Determination of the Tautomerism of 5,5-Disubstituted Analogues of 6-Amino-2-thiouracil by ^1H and ^{13}C Nuclear Magnetic Resonance Spectroscopy

Bożenna Golankiewicz,* Joanna Zeidler, and Mariusz Popenda
Institute of Bio-organic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland

5,5-Disubstituted 6-amino-2-thiouracil analogues (**1**)–(**6**), prepared by condensation of 2-acylamino-2-cyano aliphatic acids ethyl esters with thiourea, have been studied by ^1H and ^{13}C NMR. It has been found that in dimethyl sulphoxide solution the compounds exist as 5-acylamino-5-alkyl-6-imino-2-thioxo-1,2,5,6-tetrahydropyrimidin-4(3*H*)ones (**1B**)–(**6B**). The assignment of this tautomeric form has been based mainly on ^{13}C NMR spectra selectively ^1H decoupled with low-power ^1H irradiation.

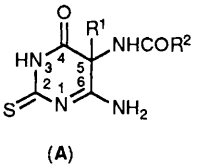
There has been continued interest in the tautomeric structures of oxo-, thioxo-, and amino-substituted pyrimidines, stimulated by their biological importance.[1] Several papers have appeared recently concerning the tautomerism of 6-amino-2-thiouracil.[2–4]

We have been interested in structurally related 5,5-disubstituted compounds (**1**)–(**6**) as the substrates undergoing cyclization–rearrangement reaction leading to derivatives of imidazo[1,5-*a*]- 1,3,5-triazine, a novel ring system isomeric with purine.[5]

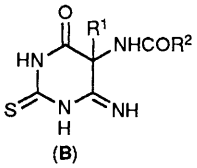
In our initial reports,[5] tautomeric form (**A**), 6-amino-4-oxo-2-thioxo-, has been tentatively assumed for compounds (**1**) and (**2**). Further application of the cyclization–rearrangement reaction, aimed towards biologically active compounds,[6] requires synthesis of other 5,5-disubstituted derivatives (**3**)–(**6**). Upon routine ^1H NMR spectral characterization of the latter, some conspicuous features were noticed, which encouraged us to undertake a detailed study of the ^1H and ^{13}C NMR spectra. Thus, we gained information on the preferred structural form of the compounds (**1**)–(**6**) in solution. The information is presented here, together with the data on the synthesis and the properties of the new compounds.

Results and Discussion

The novel 5,5-disubstituted 6-amino-2-thiouracil analogues (**3**)–(**6**) are prepared by a route involving condensation of 2-acylamino-2-cyanoaliphatic acids, ethyl esters $EtCO_2C(R^1)(CN)NHCOR^2$ with thiourea. Modifications of the procedure which we used previously[5] to obtain (**1**) and (**2**) are necessary to achieve satisfactory yields. Compounds (**3**)–(**6**) are accompanied, as already noted during the synthesis of (**1**) and (**2**),[5] by small amounts of corresponding 6-R^2-8-R^1-2-thioxoimidazo[1,5-*a*]-1,3,5-triazin-4(3*H*)-ones. This observation is of importance for the study of the mechanism of the cyclization–rearrangement of compounds (**1**)–(**6**) under silylation conditions leading to high yields of corresponding 6-R^2-8-R^1-2-thioxoimidazo[1,5-*a*]-1,3,5-triazin-4(3*H*)-ones,[5,6] which will be reported separately.



(A)



(B)

Compound	R^1	R^2
(**1**)	Me	H
(**2**)	Me	Me
(**3**)	CH_2–CH=CH_2	Me
(**4**)	Pr^i	Me
(**5**)	$PhCH_2$	Me
(**6**)	Me	Ph

Physical characteristics of the compounds (**3**)–(**6**) are given in Table 1.

The spectral data are summarized in Tables 2 and 3. In the ^1H NMR spectra (Table 2), signals deriving from the protons of the substituents at the quaternary C–5 carbon were straightforwardly assigned on the basis of chemical shifts and multiplicities. Two methyl groups of isopropyl substituent in the spectrum of (**4**) appeared as two partly overlapping doublets (*J* 6.7 Hz) centred at δ_H 0.96 and 0.92, possibly as a result of restricted rotation of this bulky group in the presence of acylamino group at the same ring carbon. No indication of the existence of two conformers due to hindered rotation around the amide C–N bond of the acylamino part was found. Interesting features were revealed in the high frequency region of the spectra. Four protons exchangeable with D_2O appeared as four separate one-proton signals. The assignment of these signals was gained from the ^{13}C spectra (Table 3), using the long-range successive selective ^1H decoupling (LSPD) technique.[7,8] Selective irradiation of the ^{13}C proton coupled spectra with particular proton frequencies at low-power decoupling field (100 Hz > γB_2 > 40 Hz) was followed by the analysis of the resulting changes in terms of amplitudes and coupling patterns. It turned out that all four exchangeable protons were located at nitrogen atoms. The broad peaks appearing at the highest frequencies [δ_H 11.28–11.64, compounds (**3**)–(**6**)], immediately removable with D_2O, were attributed to N–3 protons. Irradiation with exact frequencies of these protons resulted in ^{13}C NMR spectra in the narrowing of the broad lines of C–2 and C–4 signals indicating these carbons to be the nearest neighbours. Similarly, the unequivocal assignments of the signals deriving from the protons linked to nitrogen of acylamino groups (AANH) were possible. When compounds (**3**–**6**) were irradiated with proton frequencies corresponding to δ_H 8.86, 8.58, 9.00, and 9.40, respectively, the narrowing and multiplicity reduction of the signals deriving from C–4, C–5, C–6, and C=O of acylamino group were observed. The signals of AANH were sharp and were removed with D_2O at the slowest rate. The remaining two exchangeable protons showed up at δ_H 8.3–9.0 region as separated broad signals which partly

Table 1. Analytical data and properties of 5-acylamino-5-alkyl-6-imino-2-thioxo-1,2,5,6-tetrahydropyrimidin-4(3*H*)-ones (**3**)–(**6**).

Compound	Yield (%)	Solvent	M.p./°C (decomp.)	Found (Required) C	H	N	UV(EtOH) λ_{max}/nm ($\varepsilon 10^{-3}$dm^{3}mol^{-1}cm^{-1})	Mass spectrum *m/z* (% rel. int.)	pK_a/B$^+$H/ 80% MCS[a]
(**3**)·0.75H_2O	50	H_2O	244	42.6 (42.6	5.4 5.4	22.2 22.1)	266(8.7), 320(23.9)	240(M^+, 100), 198(90), 181(9), 151(78)	2.6
(**4**)·0.25H_2O	47	H_2O	204–205	43.8 (43.8	5.8 5.9	22.9 22.7)	268(11.4), 322(22.0)	242(M^+, 82), 183(5), 158(100), 157(25)	2.5
(**5**)	59	H_2O–EtOH 1:1	243–244	53.6 (53.7	4.7 4.9	19.5 19.3)	268(7.8), 322(22.0)	290(M^+, 39), 231(8), 157(20), 91(100)	2.6
(**6**)	64	H_2O–EtOH 1:1	250–251	52.3 (52.1	4.2 4.4	20.2 20.3)	230(13.7), 262(10.8) 317(26.8)	276(M^+, 45), 155(16), 143(13), 105(100)	4.4

[a] 2-Methoxyethanol.

Table 2. ^{1}H NMR spectra of 5-acylamino-5-alkyl-6-imino-2-thioxo-1,2,5,6-tetrahydropyrimidin-4(3*H*)-ones (**1**)–(**6**) at 90 MHz. Chemical shifts δ_H(ppm): solvent $(CD_3)_2SO$–TMS. [a] Spectral width 1.3 kHz.

Compound	N–3 H[c]	AA[b] NH[c]	N–1 H[c,d]	=NH[c,d]	5–R^1			R^2		
(**1**)[e]	11.60	(8.60)	9.20	(8.60)	1.52	s, 3H	CH_3	7.93	s, 1H	CHO
(**2**)	11.52	(8.90)	(8.90)	8.45	1.46	s, 3H	CH_3	1.87	s, 3H	CH_3
(**3**)	11.54	(8.86)	(8.96)	(8.42)	5.76–5.39	m, 1H	C*H*=CH_2	1.88	s, 3H	CH_3
					5.21–5.01	m, 2H	CH=C*H*$_2$			
					2.60[f]		CH_2			
(**4**)	11.49	8.58sh	9.01	8.29	2.03	m, 1H	$CH(CH_3)_2$	1.89	s, 3H	CH_3
					1.00	q, 6H	$CH(CH_3)_2$			
(**5**)	11.28	(9.00)sh	(8.88)sh	8.44	7.31–7.16	m,3H	C_6H_5	1.90	s, 3H	CH_3
					7.07–6.96	m,2H				
					3.12	m,2H	CH_2			
(**6**)	11.64	9.40sh	(8.97)	(8.64)	1.62	s,3H	CH_3	8.02–7.91	m, 2H	C_6H_5
								7.63–7.56	m, 3H	

[a] s, singlet; m, multiplet; q, quintet. [b] AA, acylamino. [c] Exchangeable with D_2O; mostly broad, one-proton singlets; sharp exchangeable singlets are marked sh; in parentheses overlapping or very closely located signals for which joined integration was two or three protons. [d] N–1 and =NH protons are not distinguishable by ^{13}C LSPD technique, the assignments can be interchanged. [e] Data from ref. 5 reassigned by analogy to compound (**2**). [f] Overlapping with $(CD_3)_2SO$ signal.

overlapped only in the spectrum of C–5 benzoylamino derivative (**6**). The evidence was thus provided that in 5,5-disubstituted compounds (**3**)–(**6**) not 6-amino but 6-imino N–1 H form is predominant. The chemical shifts of the latter two protons could not be unequivocally ascribed by ^{13}C LSPD technique, even when very low strength of decoupling field ($\gamma B_2$40 Hz) was applied. All three broad signals of carbon atoms C–5, C–6, and C–2 were narrowed when irradiated with the frequency of either of the two broad bands in the δ_H 8.3–9.0 region. This result was possibly due to the dynamic exchange of the discussed protons. The observed narrowing of C–2 signal excluded another possibility—the inequivalence of 6–NH_2 protons resulting from the restricted rotation around the C(6)–N(6) bond, which might be due to the proximity of bulkily disubstituted C–5. Such narrowing could not be expected to derive from the low-power irradiation with the frequency of the proton four bonds removed.

Fully unequivocal assignments of ^{13}C signals were gained from NBD, proton coupled (AP NMR), LSPD, and in some cases from gated decoupled–without NOE spectra (Table 3). Identification of the signals of carbon atoms directly bonded with protons was readily accomplished from their characteristic chemical shifts together with multiplicities and measured coupling constant values ($^1J_{CH}$ 128–130 Hz). Signals of quaternary carbon atoms, with the exception of C–5 showing up in the δ_C 55–62 range, appeared in the high frequency region. The ones having the highest chemical shift values, δ_C 186.4–187.2, were attributed to C–2 carbon atoms in the form of thioxo groups by comparison with corresponding values reported for thioxo groups in 2-thiouracil [2-thioxo-2,3-dihydropyrimidin-4(1*H*)-one, δ_C 175.9] and in 4-thiouracil [4-thioxo-3,4-dihydropyrimidin-2(1*H*)-one, δ_C 191.3].[9] The assignments for the remaining three high-frequency signals (C–4, C–6, and C=O acyl), chemical shifts of which were within a very narrow region, δ_C 166.4–171.7, were based on the analysis of long range coupling patterns. The long-range coupling constants could not be measured due to the coupling of these carbons with NH protons resulting in broadening of the multiplets. The broadening effect of NH was demonstrated by use of the proton-coupled spectrum of benzoylamino derivative (**6**) taken on a sample pre-treated with D_2O. The multiplets deriving from C–4, C–6, C=O acyl, and C–5 became simplified and 'sharpened'. Corroboration with data from proton-coupled spectra was provided by NBD spectra comprising only the regions of C–4, C–6, and C=O acyl signals measured at a spectral width of 200 Hz. Under such conditions (resolution improved to 0.1 Hz), C–4 and C=O acyl appeared as sharp singlets, C–6 was broad due to the proximity of two nitrogen atoms. The assignments were confirmed by ^{13}C LSPD NMR spectra. The

Table 3. ^{13}C NMR spectra of 5-acylamino-5-alkyl-6-imino-2-thioxo-1,2,5,6-tetrahydropyrimidin-4(3*H*)-ones (**2**)–(**6**) at 22.6 MHz. Chemical shifts δ_C(ppm): solvent $(CD_3)_2SO$–TMS.[a] Spectral width 5.0 kHz.[b]

	C-2	C-4	C-5	C-6	C=O(acyl)	R(acyl)			C-5 R		
(2)	186.8	168.2	54.4	171.1	169.7	21.6	Q	CH_3	24.2	Q	CH_3
	Ss	Sq	Sbrs	Sbrq	Sm						
(3)	186.8	167.2	57.8	169.6	169.6	21.6	Q	CH_3	128.5	Dm	$CH{=}CH_2$
	Ss	St	St	S[c]	S[c]				120.9	Tm	$CH{=}CH_2$
									46.7	Tm	CH_2
(4)	187.2	167.0	62.0	169.4	169.9	21.8	Q	CH_3	35.7	Dm	$CH(CH_3)_2$
	Ss	Sd	Sm	Sbrd	Sm				17.2	Qm	$CH(CH_3)_2$
									16.5	Qm	
(5)	186.4	167.3	58.9	168.9	169.5	21.6	Q	CH_3	131.7 Sm	129.6 Dm	Ph
	Ss	St	Sbrs	Sbrm	Sdq				128.0 Dm	127.6 Dm	
									43.4	T	CH_2
(6)	187.0	168.3	55.0	171.7	166.4	132.3	St	Ph	24.2	Q	CH_3
	Ss	Sbrq	Sm	Sbrm	Sbrm	131.9	Dt				
		q[d]	q	q	t	128.0[e]	Dd				

[a] Capital letters refer to the pattern resulting from directly bonded ^{13}C–^1H couplings, lower-case letters to those from ^{13}C–^1H couplings over more than one bond. S or s, singlet; D or d, doublet; T or t, triplet; Q or q, quartet; m, multiplet; dq, doublet of quartets; br, broad due to unresolved couplings. [b] Signals of C-2, C-4, and C=O (acyl) were additionally examined at SW 0.2Hz. [c] From integration 2C. [d] Multiplicities after exchange of NH protons with D_2O. [e] From integration 4C.

chemical shift of C-4 ranged in compounds (**3**)–(**6**) from δ_C 167.0–168.3 and was close to that reported for C-4 of 2-thiouracil (δ_C 160.9).[9]

In the ^1H NMR spectra of compounds (**1**) and (**2**) four exchangeable with D_2O protons appeared as only three separate signals. One of them, at δ_H 8.60 and 8.90 respectively, as a two proton signal, had been previously tentatively interpreted as deriving from 6-amino group.[5] During the present detailed examination of the ^1H and ^{13}C spectra of compound (**2**), it turned out that the irradiation of the ^{13}C spectrum with the frequency of protons resonating at δ_H 8.90 sharpened the intensity of C-2, C-6, C-5, and 5-acetylamino carbonyl signals. This result indicated that the signals of two separately located protons—one at AANH and the other either at C-6 =NH or N-1 H overlapped at δ 8.90.

The above data indicate that for compounds (**1**)–(**6**) tautomeric structure (**B**), 6-imino-4-oxo-2-thioxo- is preferred.* Such tautomeric preference is different from that for 6-amino-2-thiouracil, which is known to exist in the 6-amino form.[2] The data which already exist on tautomerism of uracil and related pyrimidines in solution are based on the comparative study of the properties of the molecules and their fixed tautomer forms.[4] The instability of the 5,5-disubstituted analogues of 6-amino-2-thiouracil (**1**)–(**6**) makes the access to fixed derivatives of their tautomeric forms (**A**) and (**B**) very difficult. The present work is a case in which, without recourse to model compounds, the information about the preferred tautomeric form could have been gained directly from consideration of NMR data. The long-range successive selective decoupling method used by us does not, however, provide quantitative information. Nevertheless, equal, one-proton intensities of the signals ascribed to N-1 H and C-6 =NH indicate a very high preponderance of the tautomeric form (**B**). The rare tautomers of the naturally occurring nucleobases are of great importance because of their possible biological role in mutagenesis.[10] 5,5-Disubstituted analogues, characterized in this work as having the 6-imino structural feature, may find application as model compounds for estimation of some properties of rare nucleobase tautomers.

Experimental

Synthesis.—Compound (**6**) was prepared according to the procedure described for (**2**).[5] Modifications of this procedure were necessary to obtain satisfactory yields of (**3**)–(**5**).

General procedure for the synthesis of compounds (**3**)–(**5**). To a solution of sodium ethoxide in ethanol [Na(0.69 g) in absolute ethanol, (30 cm^3, 30 mmol)] was added an appropriate ester $EtCO_2(R^1)(CN)NHCOR^2$ (10 mmol), prepared in accordance with the literature,[11] followed by thiourea (30 mmol). The resulting solution was heated at reflux for 1 h, cooled to 0 °C and adjusted to pH 5 with glacial acetic acid. Compounds (**4**) and (**5**) crystallised from that solution. Compound (**3**) only after the concentration of the solution was reduced to one third of its original volume and addition of water (10 cm^3) added. The crystals were collected by filtration, washed with cold absolute ethanol and dried *in vacuo*. Analytical samples were prepared by quick recrystallization from hot water or aqueous ethanol. Prolonged heating in these solvents resulted in decomposition. The filtrate contained an additional portion of (**3**)–(**5**) as well as the corresponding 6-R^2-8-R^1-2-thioxoimidazo[1,5-*a*]-1,3,5-triazin-4(3*H*)-ones (1–8%),[6] which could have been separated by crystallization.

Spectra.—All NMR spectra were obtained in 5 mm spinning tubes on a JEOL FX 90Q Fourier transform spectrometer operating at a frequency of 90 MHz for ^1H and 22.6 MHz for ^{13}C. The probe temperature was 22 °C. Data were accumulated with a Texas Instruments JEC-980B computer system using 8K data points. ^{13}C proton coupled spectra were obtained in the proton gated mode (alternately pulsed—AP NMR spectra). Integrations, if necessary, of some ^{13}C signals were obtained from gated decoupled-without NOE spectra. Long-range successive selective ^1H decoupling—LSPD spectra were performed for ten different decoupler frequencies at constant power level $100 > \gamma B_2 > 40$ Hz. General parameters employed were as follows (specified in the order: ^1H, ^{13}C NBD, AP NMR, LSPD, and gated decoupled-without NOE spectra): spectral width, 1.3, 5.0 or 0.2, 5.0, 5.0, and 5.0 kHz; digital resolution 0.3, 1.2 or 0.1, 1.2, 1.2, and 1.2 Hz; repetition time, 7, 1.2, 2.5, 1.2, and 15 s; pulse width, 31, 9, 17, 9, 17 μs; pulse angle 90, 40, 30, 40, and 90°. The spectra were recorded in $(CD_3)_2SO$ with the solvent providing

* IUPAC name for the investigated compounds: 5-acylamino-5-alkyl-6-imino-2-thioxo-1,2,5,6-tetrahydropyrimidin-4(3*H*)-one.

the deuterium lock. The concentrations of the samples were *ca.* 30 mg cm^{-3} for ^{1}H and 100 mg cm^{-3} for ^{13}C spectra. Chemical shifts are given in ppm relative to internal TMS.

Acknowledgements

This work was supported by the Polish Academy of Sciences Project CPBP 01.13.1.3.

References

1 For a review, see: J. Elguero, C. Marzin, A. R. Katritzky, and P. Linda, in 'Advances in Heterocyclic Chemistry Supplement I,' Academic Press. New York, NY, 1976.
2 E. S. Raper, R. E. Oughtred, and J. W. Nowell, *Acta Crystallogr., Sect. C*, 1985, **41**, 758.
3 S. Aruna and G. Shanmugam, *Spectrochim. Acta, Part A*, 1985, **41**, 531.
4 S. Aruna and G. Shanmugam, *Indian J. Chem., Sect A*, 1986, **25**, 256.
5 B. Golankiewicz, J. B. Holtwick, B. N. Holmes, E. N. Duesler, and N. J. Leonard, *J. Org. Chem.*, 1979, **44**, 1740; J. B. Holtwick, B. Golankiewicz, B. N. Holmes, and N. J. Leonard, *ibid.*, 1979, **44**, 2835.
6 B. Golankiewicz, J. Zeidler, and E. De Clercq, *Nucleosides Nucleotides*, 1987, **6**, 663.
7 L. F. Johnson, in 'Topics in Carbon-13 NMR Spectroscopy,' ed. G. C. Levy, Wiley, New York, 1979, p. 7.
8 S. Takauchi, J. Uzawa, H. Seto, and H. Yonehara, *Tetrahedron Lett.*, 1977, 2943; J. Uzawa and M. Uramoto, *e rg. Magn. Reson.*, 1979, **12**, 612.
9 J. W. Still and N. Plavac, *Can. J. Chem.*, 1978, **56**, 725.
10 B. Lippert, H. Schöllhorn, and U. Thewalt, *J. Am. Chem. Soc.*, 1986, **108**, 6616 and references cited therein.
11 N. F. Albertson, *J. Am. Chem. Soc.*, 1946, **68**, 450.

Paper 9/01950B
Received 10*th May* 1989
Accepted 16*th January* 1990