BIOORGANIC CHEMISTRY **26,** 345–355 (1998)
ARTICLE NO. BH981113

Azaguanine: A Theoretical Study of Its Tautomerism and Protonation in the Gas Phase and Aqueous Solution

J. Guillermo Contreras*,[1] and Sandra T. Madariaga†

**Facultad de Ciencias Químicas, Universidad de Concepción, Casilla 160-C, Concepción, Chile;*
and †Facultad de Ciencias Universidad Católica de la Santísima
Concepción Casilla 297, Talcahuano, Chile

Received June 15, 1998

The tautomerism and protonation of 8-azaguanine (azaG) have been studied by means of ab initio methods, both in the gas phase and in aqueous solution. An elimination procedure to choose the most stable tautomeric forms, based on AM1 and HF/6–31G* energies, has been applied. Tautomers azaG(1,9), azaG(1,7) and azaG(9,15) have been selected and their energies calculated at MP2/6–311++G**//HF/6–31G* level. Self-consistent reaction field IPCM, based on polarizable continuum model (PCM), has been applied to study the solvent effects. The stability order in the gas phase is azaG(1,7) = azag(1,9) > azaG(9,15), whereas in solution the order becomes azaG(1,7) > azaG(1,9) > azaG(9,15), the latter being just 0.76 kcal/mol over azaG(1,7). The calculations of proton affinities allowed to unambiguously determine the preferred sites of protonation of these species.

INTRODUCTION

The replacement of the CH group in the imidazolic ring in natural purines by a nitrogen atom produces a family of compounds named 8-azapurines (*1*). The presence of the nitrogen atom at position 8 provides an additional basic center which in turn affects the basicities of the other nitrogens of the ring, produces changes in the purine ring structure, and induces distinct glycosyl conformations of the corresponding nucleosides due to restricted rotation around this bond (*2*). Furthermore, the lone pair on N8 is close enough to ribose to electrostatically interact with it (*3*, *4*).

Both, azaadenine (azaA) and azaguanine (azaG) analogs of the nucleic acids constituents are important since their remarkable biological activities. Thus azaA has been studied for antileukemic (*5*) and antibiotic (*6*) activities, whereas azaG is incorporated to several RNA fractions of bacillus cereus (*6*) by substituting guanine residues. AzaG does not randomly replace guanine residues in RNA. In fact, the highest ratio azaG/*G* is found in the terminal purine nucleoside. It has been established that the shorter nucleic acid chains contain more azaG than the larger ones (*6*). At the biological level, azaG exerts a strong action on human tumor cells by inhibiting the protein synthesis as a result of its incorporation in m-RNA (*7*). Speculating on the difference between guanine and azaguanine and the failure

---

[1] To whom correspondence should be addressed. E-mail: gcontrer@udec.cl.

0045-2068/98 $25.00

of the later to function normally one can think that substitution of the CH group of guanine by a nitrogen atom either (a) considerably alters the electronic structure of the purine ring, (b) greatly affects the hydrogen bonding structure involving the exocyclic substituents, (c) shows the importance of unusual tautomeric forms in azaG in aqueous solution. These effects are supported by studies on electronic and molecular structure determinations (*8–11*), proton transfer in base pairing (*12–17*), and tautomerism in nonnatural bases (*8*, *9*, *18*).

In the present work, we have studied the prototropic tautomerism of azaguanine both in the gas phase and in aqueous solution in an attempt to determine the importance of some unusual tautomeric forms that may create unfavorable hydrogen bonding structures and hence destabilizing the RNA structure. We have also studied the effect of the CH group substitution (by a nitrogen atom) on the basicities of the 6-oxo and 2-amino groups of azaG. Since protonation could be relevant generally in terms of molecular recognition, we have also studied the protonation of the most stable azaG tautomeric forms and determined the proton affinities for each basic site.

COMPUTATIONAL METHODS

The geometry of all tautomeric forms of azaguanine were optimized first at the AM1 level of theory. Since ab initio calculations are very time consuming, those tautomers with relative energies greater than 10 kcal/mol were eliminated. Next, we applied another elimination step based on the relative energies calculated at HF/6–31G^* level. In fact, all those species with $\Delta E < 5$ kcal/mol survived and were considered for further calculations. This elimination criterion leads to consideration of just two oxo-amino and one hydroxo-amino species. The AM1 calculations were carried out using AMSOL 5.0 (*19*), whereas the ab initio calculations were performed using GAUSSIAN 94 codes (*20*).

Ab initio geometry optimizations for all three tautomeric forms of azaguanine were carried out at HF/6–31G^* level. The initial geometries were those calculated for azapurine (*8*) and azaadenine (*9*). Frequency calculations and IR intensities were predicted at the equilibrium geometries yielding all real frequencies; hence all calculated structures are local minima. The energy calculations were carried out using different basis sets to determine the effect of including polarization and diffuse functions. In all cases electronic correlation was taken into account at the second-order Moller-Plesset theory in the frozen core approximation. The calculated energy values at MP2/6–311++G^{**} were corrected for zero point vibrational energies (unscaled). In fact, scaling ZPE, as usual, by 0.9 to account for the overestimation of HF vibrational frequencies produced an almost constant value that makes no difference on the conclusions of this study. Tautomerization enthalpies were obtained by adding ZPE and thermal corrections to the relative energies calculated at the MP2/6–311++G^{**} level. The corresponding ΔG values were calculated, as usual, from $\Delta G = \Delta H - T\Delta S$.

The solute–solvent effect was taken into account by applying Tomasi's polarizable continuum model (PCM) (*21*) modified by Wiberg *et al.* (*22*, *23*). This method (IPCM) calculates the electric field analytically and the cavity is defined upon an isosurface of the total electron density calculated at the level being used. Thus, the cavity is uniquely derived from the electronic environment and just the isosurface level, i.e, the charge density (0.0004–0.001 e/B^3), needs to be specified. The solvent effect in IPCM is derived from the surface potentials and the dielectric continuum interactions. This is equivalent to going to infinite order in the electric moments. In applying IPCM method, the gas phase molecular geometries were used since structural parameters change very little in going from the gas phase to solution and hence no large effects on solvation energies, in relative terms, can be expected (*23*, *24*). The free energies in solution (G^0_{soln}) were calculated from $G^0_{\text{soln}} = \Delta G_{\text{gas}} + \Delta G^0_{\text{s}}$.

RESULTS AND DISCUSSION

Tautomerism of Neutral Azaguanine

Theoretically azaguanine is capable of existing in 15 different tautomeric forms arising from the amino ⇔ imino, hydroxo ⇔ oxo equilibria and the tropism of the protons within the two fused heterocycles; i.e., pyrimidine and triazolo rings. Inspection of semiempirical AM1 and HF/6–31*G** calculations allowed us to exclude all imino forms as well as all forms involving the two hydrogen atoms located at the same ring (annular tautomerism). Accordingly, just three tautomeric forms were considered for further calculations, namely, two oxo-amino and one hydroxo-amino species. The optimized structural parameters for these species are given in Table 1, using the atom numbering shown in Fig. 1. The experimental values (*10*) compare well with the calculated ones. The rms deviations rank from 0.07 to 0.26 Å for the bond distances, whereas for the bond angles rms deviations vary from 2 to 3^0. In the present work, we have used the notation azaG(*i*, *j*) where *i* and *j* stand for the atom number to which the tropic hydrogens are attached. In fact, we have considered the tautomers azaG(1,9), azaG(1,7), and azaG(9,15). From Table 2, which lists the energetics of these species in the gas phase, one can infer that no matter the level of calculation used, azaG(1,9) is the most stable species. For all three forms the inclusion of polarization functions for the hydrogen atoms, as well as electron correlation, is of great importance. In fact, at MP2/6–31*G*** the convergence of the results increases. However, it is also clear that the inclusion of electronic correlation produces a strong stabilization of the oxo-amino azaG(1,7) form, whereas the *cis*-hydroxo-amino azaG(9,15) species is destabilized. Polarization functions on hydrogens destabilizes the oxo-amino form azaG(1,9), whereas the hydroxo-amino species azaG(9,15) is strongly stabilized. This same effect is observed when using a large basis set like 6–311G** . Likewise in azaadenine (*9*) the use of a sufficiently extended and flexible basis set that includes diffuse functions, like 6–311++*G***, stabilizes all species and is expected to properly represent the molecular properties of systems like azaguanine (*25*). In fact, at MP2/6–311++*G*** level, azaG(1,7) and azaG(9,15) are just at 0.22 and 1.38 kcal/mol, respectively, over the most stable azaG(1,9). Therefore, the stability order in the gas phase is: azaG(1,9) $\approx$ azaG(1,7) $>$ azaG(9,15), implying that most likely azaG(1,9) and azaG(1,7) are in similar concentration, the azaG(9,15) population being smaller.

TABLE 1
Gas Phase Optimized Parameters (HF/6–31G*) for azaG(1,9), azaG(1,7), and azaG(9.15) Tautomeric Forms[a]

Parameter	azaG(1,9)	azaG(1,7)	azaG(9,15)	Experimental[b]
r(N1–C2)	1.3611	1.3766	1.3510	1.379
r(C2–N3)	1.2925	1.2776	1.3192	1.334
r(N3–C4)	1.3520	1.3642	1.3268	1.351
r(C4–C5)	1.3708	1.3618	1.3798	1.383
r(C5–C6)	1.4383	1.4375	1.4004	1.430
r(C5–N7)	1.3620	1.3502	1.3701	1.361
r(N7–N8)	1.2524	1.3057	1.2476	1.303
r(N8–N9)	1.3491	1.2736	1.3538	1.359
r(Ni–H10)[c]	0.9943	0.9949	0.9939	0.900
r(N11–C2)	1.3536	1.3671	1.3455	1.321
r(N11–H12)	0.9959	0.9975	0.9933	0.850
r(N11–H13)	0.9949	0.9969	0.9928	0.900
r(H14–N1)	0.9980	0.9981		0.860
r(H14–O15)			0.9521	
r(C6–O15)	1.1901	1.1954	1.3119	1.224
<(N1–C2–N3)	124.18	124.49	127.34	123.2
<(C2–N3–C4)	112.37	113.87	111.90	112.4
<(N3–C4–C5)	128.74	125.25	126.97	128.0
<(C4–C5–C6)	119.30	122.0	115.18	119.6
<(C4–C5–N7)	108.58	104.08	108.88	109.3
<(C5–N7–N8)	108.72	109.86	108.26	108.1
<(N7–N8–N9)	108.68	109.92	109.39	108.3
<(C4–Ni–H10)[c]	129.19	129.87	129.51	128.2
<(N3–C2–N11)	119.49	120.66	117.49	119.8
<(C2–N11–H12)	114.97	112.72	118.38	119.3
<(C2–N11–H13)	119.55	117.29	118.70	123.4
<(C6–N1–H14)	113.94	114.84		116.4
<(C6–N1–O15)	119.64	122.18	119.08	121.6
<(O15–H14–C6)			108.56	

[a] Bond distances and angles in Å and degree, respectively.
[b] From Ref. *10*.
[c] $i = 7$ or 9.

Azaguanine can be compared with azapurine (azaP) and azaadenine (azaA) since in the former the pyrimidine proton is always attached to N1. The stability orders in both azaP and azaA is (9) > (8) > (7). This implies that at least, in the gas phase, the amino group (at C6) in azaA exerts no influence on the stabilities of the three tautomeric forms or that a counterbalance between the lone pair repulsion (which favors azaA(8)) and aromaticity (which favors azaA(9)) effects takes place (*26*, *27*). In azaG the situation is different as the lone pair repulsion effect at the triazolo ring is just about the same in all species and hence azaG(1,8) is greatly destabilized and was correctly ruled out on the basis of its relative energy calculated at the AM1 level. The above-described results are in full agreement with those

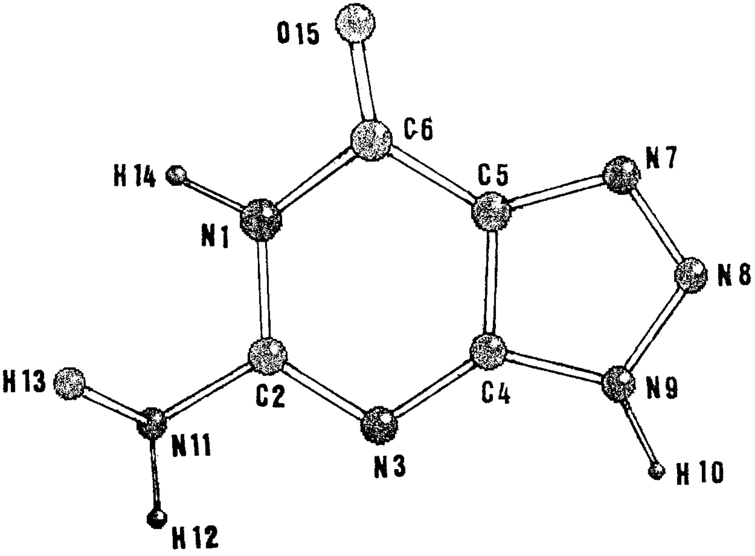

FIG. 1. Atom numbering of azaguanine tautomers. AzaG(1,7) is shown.

TABLE 2
Energetics[a,b] and Dipole Moments for the Tautomeric Forms of Azaguanine (azaG) in the Gas Phase

	azaG(1,9)	azaG(1,7)	azaG(9,15)
E(HF/6–31G*)	−555.34452	−555.33678	−555.33917
E(MP2/6–31G*)	−556.97539	−556.97486	−556.96951
E(MP2/6–31G**)	−557.01316	−557.01256	−557.00960
E(MP2/6–311G**)	−557.22097	−557.22031	−557.21780
E(MP2/6–311++G**)	−557.24685	−557.24684	−557.24436
ZPE	71.63	71.83	71.41
$H - H_0$	4.73	4.69	4.75
S	86.26	86.05	86.29
μ^c	8.28	4.78	4.62
	Relative values		
ΔE^c	0.00	0.00	1.56
$\Delta(\text{ZPE})^d$	0.00	0.20	−0.22
$\Delta E + \Delta(\text{ZPE})$	0.00	0.20	1.34
$\Delta(H - H_0)$	0.00	−0.04	0.03
ΔH	0.00	0.16	1.37
$T\Delta S$	0.00	−0.06	0.01
ΔG	0.00	0.22	1.38

[a] Calculations based on HF/6–31G* geometries.

[b] E in $^{\text{Hartrees}}$; ZPE, $H - H_0$, ΔH, ΔE, $T\Delta S$, and ΔG in kcal/mol.

[c] Calculated at MP2/6–311++G** level.

[d] Unscaled values.