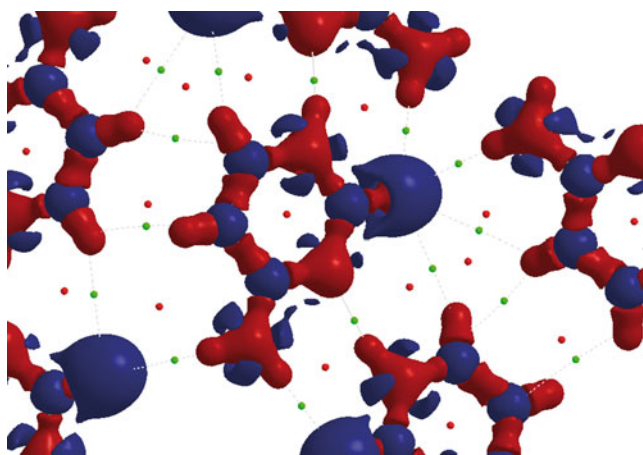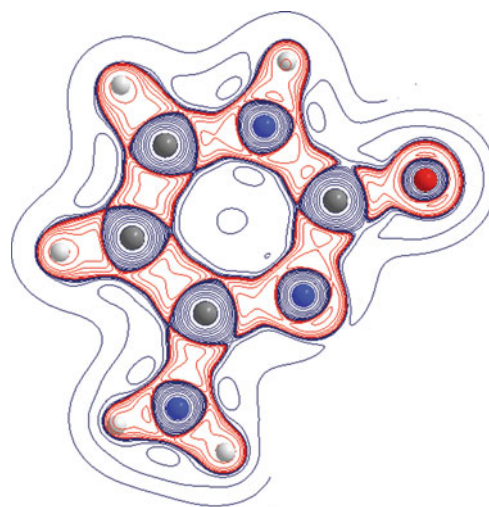

Fig. 8 Laplacian contours in cytosine (negative regions in *red*, positive regions in *blue*)

−13.53 kJ mol^{-1}) and an S...N contact of strength −5.15 kJ mol^{-1} (−0.53 kJ mol^{-1}), the hybridization and directions of the interactions are conserved, so the differences in the NQR parameters are relatively small. The –N=and –NH– sites participate in NH...N bonds of strength −16.10 kJ mol^{-1} (−35.12 kJ mol^{-1} for cytosine), and the differences in the experimental η values for 2-TH and cytosine (see Table 1) lead to the conclusion that the asymmetry parameter η increases when the H-bond is shortened and hence also strengthened, which is in agreement with Seliger's [62] observation. It is worth noting that the quality of reproduction of the ^{14}N NQR parameters by DFT—assuming one of the most stable cytosine tautomers—is weaker than that for 2-TC, which is predominantly due to the poorer geometry but also the use of the older and thus less sophisticated method of collecting NQR data.

Although stacked π–π interactions involving non-H atoms (N...N, N...C and N...O with lengths of 3.814, 3.392 and 3.668 Å and strengths of −1.85, -3.50 and −3.35 kJ mol^{-1})—which are purely van der Waals in nature—were detected in cytosine, they are weak. The stacking found in cytosine and that found in 2-TC involve generally similar overlapping, but the stacking π–π interaction in cytosine is a bit stronger (−12.05 versus −9.7 kJ mol^{-1}), which is reflected mainly in the extra contact in cytosine (N...C) in comparison to 2-TC. Indeed, the improvement in the reproduction of the NQR parameters resulting from the π–π stacking interaction is very important for cytosine (see Fig. 9) but practically negligible for 2-TC. Moreover, the size of the cluster has less of an effect on the quality of reproduction of the ^{17}O-NQR parameters (experimental data from [61]) than on that of the ^{14}N-NQR parameters (see Table 5, Fig. 9). It should be noted that the 2D or 3D cluster constructed in this manner does not

Fig. 7 *3d* distribution of the electron density Laplacian calculated by DFT for cytosine (isocontour ±0.35 a.u.); the regions of negative Laplacian are shown in *red* and the regions of positive Laplacian in *blue*; *small circles* correspond to critical points (*red* RCP, *green* BCP)
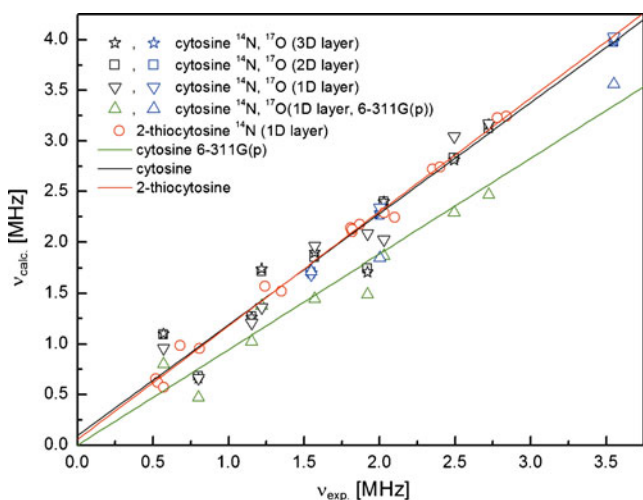
Fig. 9 The correlation between the experimental and calculated NQR frequencies for cytosine and 2-thiocytosine, showing the influence of π–π stacking

take into account periodic long-range effects. As follows from the QTAIM analysis, the differences between the electric charge distributions in the dimer and in the crystal are typically relatively small [60], so when a cluster built from 12 or 18 molecules is considered, the quality of reproduction of the electron density distribution is much improved. Moreover, according to the definition of the EFG tensor components, they decrease very quickly with increasing distance; e.g., towards the zz axis as x^2/r^5. In view of the above, it can be assumed that long-range effects on the reproduction of the EFG tensor components are smaller than those that follow when stacking π–π interactions are accounted for (the latter does not exceed 5%), they do not affect the conclusions of this work.

The above analysis of intermolecular interactions permits one more conclusion to be drawn. As mentioned earlier, cytosine and 2-thiocytosine crystals differ in the number of nonequivalent molecules in an elementary cell (one versus two). The difference between these two structures (a difference that is ultimately due to replacing oxygen with sulfur) is responsible for the effects described above, including an elongation of the C–X (X=O, S) bond and the redistribution of the electron density. These effects cause the two compounds to exhibit slightly different intermolecular bonding patterns, as characterized above. However, it should be emphasized that the hierarchy of structures in crystals of 2-TC and cytosine molecules (supramolecular synthons (dimers), ribbons, stacks), which is reflected in progressively weaker bonds, is the result of the tendency to minimize polarization and enhance structural stability. Indeed, the values of the dipole moments of the monomer of cytosine and two nonequivalent monomers of 2-TC (7.31 versus 8.67 and 8.70 D) are much higher

Table 5 NQR parameters calculated at the B3LYP/6-31 G(d,p) and B3LYP/6-311++G(d,p) levels of theory for cytosine

Site	B3LYP/6-311 G(p) 1D layer (6 molecules)					B3LYP/6-311++G(d,p) 1D layer (6 molecules)					2D layer (12 molecules)					3D layer (18 molecules)				
	e^2Qq/h	η	ν_+	ν_-	ν_0	e^2Qq/h	η	ν_+	ν_-	ν_0	e^2Qq/h	η	ν_+	ν_-	ν_0	e^2Qq/h	η	ν_+	ν_-	ν_0
−NH$_2$	2.520	0.635	2.290	1.490	0.800	3.42	0.561	3.045	2.085	0.959	3.047	0.717	2.831	1.739	1.092	3.014	0.730	2.811	1.710	1.100
−NH−	2.158	0.437	1.864	1.373	0.472	2.258	0.587	2.025	1.362	0.663	2.744	0.498	2.400	1.716	0.683	2.762	0.477	2.401	1.742	0.659
−N=	2.607	0.785	2.467	1.443	1.023	3.421	0.707	3.170	1.961	1.209	3.315	0.766	3.121	1.851	1.270	3.370	0.749	3.159	1.896	1.262
O	6.880	0.912	3.562	1.846	1.716	8.312	0.618	4.026	2.342	1.684	8.097	0.675	3.974	2.264	1.709	8.154	0.657	3.982	2.284	1.699

than those of the dimers (taking into account the most stable forms). However, the molecular synthon (dimer) of 2-thiocytosine consisting of two nonequivalent monomers is less polar than that of cytosine (4.83 versus 5.45D). This smaller polarity of the molecular synthon of 2-TC relative to that of cytosine is a consequence of the greater tendency of sulfur atoms than oxygen atoms to get involved in intermolecular interactions, which implies greater planarity of the molecular synthon of 2-TC than that of cytosine. Moreover, the arrangement of molecular synthons of 2-TC and cytosine in the neighboring ribbons guarantees a considerable—almost twofold—reduction in the dipole moment. A similar tendency to minimize polarization has been observed for the neighboring ribbons in 2-TC and cytosine. The 1D layer built from molecular synthons of 2-TC is slightly more polar than that made from molecular synthons of cytosine. Thus, the π–π stacking interaction between distinct layers in cytosine is slightly stronger than that in 2-TC. This greater steric repulsion for 2-TC leads to increased spacing between layers. It should be mentioned that, despite significant differences between their dipole moments, the differences between the orientations of their dipole moment vectors are small for monomers, dimers and 1D layers in cytosine and 2-TC. Thus, the ordering is generally similar in these solid-state structures, ignoring symmetry, which leads to a kind of duplication. It seems that the polar nature of the molecular synthons acts as a guiding template that produces a highly ordered solid-state structure in both cytosine and 2-thiocytosine.

Reactive sites in 2-thiocytosine and cytosine

In the molecules of 2-TC and cytosine, there are many sites that satisfy the criteria for reactive sites: they have one or more unshared pairs of electrons, polar bonds, and electron-deficient atoms or atoms with expandable octets. The Laplacian of the electron density can be treated as a function that probes these reactive sites. In Figs. 5 and 7, the regions in red in which the Laplacian is negative represent those in which the negative charge is concentrated, whereas the regions in blue are those in which the Laplacian is positive, characterized by a depletion of negative charge. The holes in the reactive surface defined as the isosurface, where the Laplacian changes its sign, clearly indicate that C(4) in 2-TC and C(2) and C(4) in cytosine are potential sites for nucleophilic attack. The electron-depleted regions at carbon atoms C(4)<C(2)<C(6) in 2-TC and C(2)<C(4)<C(6) in cytosine represent potential sites for nucleophilic attack. Additionally, in cytosine, the regions above and below C(2) are more vulnerable to potential nucleophilic attack than they are in 2-TC, where the claws over the nitrogen atom provide additional shielding.

The electron-rich regions at N(3)<N(4)<N(1)<S(2) in 2-TC and N(3)<N(4)<O(2)<N(1) in cytosine indicate potential sites for electrophilic attack.

It is worth noting that the value of the ^{14}N NQR frequency, which describes the degree of redistribution of the electron density from the nonbonding (lone) electron pair to equivalent bonds in which nitrogen participates, is highest for N(3), intermediate for N(2), and lowest for N(1), which clearly indicates the degree of susceptibility to electrophilic attack. Thus, the conclusions drawn from the theoretically calculated Laplacian are in a good agreement with the indications given by the NDQR results.

The amino-oxo tautomer of cytosine has a lower-lying HOMO (highest occupied molecular orbital) and a higher-lying LUMO (lowest unoccupied molecular orbital) than those of TC1 of 2-TC. The narrower HOMO–LUMO gap for the TC1 tautomer (3.56 eV) than for cytosine (5.09 eV) leads to a more stable structure for cytosine and a smaller excitation energy for TC1. In terms of the chemical reactivity descriptors (according to Parr's definition [63]), the chemical potential, absolute hardness and electrophilicity index are equal to −4.10 and −3.52 eV, 2.55 and 1.78 eV, and 3.30 and 3.47 eV for cytosine and TC1, respectively. The maximum electronic charge that cytosine can accept from the environment is equal to 0.81, while it is 0.99 for 2-TC. Consequently, cytosine—which is harder—should be less reactive than 2-TC in unimolecular reactions such as isomerization, dissociation and radical formation. This information could be helpful when attempting to rationally explain the activity of 2-TC in biological and chemical reactions. The enhanced reactivity of 2-TC compared to cytosine suggests the probable occurrence of point mutations in which cytosine is replaced by 2-thiocytosine, leading to reduced stability of DNA.

The HOMO–LUMO gap for the TC3 tautomer (3.69 eV) is only slightly wider than that for TC1 (3.56 eV), meaning that TC3 has a more kinetically stable structure than TC1. The chemical potential, absolute hardness and electrophilicity index are equal to −3.62 eV, 1.84 eV and 3.55 for TC3, while the maximum electronic charge that it can accept from the environment is 0.98. Thus, TC3—which is notably harder than TC1—should be less reactive than TC1 in unimolecular reactions. An analogous analysis performed for a cluster leads to completely different conclusions than those obtained for the isolated molecules. The cluster of TC1 tautomers of 2-TC has a lower-lying HOMO (highest occupied molecular orbital) and a higher-lying LUMO (lowest unoccupied molecular orbital) than those of the TC3 and cytosine clusters. The narrower the HOMO–LUMO gap in the TC3 tautomer (2.45 eV) is compared to that in the TC1 tautomer

(2.69 eV), the more stable the structure of TC1 and the smaller the excitation energy of TC3. In terms of the chemical reactivity descriptors, the chemical potential, absolute hardness and electrophilicity index are equal to −3.55 and −3.66 eV, 1.35 and 1.22 eV, and 4.68 and 5.49 eV for TC1 and TC3, respectively, while the maximum electronic charge that 2-TC can accept from the environment differs by 0.17 (1.32 and 1.49 e for TC1 and TC3, respectively). On the other hand, the slightly narrower HOMO–LUMO gap in cytosine (2.687 eV) compared to 2-TC means that the structure of 2-TC is more kinetically stable and the excitation energy of solid cytosine is smaller. The chemical reactivity descriptors—the chemical potential, absolute hardness and electrophilicity index—for cytosine are equal to −4.17 eV, 1.34 eV, and 6.48 eV, respectively, while the maximum electronic charge that cytosine can accept from the environment is 1.55 e. Intermolecular interactions in the solid cause the HOMO–LUMO gap for cytosine to be narrower than it is for 2-TC, which means that these interactions have a stronger stabilizing effect on the structure of cytosine than on that of 2-TC. This is also reflected in the slightly lower chemical hardness of cytosine (1.34 eV) than that of 2-TC (1.35 eV). These results also imply that in the solid phase, in contrast to the gas phase, sulfur is better than oxygen at stabilizing the structure by electron donation. Thus, solid-state properties have a profound impact on the stabilities and reactivities of both compounds.

Conclusions

1. Eighteen resonance lines detected by ^{14}N NQDR appear as a result of the nonequivalence of molecules in the elementary cell. According to the NQR spectra, in 2-TC one tautomeric form predominates, and the same is true of cytosine.
2. Scattering in the position of the proton—but not formal proton migration—is responsible for the "disordered" H-bonds detected in cytosine by FTIR and the relatively large scattering in the NQR parameters reported in the literature.
3. The combined substituent effect (replacement of oxygen with sulfur), which influences π-electron delocalization within the heterocyclic ring and changes the pattern of intermolecular interactions, is strongly manifested in the NQR spectra as a change in the number of lines in the spectrum (nonequivalent molecules) and changes in the values of NQR parameters of all ^{14}N atoms. These changes are relatively small at the –NH$_2$ site and significant at the –NH– and –N=sites.

4. In cytosine (in contrast to 2-TC), the VSCC at the oxygen (sulfur) atom points towards the carbon atom, but there is practically no VSCC at the carbon that points towards the oxygen atom; the BCP at the C=O bond is shifted towards the carbon basins (in contrast to 2-TC, in which it is located near the sulfur basins); and the electron densities at almost each BCP, excluding N (3)–C(4) and C(4)–NH$_2$ in cytosine are higher than in 2-TC. This suggests that the π-electrons from from C=O to not only neighboring bonds but also the whole heterocyclic ring.
5. The orientation of the zz axis of the EFG tensor parallel to the bond path with the BCP at which the change of electron density is the highest explains the noticeable decrease in quadrupole coupling constant that occurs only at the –N=site upon replacing sulfur with oxygen, while the more symmetrical distribution of the electron density in 2-TC compared to that in cytosine explains the increase in the NQR asymmetry parameter at the –NH site.
6. The stacking π–π interaction in cytosine is stronger, due to the larger spacing between layers in 2-TC compared with cytosine. Thus, the improvement in the reproduction of NQR parameters resulting from the π–π stacking repulsive interaction is substantial for cytosine, but practically negligible for 2-TC.
7. The presence of a sulfur instead of an oxygen substituent at the C(2) position is crucial, because the presence of this sulfur allows 2-TC to form much weaker hydrogen bonds that are reversible, which in turn facilitates biological processes (e.g., enzyme–substrate interactions) and can be used for regulation in metabolic processes. In the solid phase, in contrast to the gas phase, sulfur is better than oxygen at stabilizing the structure by electron donation.

References

1. Carbon J, David H, Studier MH (1968) Science 161:1146–1147
2. Roosaf RA, DeLamater ED (1960) Cancer Res 20:1543–1554
3. Timson J, Price DJ (1972) In: Darlington CD, Lewis KR, Hafner PF (eds) Chromosomes today, vol 3. Wiley, New York, p 118
4. Timson J, Price DJ, Walker JS (1972) Cytobios 5:97–100
5. Lozzio CB, Wigler PW (1971) J Cell Physiol 78:25–31

6. Lozzio CB (1971) Exp Cell Res 69:377–383
7. Szinicz L, Albrecht GJ, Weger N (1981) Arzneimittel Forschg 31:1713–1717
8. Singh K, Groth-Vasselli B, Farnsworth PN, Rai DK (1996) Res Commun Mol Pathol Pharmacol 94:129–140
9. Podolyan Y, Gorb L, Blue A, Leszczynski J (2001) J Mol Struct THEOCHEM 549:101–109
10. Rostkowska H, Nowak MJ, Lapiński L, Bretner M, Kulikowski T, Les A, Adamowicz L (1993) Biochim Biophys Acta 1172:239–246
11. Kawaguchi T, Ichikawa T, Hasegawa T, Saneyoshi M, Yukita A, Asano M, Wakayama T, Kato H, Nagata T (1999) Biol Pharm Bull 22:100–102
12. Kawaguchi T, Ichikawa T, Hasegawa T, Saneyoshi M, Wakayama T, Kato H, Yukita A, Nagata T (2000) Chem Pharm Bull 48:454–457
13. Vetter C, Wagner Ch, Kaluđerović GN, Paschke BR, Steinborn D (2009) Inorg Chim Acta 362:189–195
14. Beetz CP Jr, Ascarelli G (1980) Spectrochim Acta A 36:525–534
15. Yadav RA, Yadav PNS, Yadav JS (1988) Spectrochim Acta A 44:12011–1206
16. Shigeta S, Mori S, Kira T, Takahashi K, Kodama E, Konno K, Nagata T, Kato H, Wakayama T, Koike N, Saneyoshi M (1999) Antivir Chem Chemother 10:195–209
17. Bhave DP, Muse WB, Carroll KS (2007) Infect Disord Drug Targets 7:140–158
18. Nowak MJ, Łapiński L, Fulara J (1989) Spectrochim Acta A 45:229–242
19. Gould IR, Vincent MA, Hiller IH, Lapinski L, Nowak MJ (1992) Spectrochim Acta A 48:811–818
20. Jaworski A, Szczepaniak M, KiBulat K, Person WB (1990) J Mol Struct 223:63–92
21. Kobayashi R (1998) J Phys Chem A 102:10813–10817
22. Les A, Adamowicz L, Bartlett RJ (1989) J Phys Chem 93:4001–4007
23. Podolyan Y, Gorb L, Leszczynski J (2003) Int J Mol Sci 4:410–421
24. McClure RJ, Craven BM (1973) Acta Crystallogr B 29:1234–1238
25. Latosińska JN, Latosińska M, Koput J (2003) J Mol Struct 648:9–18
26. Rostkowska H, Nowak MJ, Lapiński L, Brenner M, Kulikowski T, Les A, Adamowicz L (1993) Spectrochim Acta A 49:551–565
27. Rostkowska H, Nowak MJ, Lapiński L, Brenner M, Kulikowski T, Les A, Adamowicz L (1993) Biochim Biophys Acta 1172:239
28. Kwiatkowski JS, Leszczyński J (1996) J Phys Chem 100:941–953
29. Kwiatkowski JS, Pullman B (1975) Adv Heterocycl Chem 18:199–335
30. Furberg S, Jensen LH (1970) Acta Crystallogr B 26:1260–1268
31. Rozenberg M, Shoham G, Reva I, Fausto R (2004) Spectrochim Acta A 60:463–470
32. Latosińska JN, Seliger J, Žagar V, Burchardt DV (2009) J Phys Chem A 113:8781–8790
33. Latosińska JN, Latosińska M, Seliger J, Žagar V, Kazimierczuk Z (2009) Chem Phys Lett 476:293–302
34. Latosińska JN, Latosińska M, Seliger J, Žagar V, Maurin J, Orzeszko A, Kazimierczuk Z (2010) J Phys Chem A 114:563–575
35. Hansen NK, Coppens P (1978) Acta Crystallogr A 34:909–921
36. Allen FH (1986) Acta Crystallogr B 42:515–522
37. Hammond GS (1955) J Am Chem Soc 77:334–338
38. Seliger J (2000) In: Lindon JC, Tranter GE, Holmes JL (eds) NQR theory in encyclopedia of spectroscopy and spectrometry. Academic, San Diego, pp 1672–1680
39. Seliger J, Osredkar R, Mali M, Blinc R (1976) J Chem Phys 65:2887–2893
40. Seliger J, Blinc R, Arend H, Kind R (1976) Z Phys B 25:185–189
41. Stephenson D, Smith JAS (1988) Proc R Soc Lond A 416:149–178
42. Seliger J, Žagar V (2008) J Magn Reson 193:54–62
43. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Vreven T Jr, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, M. A. Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, revision D.01. Gaussian Inc., Wallingford
44. Becke AD (1993) J Chem Phys 98:1372–1377
45. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789
46. Latosińska JN (2000) Magn Reson Chem 38:192–196
47. Bader RFW (1994) Atoms in molecules: a quantum theory. Oxford University Press, Oxford
48. Espinosa E, Molins E, Lecomte C (1998) Chem Phys Lett 285:170–173
49. Garcia MLS, Smith JAS (1983) J Chem Soc Perkin Trans II 1401–1408
50. Blinc R, Mali M, Osredkar R, Prelesnik A, Seliger J, Zupancic I, Ehrenberg L (1972) J Chem Phys 57:5087–5093
51. Safin IA, Osokin DYa (1977) Yadernyj kvadrupolnyj rezonans v soedinieniach azota (Nuclear quadrupole resonance in nitrogen compounds). Nauka, Moskow
52. Klein O, Aguilar-Parrilla F, del Lopez Amo JM, Jagerovic N, Elguero J, Limbach HH (2004) J Am Chem Soc 126:11718–11732
53. Desiraju GR (1995) Angew Chem Int 34:2311–2327
54. Koch PLA, Popelier A (1995) J Phys Chem 99:9747–9754
55. Gatti C (2005) Z Kristallogr 220:399–457
56. Bader RFW, Esse H (1984) J Chem Phys 80:1943–1960
57. Espinosa E, Alkorta I, Elguero J, Molins E (2002) J Chem Phys 117:5529–5542
58. Macchi P, Proserpio DM, Sironi A (1998) J Am Chem Soc 120:13429–13435
59. Jeffrey GA (1997) An introduction to hydrogen bonding. In: Truhlar DG (ed) Topics in physical chemistry. Oxford University Press, New York, p 303
60. Wu G, Dong S, Ida R, Reen N (2002) J Am Chem Soc 124:1768–1777
61. Munshi P, Cameron E, Cameron TS, Row TNG, Ferrara JD (2007) J Phys Chem A 111:7888–7897
62. Seliger J (1998) Chem Phys 231:81–86
63. Parr RG, Pearson RG (1983) J Am Chem Soc 105:7512–7516