Comparative Assessment of Structure and Reactivity of Wyosine by Chemistry, Spectroscopy and ab initio Calculations

ARTICLE *in* TETRAHEDRON · JANUARY 1996

Impact Factor: 2.64 · DOI: 10.1016/0040-4020(95)00988-4

Pergamon

Tetrahedron, Vol. 52, No. 5, pp. 1597-1608, 1996

0040-4020/96 $15.00 + 0.00

0040-4020(95)00988-4

Comparative Assessment of Structure and Reactivity of Wyosine by Chemistry, Spectroscopy and *ab initio* Calculations

Janez Plavec and Jyoti Chattopadhyaya*

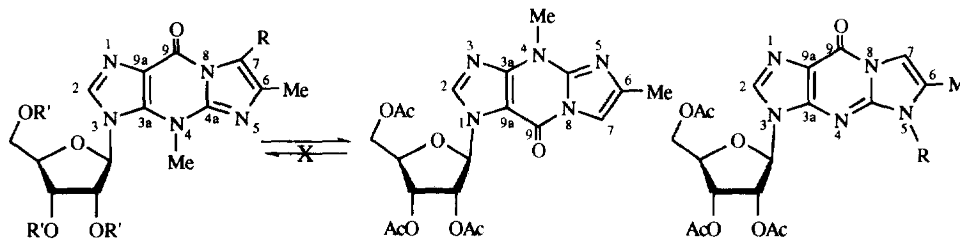
Department of Bioorganic Chemistry, Box 581, Biomedical Center,
University of Uppsala, S-751 23 Uppsala, Sweden

Dedicated to Professor T Hata on the occasion of his 60th birthday

Abstract: *The hypermodified fluorescent Y-nucleosides (Wyosine,* ***1a****) and its 7-substituted congeners occur naturally adjacent to the 3'-end of the anticodon loop in yeast phenylalanine transfer RNA (tRNA^{Phe}). The biological importance of Wyosine is emphasised by the fact that the removal of the aglycone (Y- or wye-base) from Y-nucleoside in tRNA^{Phe} by mild acidic treatment causes the loss of its codon recognition property required for protein biosynthesis. In this work, we have performed detailed ab initio calculations on several isomerically methylated Y-base analogs modelling the unique N4-methylation site of Y-base and the N5-methylated site of the N4-desmethyl counterpart. These results have been compared with our earlier spectroscopic and chemical studies on Wyosine (**1a**) (refs. 1,3), which can be summarized as follows: (i) HF/6-31G** calculations on model compounds **4** and **5** show that the former is preferred by 12.8 (vacuo), 10.3 (ε = 8.9) and 9.9 (ε = 78.3) kcal/mol which is in excellent agreement with our facile Lewis acid promoted isomerization of **1b** (R=H, R'=Ac) into its thermodynamically preferred N^1 isomer (**2**) (ref. 1). (ii) N3,N5-dimethyl derivative **6** exhibits an energetic preference (HF/6-31G** level) of 12.0 kcal/mol over the N3,N4-isomer **5**, which is consistent with the chemical methylation studies of 4-desmethylwyosine (**3a**) which gives N5-methyl-4-desmethylwyosine (**3b**) as the dominant product whereas wyosine (**1b**) is formed only in 3% yield (ref. 2). (iii) Relative HF/6-31G** energies show that the first protonation site is N5 for both N3,N4-dimethyl analog **8** and its N1,N4 isomer **10** (N5-H$^+$ species is preferred by 1.2 - 3.0 and 9.3 - 16.4 kcal/mol over N1-H$^+$ counterparts **9** and **11**), which is in excellent agreement with our ^{15}N-NMR titration studies with CF_3COOH (ref. 1). (iv) The examination of N^1-H (**14**) $\rightleftarrows$ N^3-H (**15**) tautomerism has shown that N^1-H tautomer **14** is energetically preferred by 5.8 to 8.8 kcal/mol which is consistent with our studies based on ^{13}C-NMR (ref. 4) and fluorescence decay time measurements (ref. 6) of wye-base (**14**). These studies suggest that the wyosine is a derivative of the thermodynamically unfavoured N^3-H tautomer (**15**). The biological consequence of the thermodynamically unstable natural wyosine is the extreme lability of its glycosyl bond under mild acidic conditions, which thus constitutes a switch for the deactivation of the codon function of tRNA^{Phe}. This finding of the inherent thermodynamic unstablility of natural wyosine also suggests that guanosine is most probably the building block for the biosynthesis of wyosine.*

The hypermodified fluorescent Y-nucleosides [4,9-dihydro-4,6-dimethyl-9-oxo-3-(β-D-ribofuranosyl)imidazo[1,2-a]purine] (Wyosine, **1a**) and its 7-substituted congeners (general formula: **1a**) occur naturally adjacent to the 3'-end of the anticodon loop in yeast phenylalanine transfer RNA (tRNA^{Phe}) and are characterized by their fluorescence properties and the extremely acid labile glycosidic bond. The removal of the aglycone (Y- or wye-base) from Y-nucleoside in tRNA^{Phe} by mild acidic treatment causes the loss of its codon recognition property required for protein biosynthesis. The physical and chemical properties of these Y-nucleosides have been extensively studied because of their distinctive biological properties.[1] The unique reactivity of Y-nucleosides posed a considerable challenge for a high yielding synthesis, and the complexity encountered are the following: (1) the direct glycosylation of wye-base [(**14**) $\rightleftarrows$ (**15**)] takes place exclusively at the undesired N^1 position, and (2) the methylation of 4-

desmethylwyosine-triacetate [5,9-dihydro-6-methyl-9-oxo-3-(2',3'-5'-tri-O-acetyl-β-D-ribofuranosyl)-5H-imidazo[1,2-a] purine] (**3a**) gives mainly N5-methyl-4-desmethylwyosine-triacetate (**3b**), while wyosine-triacetate (**1b**) is formed in only 3% yield.[2] Recently, however the regiospecific one-step N4-methylation of 4-desmethylwyosine-triacetate (**3a**) to wyosine-triacetate (**1b**) has been achieved in our laboratory in 74% yield in gram quantities.[3] This has enabled us to examine the acidic stability of its glycosidic bond compared to its analogs and other purine nucleosides and its chemical reactivity as well as the detailed ^{13}C and ^{15}N-NMR properties, thus exploring the structural uniqueness of the hypermodified aglycone wye-base in **1a**.[1,4-6] In this paper we extend those chemical and spectroscopic studies by standard *ab initio* molecular orbital (MO) calculations (Gaussiaan 92)[7] to further explore the basis of the chemical uniqueness of wyosine and wye-base to understand the importance of Y-nucleosides for codon recognition in the biologically ubiquitous $tRNA^{Phe}$. These *ab initio* calculations were initiated basing on the numerous examples in the literature that *ab initio* calculations at the Hartree-Fock (HF) level do reproduce the energetic trends as well as all bond lengths, bond angles and charge density effects associated with the structure of a tautomer.



1a: R' = H, R = H or Me or aminoacid*
1b: R = H or Me or aminoacid*
R' = Ac

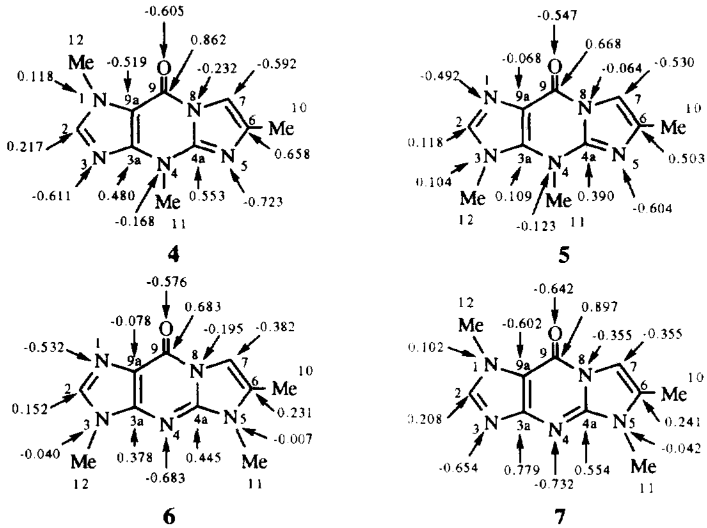
2

3a: R = H
3b: R = Me

* $CH_2CH(H/OH)CH(NHCO_2Me)CO_2Me$

Scheme 1. The Y-nucleosides and its 7-substituted congeners

The present *ab initio* examination of the observed chemical reactivities and spectroscopic properties of wyosine and its analogs has been based on the model systems **4** - **7** in which 2',3',5'-tri-O-acetyl-β-D-ribofuranosyl moiety at N1 or N3 has been replaced by the methyl group in order to reduce the computational time. Using methylated models **8** - **13**, we show that the preferred sites of protonation in **1b** and **2** monitored by ^{15}N-NMR chemical shifts are well reproduced by the energetics in *ab initio* calculations (HF/6-31G** level). We have subsequently examined the tautomerism problem of wye-base [*i.e.* N^1-H (**14**) ⇄ N^3-H (**15**)] and of the aglycone of N3-methyl-N4-desmethylwyosine [N^5-H (**16**) ⇄ N^4-H (**17**)] and show by *ab initio* calculations that the observed chemical reactivities of the above aglycones are related to the thermodynamic stability observed by ^{13}C-NMR[4] and fluorescence spectroscopy[6] of the participating tautomer at the equilibrium in solution (Scheme 5). All N1 and N3-methylated derivatives **4** -**7**, their protonated counterparts **8** - **13** and different tautomeric forms of the aglycones **14** - **19** were first completely optimized using 3-21G basis set. The HF/3-21G optimized geometries were used as input for



Scheme 2. Numbering of the atoms and their ESP point charges (shown by arrows)

full optimization at HF/6-31G** level. Three geometry optimizations were performed for **4** - **7**, two of which by self-consistent reaction field (SCRF) treatment with dielectric constants of 8.9 (to mimic ^{15}N-NMR protonation studies in CH_2Cl_2) and 78.3 (to mimic the aqueous environment).

(A) Structure of N1 and N3-methyl wye-base 4 and 5, respectively, modelling wyosine-triacetate 1b and its N1 isomer 2.

During our studies on C-7 functionalization of wyosine-triacetate (**1b**), we observed that it could be quantitatively transformed into its N^1 isomer (**2**) when treated with anhydrous $AlCl_3$ in dry CH_2Cl_2, whereas the latter was completely stable under this reaction condition.[1] That suggested that the N^1 isomer **2** is thermodynamically more stable than the natural N^3 counterpart **1b**. In the present studies we have tried to understand this thermodynamic preference of N^1 isomer **2** over N^3 isomer **1a** or **1b** using *ab initio* calculations.

The HF/6-31G** relative energies of **4** - **7** (Table 1) show that (i) N1,N4-dimethylimidazo[1,2-a]purine derivative (**4**) is energetically favoured over its N^3-Me isomer **5** by 12.8, 10.3 and 9.9 kcal/mol in *vacuo*, with SCRF ε = 8.9 and ε = 78.3, respectively, which is in excellent agreement with our isomerization studies on N3-(2',3',5'-tri-O-acetyl-β-D-ribofuranosyl)wyosine (**1b**) which is upon treatment with anhydrous $AlCl_3$ in dry CH_2Cl_2 quantitatively transformed into its N^1 isomer (**2**) while the reverse reaction is virtually non-existent. (ii) N3,N5-dimethylimidazo[1,2-a]purine derivative (**6**) is thermodynamically more stable, and therefore shows a high energetic preference (12.0 kcal/mol in *vacuo*, with SCRF ε = 8.9 and ε = 78.3) over the isomeric N3,N4-dimethylimidazo[1,2-a]purine (**5**), which is consistent with the result of the chemical methylation studies showing that 4-desmethylwyosine (**3a**) gives mainly N5-methyl-4-desmethylwyosine (**3b**) as the dominant product whereas wyosine (**1b**) is

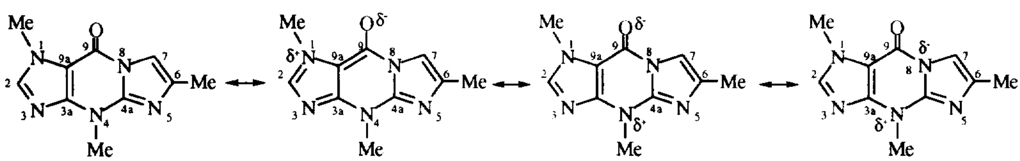
Table 1. Total and relative HF/6-31G** energies[a] and dipole moments[b] of optimized structures of **4** - **7**

Compound	Optimization *in vacuo*			SCRF (ε = 8.9) optimization			SCRF (ε = 78.3) optimization		
	HF energy	Relative energy[c]	Dipole moment	SCRF energy	Relative energy[c]	Dipole moment	SCRF energy	Relative energy[c]	Dipole moment
4	-732.24291	0.0[c]	3.57	-732.24389	1.7	4.27	-732.24411	2.1	4.42
5	-732.22256	12.8	8.03	-732.22742	12.0	9.30	-732.22839	12.0	9.56
6	-732.24158	0.8	8.17	-732.24655	0.0[c]	9.45	-732.24753	0.0[c]	9.70
7	-732.23980	2.0	2.87	-732.24043	3.8	3.44	-732.24057	4.4	3.54

[a] Energies in au (= 627.5095 kcal/mol), relative energies in kcal/mol. [b] In units of Debye. [c] The lowest energy structure is taken as zero for the calculation of relative energies.

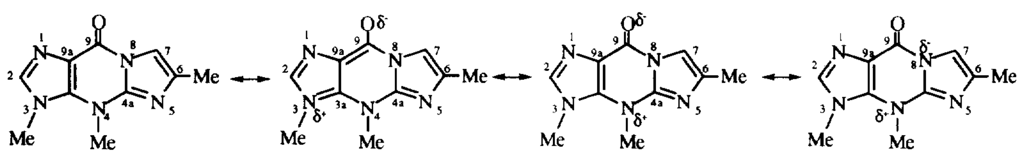
formed only in 3% yield.[2] (iii) The above direct correlation of HF/6-31G** energy trends with the observed chemical reactivities along with the consideration of the relative HF energies found on N1, N4-dimethylimidazo[1,2-a]purine derivative (**4**) and N1, N5-dimethylimidazo[1,2-a]purine derivative (**7**) show that the former is more stable than the latter by 2.0, 2.1 and 2.3 kcal/mol *in vacuo*, with SCRF ε = 8.9 and ε = 78.3 (>95% preference of **4** over **7** in terms of Boltzman distribution), which suggest that the methylation of N1-(2',3',5'-tri-O-acetyl-β-D-ribofuranosyl)-N4-desmethylwyosine should give mainly the corresponding N4-methylated product **2**. (iv) The energetic preference of N^1-Me derivative **4** over the isomeric N^3-Me counterpart **5** has been attributed to different degree of delocalization originating from the following factors, which suggest that various degree of resonance stabilization is produced by different participating resonance forms: (a) The exchange of the lone pair/σ-bond (n/σ) and σ-bond/σ-bond (σ/σ) interactions[10,11] in the 'left' imidazole part of N^1-Me (**4**) and N^3-Me (**5**) isomers result in a more facile delocalization of π-excessive imidazole part into π-deficient central pyrimidine part in the former which is reflected in smaller difference between N1-C2 and C2-N3 bond distances (Table 5) in **4** (0.02 Å) compared to those of **5** (0.11 Å). (b) The comparison of ESP charges (Besler-Merz-Kollman charges)[8] on N^1 (0.118) and N^3 (-0.611) in **4** with N^3 (0.104) and N^1 (-0.492) in **5**, respectively, shows that the σ/σ interactions of N1 and n/σ interactions of N3 is considerably more delocalized through the sp^2 hybridized C^2 and pyrrole-nitrogen in **4** (resonance structure A) than in **5** (resonance structure A') in Scheme 3, resulting into larger delocalization of π-charge which is evident from the larger negative charge on N^3 in the former. This also results in a slightly larger negative charge on O^9 (-0.605) and N^8 (-0.232) in **4** compared to those of **5** (-0.547 and -0.064, respectively). Therefore the canonical structure B (Scheme 3) should play more important role in the delocalization of π-excessive imidazole part into π-deficient central pyrimidine in **4** than the canonical structure B' for **5**. The lone pair of N^4 may also be delocalized either through the pyrimidine part as represented in the canonical structures C or C' or towards 'right' imidazole part through the canonical structure D or D' as shown in Scheme 3. The delocalization pathways in C/C' or D/D' should also result in the overall decrease of the net negative charge on N^4 and consequently increase of the negative charge on O^9 or N^8. The HF/6-31G** calculations however show that N^4 has a slightly more negative charge in N^1-Me derivative **4** (-0.168) than in the isomeric N^3-Me counterpart **5** (-0.123) whereas the charge densities of O^9 and N^8 are larger in **4** than **5**. This effective charge separation in **4** is evident from the experimental single bond character of its C^9=O (IR: $\nu_{C=O}^{CCl_4}$ 1689 cm^{-1} for **4** and 1701 cm^{-1} for **5**) and hence canonical structures C and D for **4** and B', C' and D' for **5** are probably of minor importance. The ^{15}N-NMR evidence[1] also supports the relatively larger delocalization of the 'left' imidazole system in **4** than **5**, which can be seen from the larger deshielding of the N1 (δ -213.7 ppm) and shielding of N3 (δ -151.3 ppm) in the former compared to the counterparts (N1: δ -130.5 ppm and N3: δ -220.6 ppm) in the latter. Thus the energy gained by the preferential resonance stabilization through the canonical structures A↔B for **4** is presumably one of the main reasons for the overall relative energetic preference of **4** compared to **5** (Table 1). (c) It is noteworthy that we have also observed a good qualitative correlation of the nuclear magnetic resonance absorption (chemical shift δ in ppm)) of the carbon, proton or nitrogen atom (Q) in the aromatic systems with that of the sum of the atomic charges around the central atom (Q) and the central atom itself, Σe {Σe = Σ(ESP charges of all neighbouring atoms plus on the central atom, Q)}. Thus the comparison of the relative negative or positive charges on the nitrogen atoms in **4** and **5** shows that as the total negative charge, Σe, increases, the more magnetic shielding of the central atom, Q, is observed in the NMR which is also consistent with our observation that a reduction of the negative charge in Σe deshields the central atom, Q (see Table 2). (d) A perusal of total ESP charges on the N1-C2-N3 atoms of the 'left' imidazole (-0.276) with N5-C6-C7 atoms of the 'right' imidazole (-0.657) also clearly shows that the former is more deactivated than the latter in **4**, which is also the true in **5** (-0.270 and -0.631, respectively). This is also consistent with the general chemical reactivities observed for **1b** that while a carbanion could be generated easily at C2 of the 'left' imidazole, electrophilic substitution reactions at C7 of the 'right' imidazole were found to be the dominant features.[1]

Possible resonance structures for **4**



A **B** **C** **D**

Possible resonance structures for **5**



A' **B'** **C'** **D'**

Scheme 3

Table 2. ^1H, ^{13}C and ^{15}N-NMR chemical shift (δ) of **2** and **1b** (R=H, R'=Ac) changes as a function of the total charge[a] (Σe) calculated from HF/6-31G** optimized **4** and **5**

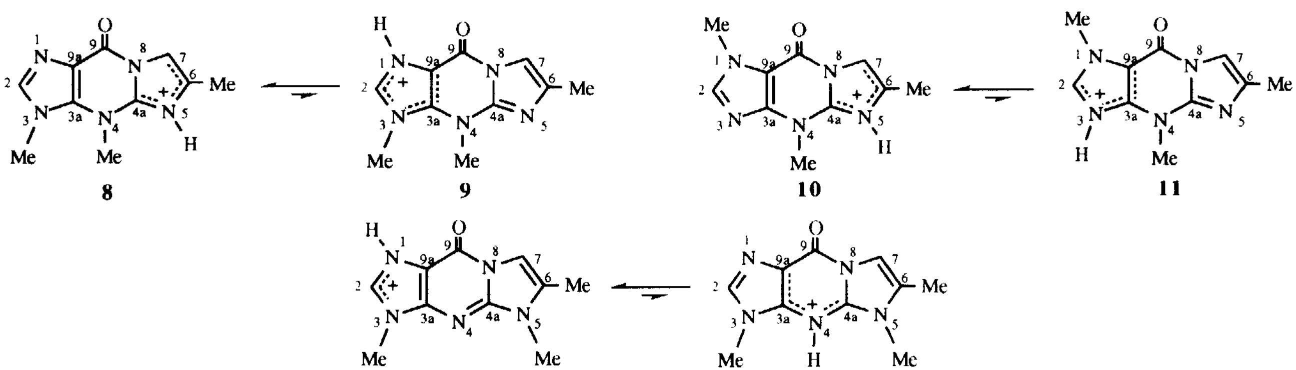
	C-2		H-2	N^4-CH_3		C9		N4		N5	
Compd.	Σe	$\delta^{13}C$	δ^1H	Σe	$\delta^{13}C$	Σe	$\delta^{13}C$	Σe	$\delta^{15}N$	Σe	$\delta^{15}N$
4	-0.129	140.3	8.07	-0.014	31.0	-0.494	149.4	0.698	-283.0	0.488	-152.4
5	-0.120	133.6	7.73	0.029	33.9	-0.011	151.8	0.225	-290.3	0.289	-158.2

[a] Total charge (Σe) signifies the sum of Besler-Merz-Kollman[8] point atomic charges of neighbors around Q atom and the central Q itself whose chemical shift[1,3] is given.

Table 3. Total and relative HF/6-31G** energies[a] and dipole moments[b] of completely optimized protonated structures **8** - **13** and changes in ^{15}N-NMR chemical shifts (Δδ) upon protonation[c]

	Ab initio calculations												Spectroscopic data			
	Optimization *in vacuo*			SCRF (ε = 8.9) optim'n			SCRF (ε = 46.45) optim'n			SCRF (ε = 78.3) optim'n			^{15}N-NMR [c]			
Compound	HF energy	Relative energy	Dipole moment	SCRF energy	Relative energy	Dipole moment	SCRF energy	Relative energy	Dipole moment	SCRF energy	Relative energy	Dipole moment	ΔδN1	ΔδN5	ΔδN3	ΔδN4
8	-732.60895	21.5	10.91	-732.61795	17.6	12.52	-732.61949	17.5	12.79	-732.62009	16.7	12.90	2.4	39.9	-2.9	-2.4
9	-732.60410	24.5	12.05	-732.61545	19.2	14.29	-732.61755	18.7	14.71	-732.61787	18.1	14.77				
10	-732.63805	3.2	4.51	-732.63961	4.0	5.26	-732.63990	4.7	5.40	-732.63994	4.3	5.44	-1.2	45.7	0.1	-1.1
11	-732.61196	19.6	11.80	-732.62279	14.6	13.92	-732.62478	14.2	14.31	-732.62499	13.6	14.35				
12	-732.64318	0.0	6.78	-732.64598	0.0	7.98	-732.64738	0.0	8.21	-732.64673	0.0	8.24	36.6	-2.8	-4.0	-1.2
13	-732.61617	16.9	10.95	-732.62525	13.0	12.58	-732.62687	12.9	12.88	-732.62704	12.4	12.91				

[a] Energies in au (= 627.5095 kcal/mol), relative energies in kcal/mol. The lowest energy structure is taken as zero for the calculation of relative energies. [b] In Debyes. [c] Upfield shift of the corresponding signal in ^{15}N-NMR spectrum of N1- and N3-(2',3',5'-tri-O-acetyl-β-D-ribofuranosyl) substituted compounds **2** and **1b** (R=H, R'=Ac), respectively in CH_2Cl_2 upon addition of 1 equiv. of TFA.[1]



8 **9** **10** **11** **12** **13**

Scheme 4

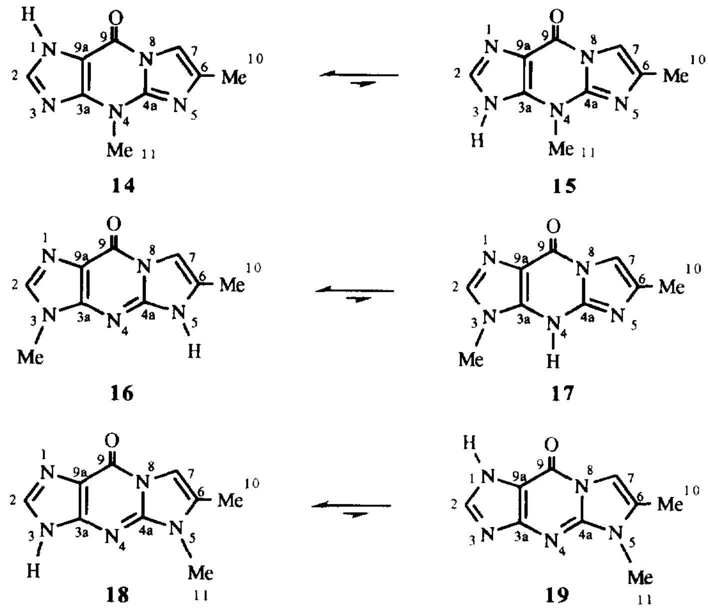
(B) Protonation behavior of N1 and N3-methylated isomers 4 and 5.

^{15}N-NMR spectroscopy has been used[1,5] to establish the site and the magnitude of protonation by the comparison of ^{15}N chemical shifts between protonated and neutral species ($\Delta\delta$). These ^{15}N-NMR titration studies[1] with trifluoroacetic acid (TFA) have shown that N5 is almost exclusively protonated in both N1-(2',3',5'-tri-O-acetyl-β-D-ribofuranosyl)wyosine (**2**) or N3-(2',3',5'-tri-O-acetyl-β-D-ribofuranosyl)wyosine (**1b**), but the magnitude of $\Delta\delta$N5-H$^+$ are indeed considerably different ($\Delta\delta$N5-H$^+$ = 39.9 ppm for **1b** and 45.7 ppm for **2**). Our *ab initio* caluclations also supports these ^{15}N-NMR spectroscopic trends on **1b** and **2** very nicely through our studies of the various protonated forms of the model methylated systems **8** -**11**, showing that N5-H$^+$ species **8** and **10** are the principal site of protonation, which is evident from the relative HF/6-31G** energies of freely optimized isomeric protonated pairs **8** ⇄ **9** and **10** ⇄ **11** (Scheme 4). A comparison of HF/6-31G** relative energies show that N5-H$^+$ species **8** is indeed preferred by 1.6 (ε = 8.9), 1.2 (ε = 46.45) and 1.4 (ε = 78.3) kcal/mol over N1-H$^+$ counterpart **9** with SCRF whereas the energy difference is 3.0 kcal/mol in vacuum (Table 3). Protonated N5-H$^+$ species in **10** is energetically preferred over N3-H$^+$ species **11** by 10.6 (ε = 8.9), 9.5 (ε = 46.45) and 9.3 (ε = 78.3) kcal/mol with SCRF and by 16.4 kcal/mol in vacuum which explains $\Delta\delta$(N5) of 45.7 ppm and negligible $\Delta\delta$(N3) (0.1 ppm; Table 3) as the result of N5 protonation. Hence, it can be seen that the model N5-protonated methyl derivative **10** (*i.e.* an analogue of N5-protonated **2**) is more stable than the isomeric **8** (*i.e.* an analogue of N5-protonated **1b**) by 13.6 (ε = 8.9), 12.8 (ε = 46.45), 12.4 (ε = 78.3) and 18.3 kcal/mol in vacuum, which is consistent with the relative $\Delta\delta$(N5) of **2** and **1b** (Table 3). The ^{15}N-NMR protonation studies[1] showed that N1 is exclusively protonated in N3-(2',3',5'-tri-O-acetyl-β-D-ribofuranosyl)-N4-desmethyl-N5-methylwyosine (**3b**) (Table 3), which is evident from its large upfield shift: $\Delta\delta$(N1) = 36.6 ppm. HF/6-31G** relative energies clearly show that N1-H$^+$ species in model **12** (*i.e.* an analogue of N5-protonated **3b**) is indeed preferred by 13.0 (ε = 8.9), 12.9 (ε = 46.45) and 12.4 (ε = 78.3) kcal/mol over N4-H$^+$ species **13** (*i.e.* an analogue of N4-protonated **3b**) with SCRF and that energy preference rises to 16.9 kcal/mol in vacuum which is in excellent agreement with $\Delta\delta$(N1) of 36.6 ppm and small downfield shift of 1.2 ppm for N4 (Table 3).

(C) Tautomerism of wye-base

Detailed ^{13}C-NMR study[4] of the prototropic equilibrium of N^1-H (**14**) ⇄ N^3-H (**15**) tautomerism in solution, through the measurement of time-averaged δ_{C3a} and δ_{C9a} chemical shifts and $^3J_{C9a,H2}$ and $^3J_{C3a,H2}$ coupling constants, has shown that the population of N^1-H tautomer of wye-base (**14**) is at least >95%.[4] These data are also consistent with our fluorescence decay time measurements of N^1-H (**14**) ⇄ N^3-H (**15**) tautomeric equilibrium, which have shown that N^1-H (**14**) tautomer is by far the predominant (>95%) component in DMSO solution.[4] It was clear to us that the chemical consequence of the thermodynamical preference of N^1-H (**14**) tautomeric form is that wye-base is regioselectively glycosylated or methylated under the neutral condition (MeI in DMF at RT) at N^1 position. Therefore, the biological consequence of natural wyosine being glycosylated at the N^3 position [*i.e.* wyosine is a derivative of the thermodynamically unfavoured N^3-H tautomer (**15**)] is that the extreme lability of its glycosyl bond under mild acidic conditions constitutes the switch for the deactivation of the codon function of tRNA^{Phe}.

In 14 ⇄ 15 prototropism (Scheme 5), N^1-H tautomer **14** is preferred by 5.8 to 8.6 kcal/mol depending on the dielectric constant applied in SCRF calculation (Table 4). The energy difference between **14** and **15** is the largest in vacuum and decreases by the increase of the dielectric constant in the SCRF treatment.



Scheme 5. Tautomerism in imidazo[1,2-a]purines

Table 4. Total and relative MP2/6-31G**//6-31G** and HF/6-31G** energies[a] and dipole moments[b] of completely optimized structures of **14** - **19**

Compd.	SCRF (ε = 78.3) optimization			SCRF (ε = 8.9) optimization			Optimization in *vacuo*			
	SCRF energy	Relative energy	Dipole moment	SCRF energy	Relative energy	Dipole moment	MP2[c] energy	Relative energy	HF energy	Relative energy
14	-693.20979	0.0	4.09	-693.20960	0.0	3.96	-695.37854	0.0	-693.20865	0.0
15	-693.20051	5.8	8.88	-693.19958	6.3	8.65	-695.36445	8.8	-693.19497	8.6
16	-693.21486	0.0	9.13	-693.21391	0.0	8.90	-695.37611	0.0	-693.20895	0.0
17	-693.20141	8.4	9.13	-693.20046	8.4	8.90	-695.36346	7.9	-693.19558	8.4
18	-693.21399	0.0	9.37	-693.21300	0.0	9.12	-695.37627	0.5	-693.20787	0.0
19	-693.20827	3.6	3.71	-693.20812	3.1	3.58	-695.37700	0.0	-693.20734	0.3

[a] Energies in au (= 627.5095 kcal/mol), relative energies in kcal/mol. [b] In units of Debye. [c] MP2/6-31G**//6-31G** level.

The tautomeric equilibrium of 16 ⇄ 17 was studied in a similar manner since we wished to see if the *ab initio* theory can model the result of chemical methylation of 4-desmethylwyosine (**3a**) which gives mainly N5-methyl-4-desmethylwyosine (**3b**), while wyosine (**1b**) is formed in only 3% yield.[2] Note that the full energy minimizations of **16** and **17** at HF/6-31G** level of the theory showed that N^5-H tautomer **16** is energetically preferred by 8.4 kcal/mol (Table 4) and the energy difference between **16** and **17** does not change with the change of dielectric constant in the SCRF calculations. The preference of the tautomeric form **16** over **17** is presumably related to the relatively larger loss of the aromatic stabilization energy owing to the protonated imidazole character in **17** compared to the protonated pyrimidine in **16**.

The **18** $\rightleftarrows$ **19** tautomeric equilibrium was also studied in a similar manner. In the optimizations with SCRF with $\varepsilon = 78.3$ and 8.9, it is found that N^3-H tautomer **18** is preferred by 3.1 - 3.6 kcal/mol, which is again consistent with the chemospecific glycosylation at N1 over N3 (Table 4) in N4-desmethyl-5-methylwyosine.

Table 5. Bond lengths[a] of HF/6-31G** optimized **4** - **7** and **14**, **15**, **18** and **19**

Compound	4	14	5	15	6	18	7	19
N1-C2	1.328	1.331	1.272	1.274	1.276	1.276	1.335	1.338
C2-N3	1.308	1.305	1.381	1.380	1.376	1.374	1.300	1.297
N3-C3a	1.346	1.350	1.359	1.350	1.354	1.354	1.357	1.364
C3a-N4	1.363	1.363	1.368	1.357	1.349	1.343	1.355	1.353
N4-C4a	1.366	1.366	1.378	1.374	1.299	1.300	1.289	1.292
C4a-N5	1.283	1.284	1.283	1.282	1.349	1.344	1.354	1.352
N5-C6	1.387	1.386	1.388	1.386	1.401	1.400	1.397	1.398
C6-C7	1.341	1.341	1.341	1.341	1.329	1.331	1.329	1.329
C7-N8	1.397	1.396	1.395	1.396	1.397	1.400	1.400	1.401
N8-C9	1.394	1.396	1.405	1.408	1.421	1.416	1.399	1.399
C9-O9	1.199	1.197	1.190	1.193	1.195	1.198	1.206	1.203
C9-C9a	1.432	1.430	1.446	1.443	1.431	1.431	1.421	1.419
C4a-N8	1.372	1.372	1.363	1.366	1.359	1.360	1.368	1.369
C3a-C9a	1.369	1.365	1.369	1.366	1.378	1.380	1.382	1.377
C6-C10	1.493	1.493	1.493	1.493	1.492	1.492	1.492	1.492
C10-H10a	1.085	1.085	1.085	1.085	1.082	1.085	1.085	1.085
C10-H10b	1.083	1.084	1.083	1.084	1.085	1.082	1.082	1.081
C10-H10c	1.085	1.085	1.085	1.085	1.085	1.085	1.085	1.085
N4/5-C11	1.455 ^{N4}	1.455 ^{N4}	1.456 ^{N4}	1.455 ^{N4}	1.447 ^{N5}	1.452 ^{N5}	1.447 ^{N5}	1.449 ^{N5}
C11-H11a	1.077	1.077	1.082	1.084	1.078	1.081	1.078	1.078
C11-H11b	1.081	1.081	1.082	1.076	1.083	1.078	1.083	1.083
C11-H11c	1.081	1.081	1.076	1.084	1.083	1.081	1.083	1.083
C2-H2	1.071	1.071	1.071	1.070	1.072	1.071	1.072	1.072
C7-H7	1.066	1.066	1.066	1.066	1.066	1.065	1.065	1.066
N1/3-C12/H	1.451 ^{N1}	0.993 ^{N1}	1.451 ^{N3}	0.993 ^{N3}	1.446 ^{N3}	0.993 ^{N3}	1.452 ^{N1}	0.992 ^{N1}
C12-H12a	1.081		1.080		1.083		1.077	
C12-H12b	1.080		1.081		1.078		1.083	
C12-H12c	1.080		1.081		1.083		1.083	

[a] In Å.

The effects of electron correlation as estimated by MP2/6-31G**//6-31G** calculations on **14** - **19** do not affect the observed energetic preferences in **14** $\rightleftarrows$ **15**, **16** $\rightleftarrows$ **17** and **18** $\rightleftarrows$ **19** tautomeric equilibria considerably. Table 4 shows that the relative energies of **15** compared to **14** and **16** compared to **17** differ by < 0.5 kcal/mol with inclusion of electron correlation effects. In the case of **18** $\rightleftarrows$ **19** however, **18** is preferred over **19** by 0.3 kcal/mol within HF treatment, whereas **19** is energetically favoured over **18** by 0.5 kcal/mol at MP2/6-31G** level (Table 4).

Why is the N^1-H tautomer of wye-base is energetically more preferred over the N^3-H tautomer? The reason for this energetic preference can be traced back to the resonance energy gained by having the N3 locked in the N^3-H protonated form. Here again the basic electron delocalization argument given for the canonical structure A for the methylated structure **4** is valid compared to **5** (Scheme 3). The supporting evidence for the resonance energy gained by a more extensive delocalization of the 'left' imidazole with the electron-deficient central pyrimidine ring is evident from the comparison of negative charges on N1 (-0. 483), N3 (-0.646), N4 (-0. 386) and C9a (-0.319) in N^1-H tautomer **14** *versus* N1 (-0. 584), N3 (-0.506), N4 (-0. 231) and C9a (0.102) in N^3-H tautomer **15**, showing that there are increases of the negative charge at N3, N4 and C9a as the negative charge decreases at N1 in **14** in comparison with **15**.

(D) HF/6-31G optimized bond lengths, bond angles and torsion angles.**

An inspection of HF/6-31G** optimized bond lengths in **4** and **5** (Table 5) show the following trends: (i) The N1-C2 bond (1.328 Å) in **4** is considerably shorter than its equivalent C2-N3 bond (1.381 Å) in **5**. (ii) The C2 -N3 bond (1.308 Å) in **4** is considerably longer than its equivalent N1-C2 bond (1.272 Å) in **5**. (iii) The N3 -C3a bond (1.346 Å) in **4** is slightly shorter than the corresponding bond (1.359 Å) in **5**. (iv) The N4-C4a bond (1.366 Å) in **4** is slightly shorter than the corresponding bond (1.378 Å) in **5**. (v) The C9-C9a bond (1.432 Å) in **4** is slightly shorter than the corresponding bond (1.446 Å) in **5**. The shortening of N1-C2, N3-C3a, N4-C4a, C9-C9a single bonds in **4** compared to **5** are directly the result of larger electron delocalization of the π-excessive 'left' imidazole. In fact, the lengthening of C2-N3 double bond in **4** compared to N1-C2 double bond in **5** is also consistent with a larger delocalization in the former. In contrast, the C4a-N8, C4a-N5, N5-C6, C6-C7, C7-N8 bonds of the 'right' imidazole show only very small (< 0.01Å) difference in their bond lengths in **4** and **5**, which suggest, as

Table 6. Bond angles[a] of HF/6-31G** optimized **4** - **7** and **14**, **15**, **18** and **19**

Compound	4	14	5	15	6	18	7	19
N3-C2-N1	114.35	113.55	113.63	112.64	113.61	113.07	114.85	114.13
C3a-N3-C2	104.01	104.33	105.08	105.97	105.86	106.70	104.56	104.84
N4-C3a-N3	126.22	126.24	129.22	127.53	125.54	125.72	125.10	125.46
C4a-N4-C3a	116.39	116.31	115.29	115.09	110.61	110.36	112.16	112.08
N5-C4a-N4	126.57	126.59	126.29	126.80	126.73	126.17	126.77	126.67
C6-N5-C4a	105.19	105.17	105.12	105.06	109.03	108.97	109.25	109.34
C7-C6-N5	110.87	110.93	110.56	110.64	107.60	107.61	108.00	107.95
N8-C7-C6	105.68	105.65	105.78	105.86	107.54	107.60	107.31	107.31
C9-N8-C7	127.59	127.55	127.16	126.98	125.85	126.39	126.52	126.68
O9-C9-N8	122.33	122.59	121.12	120.75	119.51	119.88	120.94	121.91
C9a-C9-N8	109.29	109.05	109.88	109.84	108.75	108.85	108.56	108.25
C10-C6-C7	128.83	128.79	129.05	129.11	130.27	129.67	129.93	129.97
H10a-C10-C6	110.46	110.47	110.56	110.51	109.50	111.47	111.32	111.31
H10b-C10-C6	111.02	111.01	110.92	110.95	111.44	109.34	109.63	109.64
H10c-C10-C6	110.46	110.47	110.56	110.51	111.44	111.47	111.32	111.31
C11-N4/5-C3a/4a	123.68 [b]	123.64 [b]	125.94 [b]	123.44 [b]	124.05 [c]	122.45 [c]	123.57 [c]	123.69 [c]
H11a-C11-N	108.00 ^{N4}	108.00 ^{N4}	110.96 ^{N4}	110.41 ^{N4}	107.95 ^{N5}	110.11 ^{N5}	107.74 ^{N5}	107.79 ^{N5}
H11b-C11-N	109.95 ^{N4}	109.94 ^{N4}	110.96 ^{N4}	107.93 ^{N4}	110.69 ^{N5}	109.13 ^{N5}	110.73 ^{N5}	110.65 ^{N5}
H11c-C11-N	109.95 ^{N4}	109.94 ^{N4}	107.20 ^{N4}	110.41 ^{N4}	110.69 ^{N5}	110.19 ^{N5}	110.73 ^{N5}	110.65 ^{N5}
H2-C2-N1	121.78	122.34	126.01	126.33	125.50	125.66	121.18	121.67
H7-C7-N8	120.96	120.94	120.83	120.80	120.64	120.79	120.85	120.87
C12 / H-N1/3-C2	127.69 ^{N1}	127.51 ^{N1}	123.56 ^{N3}	126.49 ^{N3}	127.18 ^{N3}	127.83 ^{N3}	127.28 ^{N1}	127.89 ^{N1}
H12a-C12-N	108.42 ^{N1}		107.54 ^{N3}		110.64 ^{N3}		108.30 ^{N1}	
H12a-C12-N	110.16 ^{N1}		111.47 ^{N3}		108.23 ^{N3}		110.40 ^{N1}	
H12a-C12-N	110.16 ^{N1}		111.47 ^{N3}		110.61 ^{N3}		110.40 ^{N1}	

[a] In degrees. [b] C11-N4-C3a. [c] C11-N5-C4a.

pointed out above, that the aromaticity of the 'right' imidazole is relatively unperturbed. A perusal of torsion angles in **4** and **5** (Table 7) shows that they are all perfectly planar (the deviation from planarity is < ±0.03°). Note that there is a very small difference in the optimized bond lenghts (Table 5), bond angles (Table 6) and torsion angles (Table 7) between **4** and **5** and the corresponding tautomers **14** and **15**, respectively.

An inspection of HF/6-31G** optimized bond lengths in **6** and **7** (Table 5) show the following trends: (i) The N1-C2 bond (1.276 Å) in **6** is considerably shorter than its equivalent C2-N3 bond (1.300 Å) in **7**. (ii) The C2 -N3 bond (1.376 Å) in **6** is longer than N1-C2 bond (1.335 Å) in **7**. (iii) The C3a-N4 bond (1.349 Å) in **6** is slightly shorter than the corresponding bond (1.355 Å) in **7**. (iv) The C4a-N5 bond (1.349 Å) in **6** is slightly shorter than the corresponding bond (1.354 Å) in **7**. (v) The N8-C9 bond (1.420 Å) in **6** is slightly longer than in **7** (1.399 Å). There is a very small difference in the optimized bond lenghts (Table 5), bond angles (Table 6) and torsion angles (Table 7) between **6** and **7** and the corresponding tautomers **18** and **19**, respectively.

Table 7. Torsion angles[a] of HF/6-31G** optimized **4** - **7** and **14**, **15**, **18** and **19**

Compound	4	14	5	15	6	18	7	19
C3a-N3-C2-N1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N4-C3a-N3-C2	179.99	180.00	180.01	180.00	180.00	179.99	180.00	179.99
C4a-N4-C3a-N3	180.00	-179.99	-179.99	-179.99	-179.99	-179.99	180.00	180.00
N5-C4a-N4-C3a	179.99	180.00	179.98	180.00	179.99	180.01	180.00	180.00
C6-N5-C4a-N4	180.00	180.00	179.99	180.00	179.99	180.03	180.00	180.00
C7-C6-N5-C4a	0.00	0.00	0.00	0.00	0.00	-0.03	0.00	0.00
N8-C7-C6-N5	0.00	0.00	0.00	-0.01	0.00	0.01	0.00	0.00
C9-N8-C7-C6	179.99	179.99	179.97	180.00	180.00	179.99	179.99	179.99
O9-C9-N8-C7	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00
C9a-C9-N8-C7	180.00	180.00	180.01	180.00	179.99	180.02	180.00	180.00
C10-C6-C7-N8	179.99	180.00	179.98	179.99	180.00	179.99	180.00	179.99
H10a-C10-C6-C7	-120.56	-120.55	-120.38	-120.54	0.00	-119.45	-119.74	-119.71
H10b-C10-C6-H10a	120.56	120.55	120.47	120.55	119.64	119.53	119.74	119.72
H10c-C10-C6-H10a	-118.87	-118.88	-119.04	-118.89	-119.64	-120.93	-120.51	-120.55
C11-N-C-N	0.00 [b]	0.00 [b]	-0.10 [b]	0.00 [b]	0.00 [c]	-0.09 [c]	0.00 [c]	0.00 [c]
H11a-C11-N-C	0.01 [d]	0.02 [d]	-61.25 [d]	60.76 [d]	0.01 [e]	59.35 [e]	0.03 [e]	0.04 [e]
H11b-C11-N-C11a	120.14	120.10	122.48	119.19	119.37	119.92	119.33	119.37
H11c-C11-N-C11a	-120.14	-120.10	-118.76	-121.60	-119.37	-120.09	-119.33	-119.37
H2-C2-N1-C9a	180.00	179.99	180.00	179.99	179.99	180.00	179.99	180.00
H7-C7-N8-C9	0.00	0.00	-0.01	0.00	0.00	-0.01	0.00	0.00
C12 / H-N-C-N	180.00 [f]	179.99 [f]	180.04 [g]	180.00 [g]	180.04 [g]	179.99 [g]	180.00 [f]	179.99 [f]
H12a-C12-N-C	0.00 [h]		-0.10 [i]		60.79 [i]		180.02 [h]	
H12b-C12-N-H12a	119.98		118.49		119.45		119.53	
H12c-C12-N-H12a	-119.98		-118.49		-121.11		-119.53	

[a] In degrees. [b] C11-N4-C3a-N3. [c] C11-N5-C4a-N4. [d] H11a-C11-N4-C3a. [e] H11a-C11-N5-C4a. [f] C12-N1-C2-N3. [g] C12-N3-C2-N1. [h] H12a-C12-N1-C2. [i] H12a-C12-N3-C2.

The solvent reaction field (polar *vs.* non-polar medium) has a strong influence on the molecular geometry if the compound exists mainly in the ionic form. We have in this work employed SCRF to mimic solvent environment of polar water ($\varepsilon = 78.3$) or non-polar CH_2Cl_2 ($\varepsilon = 8.9$) for **4** - **7** and **14** - **19** for full geometry optimizations at HF/6-31G** level. In addition the complete optimizations with SCRF of DMSO ($\varepsilon = 46.45$) were performed for **8** - **13**. These studies have shown that the medium has virtually no effect on the molecular geometries, which is also evident from virtually no change in the bond lengths, angles and torsion angles and also from small change in the resultant dipole moments and energies of the fully minimized structures at HF/6-31G** level of calculations. The small energy differences and the dipole moments between isomeric or tautomeric sets of compounds optimized with different SCRF (ε = 8.9 or 78.3) are therefore originating mainly from the minor differences in the solvation behaviour of the neutral structures **4** - **7** and **14** - **19**.

Experimental

Ab initio calculations were performed using the GAUSSIAN 92 program[7] with Silicon Graphics Indigo R4000 computers. For all the compounds all internal degrees of freedom were freely optimized with 3-21G basis set in the first step and optimized values were used as an input for the geometry optimizations at HF/6-31G** level. The self-consistent reaction field (SCRF) was used to mimic the presence of a solvent. In the reaction field the solvent is represented by a continuous dielectric characterized with dielectric constant (8.9 for CH_2Cl_2 and 78.3 for H_2O)[9,12]. The radius of spherical cavity (a_0) was obtained from the molecular volume to which 0.5 Å was added to account for the nearest approach of the solvent molecules according to the suggestions of Wiberg *et al.*[9] Net atomic point charges were calculated using a Besler-Merz-Kollman method[8] which performs a least squares fit of the quantum mechanically calculated electrostatic potential (ESP) to that of the charge model. A set of atomic charges was calculated through charge fitting with 6-31G** basis set on 6-31G** optimized geometries. The effects of electron correlation on energy difference between tautomeric pairs **14** - **19** were calculated through single-point energy calculations at MP2 level using 6-31G** optimized structures.

Acknowledgments

We thank Swedish Natural Science Research Council, Swedish Board for Technical Development (NUTEK), Wallenbergsstiftelsen, Forskningsrådsnämnden and Uppsala University for generous financial support.

References

1. Glemarec, C; Wu, J.-C.; Remaud, G.; Bazin, H.; Oivanen, M.; Lönnberg, H.; Chattopadhyaya, J. *Tetrahedron* **1988**, *44*, 1273 and references therein.
2. Golankiewicz, B.; Folkman, W. *Nucleic Acids Res.* **1983**, *11*, 5243.
3. Bazin, H.; Zhou, X.-X.; Glemarec, C; Chattopadhyaya, J. *Tetrahedron Lett.* **1987**, *28*, 3275.
4. Glemarec, C; Remaud, G.; Chattopadhyaya, J. *Nucleosides & Nucleotides* **1989**, *8*, 1513.
5. Glemarec, C; Remaud, G.; Chattopadhyaya, J. *Magn. Reson. Chem.* **1988**, *26*, 435.
6. Nordlund, T.M.; Rigler, R.; Glemarec, C; Wu, J.-C.; Bazin, H.; Remaud, G.; Chattopadhyaya, J. *Nucleosides & Nucleotides* **1988**, *7*, 805.
7. Frish, M.J.; Trucks, G.W.; Head-Gordon, M.; Gill, P.M.W.; Wong, M.W.; Foresman, J.B.; Johnson, B.G.; Schlegel, H.B.; Robb, M.A.; Replogle, E.S.; Gomperts, R.; Andres, J.L.; Raghavachari, K.; Binkley, J.S.; Gonzalez, C.; Martin, R.L.; Fox, D.J.; Defrees, D.J.; Baker, J.; Stewart, J.J.P.; Pople, J.A. Gaussian 92, Revision A, Gaussian, Inc., Pittsburgh PA, 1992.
8. Besler, B.H.; Merz, K.M. Jr.; Kollman, P.A. *J.Comp. Chem.* **1990**, *11*, 431.
9. Wong, M.W.; Frisch, M.J.; Wiberg, K.B. *J. Am. Chem. Soc.* **1991**, *113*, 4776.
10. Anders, E.; Katrizky, A.R.; Malhotra, N.; Stevens, J. *J. Org. Chem.* **1992**, *57*, 3698.
11. Katritzky, A.R.; Yannakopoulou, K.; Anders, E.; Stevens, J.; Szafran, M. *J. Org. Chem.* **1990**, *55*, 5683.
12. Hehre, W.J.; Radom, L.; Schleyer, P.v.R.; Pople, J.A. *Ab initio Molecular Orbital Theory*; Wiley: New York, **1986**.

(*Received in UK* 21 *August* 1995; *accepted* 9 *November* 1995)