JOURNAL OF PHYSICAL ORGANIC CHEMISTRY, VOL. 11, 392–396 (1998)

Theoretical study of the tautomerism of 8-azaadenine

J. Guillermo Contreras,* Sandra T. Madariaga and Joel B. Alderete

Facultad de Ciencias Quimicas, Universidad de Concepcion, Casilla 3-C, Concepcion, Chile

Received 20 March 1997; revised 15 July 1997; accepted 15 September 1997

ABSTRACT: The prototropic tautomerism of 8-azaadenine (azaade) was studied theoretically by means of *ab initio* methods, in both the gas phase and aqueous solution. A number of tautomeric forms were not included in the calculations after applying a stepwise elimination procedure based on both AM1 and HF/6–31G* energy values. The tautomers 9*H*-azaade, 8*H*-azaade and 7*H*-azaade survived to this elimination and their optimized geometries and energies were calculated at the MP2/6–31*//HF/6–31G* level. To include the solvent effects, two self-consistent reaction field method were used: (1) Onsager's SCRF with multipolar expansion up to the hexadecapolar term and (2) the isodensity polarizable continuum method (IPCM). Both methods produce similar results, although the latter represents better the situation in aqueous solution. The stability order in solution, 8*H*- > 9*H*- > 7H-azaade, differs slightly from that found in the gas phase, implying that in general the electrostatic effects in solution are important, but the intrinsic stability of these species in the gas phase overcomes the solvent effect. © 1998 John Wiley & Sons, Ltd.

KEYWORDS: 8-azaadenine; tautomerism; theoretical study

INTRODUCTION

8-Azapurines [azapurine, azaadenine (azaade) and azaguanine] differ from natural purines in that the CH group of the imidazolic ring has been replaced by a nitrogen atom. This replacement produces distinct glycosyl conformations of the corresponding nucleosides as a result of changes in the possible rotations around this bond.[1] The presence of this nitrogen atom induces interesting biological activities, e.g. 8-azapurines possess antipurine, antifungal, antiviral and anticancer properties.[2,3] In this sense, azapurines work either by replacing adenine or guanine in nucleosides or interfering with some enzymatic process. The insertion of azapurines in RNA implies that the appropriate hydrogen bonding between the azapurine derivative and the pyrimidine base must be formed.

It is well known that hydrogen bond complex formation plays an important role in the stabilization of the double helix structure of nucleic acids.[4] Accordingly, the adenine (A)–thymine (T) pair is stabilized by two hydrogen bonds, whereas the stabilization of the guanine (G)–cytosine (C) pair is achieved by three hydrogen bonds. The recognition of a pyrimidine base by its complementary purine is determined by the appropriate hydrogen bond formation, which in turn preserves the genetic code. There has been some theoretical speculation about the possibility of expanding the genetic code by inserting non-natural bases in the nucleic acids. In fact, Piccirilli *et al.*[5] prepared a set of new molecules that form mutually compatible hydrogen bonding patterns with each other and yield unstable pairs with the natural bases.[6]

In the present work, we studied the prototropic tautomerism of 8-azaadenine both in the gas phase and in aqueous solution in an attempt to determine the major tautomeric forms present in solution for further studies on pair formation with both natural (thymine, uracil) and non-natural (2,6-diaminopyrimidine) bases.

COMPUTATIONAL METHODS

The geometry of the various tautomers of 8-azaadenine were optimized first at the AM1 level of theory[7] and their relative energies determined. Since *ab initio* calculations on molecular systems such as azaadenine are exceedingly time consuming, all those tautomeric forms with gas-phase relative energies over 15 kcal mol^{-1} were not considered further for *ab initio* calculations. In fact, AM1 calculations revealed that the tautomeric forms 9*H*, 8*H*, 7*H*- and 3*H*-azaade possess relative energies in the range

**Correspondence to:* J. Guillermo Contreras, Facultad de Ciencias Quimicas, Universidad de Concepcion, Casilla 3-C, Concepcion, Chile. Email: gcontrer@buho.dpi.udec.cl
Contract/grant sponsor: Universidad de Concepcion; *contract/grant number:* 96.021.003-1; *contract/grant number:* 97.21.006-1.2; *contract/grant number:* 97.023.015-11D.

CCC 0894–3230/98/060392–05 $17.50

0–15 kcal mol^{-1}. In this sense, azaadenine behaves as azapurine.[8] Since it has been found that the AM1 method does not provide a good description of atomic charge distributions and energies in this kind of molecule,[9–11] we applied another elimination criterion, i.e. all species possessing relative energies less than 7 kcal mol^{-1} at the HF/6–31G* //HF/6–31G* level will survive. In fact, the *ab initio* calculations at this level revealed that the 9*H*-, 8*H*- and 7*H*- tautomers would be the only ones to be considered for further calculations. The present case is in agreement with our previous results.[8] The *ab initio* calculations were carried out using Gaussian 94 codes,[12] whereas the AM1 calculations were performed using the Mopac 5.0 program package.[13]

Ab initio geometry optimization for all amino forms was performed at the HF/6–31G* level. Initial geometries were taken from the corresponding azapurine and adenine tautomeric forms.[8,14] Frequency calculations and IR intensities predicted at the equilibrium geometries produce all real frequencies and hence all structures are local minima. For the selected tautomers of azaadenine, energies were calculated using various basis sets, including polarization effects on both heavy and hydrogen atoms and diffuse functions and inclusion of electron correlation at the MP2 level in the frozen core approximation. For comparison purposes, we also applied the BLYP correlated function. The calculated energies were corrected for zero-point vibrational energies (unscaled). Scaling the ZPE correction by 0.9, as usual, to account for the overestimation of vibrational frequencies at the HF level yielded an almost constant value for all species studied here and hence it had no effect on the conclusions of this study. To obtain the free energy changes of tautomerization, enthalpies were calculated by adding ZPE and the thermal corrections ($H - H_o$) to the relative energies calculated at the MP2/6–311++G** level. ΔG values were obtained, as usual, from $\Delta G = \Delta H - T\Delta S$. Since in azapurine all three tautomeric forms originated at the triazole ring were found to have similar free energies in solution ($G°_{soln}$),[8] we also included the 7*H*-azaade tautomer. In addition, this species possesses the largest dipole moment and hence it may be greatly stabilized in aqueous solution.

The solute–solvent (water) effect was taken into account by using two self-consistent reaction field models (SCRF). In SCRF models, the solute charge distribution immersed in an unstructured medium of dielectric constant ε will induce an electric field in the solvent, which in turn interacts, stabilizing the solute. The simplest Onsager SCRF model considers a spherical cavity and the solute charge distribution is truncated at the dipole term ($l = 1$). The cavity radius is determined from the electronic wave function following Wong *et al.*[15] We used an extension of the Onsager method by including other terms in the charge distribution expansion and with the cavity still being spherical. In fact, Foresman *et al.*[16] have shown that an ellipsoidal cavity does very little to improve the basic method. The solute charge distribution is described by a single center multipole expansion up to the hexadecapole term ($l = 4$), as implemented in Gaussian 94 following Rinaldi and co-workers' approach.[17] In order to compare the results obtained by the above method, we also applied Tomasi and co-workers' polarizable continuum model (PCM)[18] modified by Wiberg and co-workers'.[16,19] This method (IPCM) calculates the electric field analytically instead of numerically and the cavity is defined upon an isosurface of the total electron density calculated at the level of theory being used. The cavity is derived uniquely from the electronic environment and one needs to specify just the isosurface level, i.e. charge density to be used to define the surface and that ranks between 0.0004 and 0.001 e/B^3. The solvent effect in IPCM is derived from the interaction of the surface potentials with the dielectric continuum. This is equivalent to going to infinite order in the electric moments.

In both types of SCRF methods, the gas-phase molecular geometries optimized at the HF/6–31G* level were used. In fact, it is well known that the structure parameters change very little on going from the gas phase to solution and hence one can expect not to produce large effects on solvation energies.[11,19a,b,20] The solvation free energies were taken as the difference between the energies in solution and in the gas phase. The free energies in solution [$G°_{soln}$] were calculated, as usual, from the relationship $G°_{soln} = \Delta G°_{gas} + \Delta G°_{s}$.

RESULTS AND DISCUSSION

Figure 1 shows the atom numbering used in Table 1. Table 1 lists the gas-phase most relevant geometric parameters for the tautomers calculated here. The derived structural parameters compare well with those reported for 7-methylazaadenine[21] with rms deviations of *ca* 0.04 Å for the bond lengths and 1.2–5.2° for bond angles. The average N7— N8 and N8— N9 distances for all three tautomeric forms are 1.287 and 1.304 Å , i.e. surprisingly close to the corresponding bond distances in 8*H*-azaade. This same trend is detected in the AM1 bond distances and the calculated bond orders reveal that in 8*H*-azaade there must exist a slight electron delocalization on the N7 — N8 — N9 group. The gas-phase *ab initio* calculations are presented in Table 2. The energy values are those at the MP2(fc)/6–311++G**//HF/6–31G* level, all relative values are referred to the 9*H*-azaade values, since this species appears to be the most stable. The ΔG values show that all species tend to produce the more stable 9*H*-azaade. 8*H*-Azaade and 7*H*-azaade lie at 0.78 and 6.67 kcal mol^{-1}, respectively. The free energy changes for the tautomerization reactions of 3*H*- and 1H-azaade at the HF/6–31G* level to produce 9*H*-azaade are −12.60 and −18.46 kcal mol^{-1}, respectively, supporting their earlier exclusion in the *ab initio* calculations. The

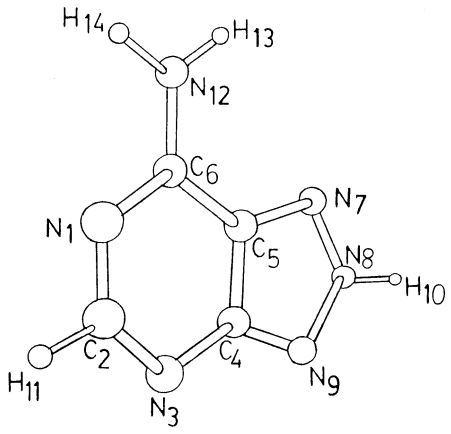

Figure 1. Atom numbering of azaadenine tautomers. 8*H*-Azaade is shown

inclusion of 7*H*-azaade in the solution calculation can be justified by its high dipole moment, which could lead to strong stabilization in aqueous solution. From Table 2 it can also be inferred that when using second-order Möller–Plesset electron correlation, the relative energies change very little on going from the 6–31G* basis set to the more flexible 6–311++G** basis set. In fact, for 8*H*-azaade, ΔG using the latter basis set is just 0.28 kcal mol^{-1} smaller than the value obtained with the 6–31G* basis set, whereas for 7*H*-azaade ΔG decreases by 0.73 kcal mol.$^{-1}$ Therefore, despite the use of the highly flexible 6–311++G** basis set, the 9*H*- and 8*H*-azaade tautomeric forms remain being the only species present in the gas phase. This effect has also been observed for 1,2,3-triazole.[22,23] The stabilization of 9*H*- and 8*H*-azaade can be rationalized in term of the balance between lone pair repulsion (which favors the 8*H*-tautomer) and the aromaticity (which favors the 9*H*-tautomer). Since the separation between the two most stable tautomers is just 0.78 kcal mol^{-1}, the unfavorable lone pair repulsion and resonance effects are approximately counterbalanced, allowing both species to exist in the gas phase. This statement should be treated with caution, owing to the disagreement between the MP2 and DFT results. In fact, the energy separation between the 8*H*-and 9*H*-tautomers at the BLYP/6–311++G** level is *ca* 3.32 kcal mol.$^{-1}$ This means that the results should also be treated with caution as no satisfactory explanation for the different DFT and MP2 results could be found.

The solvent effect on the prototropic tautomerism is given in Table 3, where the solvated energies of the *iH*-azaade (i = 7,8,9) tautomeric species in water (ε = 78.5) are listed. From this table, it can be inferred that the stability order of the tautomers in aqueous solution depends on the method used to simulate the solvent. Thus, using the SCRF (l = 4) method the stability order is 9*H*-> 8*H*-> 7*H*-azaade, although the energy difference between 9*H*-and 8*H*-azaade is just 0.31 kcal mol^{-1},

Table 1. Gas-phase optimized geometries (HF/6–31G*) for the *iH*-azaade (i = 7, 8, 9) tautomeric forms[a]

Parameter	7*H*-Azaade	8*H*-Azaade	9*H*-Azaade	Exp.[b]
r(N1—C2)	1.346	1.362	1.334	1.349
r(C2—N3)	1.297	1.288	1.311	1.319
r(N3—C4)	1.337	1.356	1.332	1.354
r(C4—C5)	1.371	1.397	1.373	1.384
r(C5—C6)	1.408	1.431	1.408	1.423
r(C5—N7)	1.359	1.322	1.368	1.367
r(N7—N8)	1.320	1.289	1.252	1.346
r(N8—N9)	1.259	1.308	1.344	1.306
r(N*i*—H10)[c]	0.993	0.995	0.994	
r(C2—H11)	1.074	1.075	1.075	0.900
r(C6—N12)	1.363	1.337	1.335	1.325
r(N12—H13)	0.995	0.993	0.994	0.930
r(N12—H14)	0.997	0.993	0.994	0.890
<N1C2N3	128.2	129.7	129.2	129.1
<C2N3C4	112.9	112.5	111.1	117.1
<N3C4C5	124.1	123.6	126.4	124.9
<C4C5C6	118.1	117.7	116.7	118.7
<C4C5N7	103.8	109.0	108.9	104.2
<C5N7N8	109.5	102.3	108.2	109.4
<N7N8N9	110.6	118.2	109.4	109.8
<CNiH10	131.5		129.5	
<N3C2H11	117.1	116.7	116.0	114.0
<C5C6N12	123.2	121.6	122.8	125.5
<C6N12H13	118.1	121.3	121.0	124.0
<C6N12H14	114.3	119.0	119.1	118.0

[a] Bond distances in Å and angles in degrees.
[b] Experimental data from Ref. 18.
[c] i = 7, 8, 9.

Table 2. Calculated energies[a,b] and dipole moments (μ) for the three tautomers of azaadenine in the gas phase

Parameter	9*H*-Azaade	8*H*-Azaade	7*H*-Azaade
E(HF/6–31G*)	−480.46569	−480.45437	−480.44701
E(MP2/6–31G*)	−481.92491	−481.92331	−481.97337
E(MP2/6–31G**)	−481.96102	−481.95933	−481.94956
E(MP2/6–311G**)	−482.12772	−482.12632	−482.11676
E(MP2/6–311+G**)	−482.14946	−482.14825	−482.13905
E(MP2/6–311++G**)	−482.14992	−482.14877	−482.13955
E(BLYP/6–311++G**)	−483.32611	−483.32082	−483.31230
ZPE	68.08	68.46	68.04
$H - H_o$	4.25	4.30	4.14
S	82.62	83.85	81.56
μ[c]	1.05	4.72	6.67
Relative values:			
Δ*E*(HF/6–31G*)	0.00	7.10	11.72
Δ*E*(MP2/6–31G*)	0.00	1.00	7.24
Δ*E*(MP2/6–31G**)	0.00	1.06	7.19
Δ*E*(MP2/6–311G**)	0.00	0.88	6.88
Δ*E*(MP2/6–311+G**)	0.00	0.76	6.53
Δ*E*(MP2/6–311++G**)	0.00	0.72	6.51
Δ*E*(BLYP/6–311++G**)	0.00	3.32	8.67
Δ(ZPE)[d]	0.00	0.38	−0.04
Δ*E* + Δ(ZPE)*	0.00	1.10	6.47
$\Delta(H - H_o)$	0.00	0.05	−0.12
Δ*H*	0.00	1.15	6.35
*T*Δ*S*	0.00	0.37	−0.32
Δ*G*[e]	0.00	0.78	6.67

[a] Based on HF/6–31G* geometries.
[b] *E* in hartree; ZPE, $H - H_0$, Δ*H*, Δ*E*, *T*Δ*S* and Δ*G* in kcal mol^{-1}; *S* in cal mol^{-1} K^{-1}.
[c] MP2/6–31G* values (D).
[d] Unscaled values.
[e] Based on Δ*E* values calculated at the MP2/6–311++G** level

whereas IPCM stabilizes 8*H*-azaade by 0.36 kcal mol^{-1} over 9*H*-azaade, in good agreement with the larger dipole moment of the former. In both methods 7*H*-azaade appears greatly destabilized despite having the largest dipole moment. This effect is more likely due to the intrinsic instability of this species in the gas phase. We

Table 3. Energies[a] of solvated azaadenine tautomers (ε = 78.5)

Tautomer	SCRF(l = 4)[b]	IPCM
9*H*-Azaade	−480.47955	−480.49353
8*H*-Azaade	−480.46941	−480.48446
7*H*-Azaade	−480.46428	−489.47753
Solvation free energies ($\Delta G°_s$)[d]		
9*H*-Azaade	−8.70	−17.46
8*H*-Azaade	−9.44	−18.88
7*H*-Azaade	−10.57	−19.15
Free energies in solution ($G°_{soln}$)[e]		
9*H*-Azaade	−8.70	−17.46
8*H*-Azaade	−8.38	−17.82
7*H*-Azaade	−3.17	−11.75
Relative values		
9*H*-Azaade	0.00	0.00
8*H*-Azaade	0.32	−0.36
7*H*-Azaade	5.53	6.07

[a] Energies in hartree.
[b] SCRF (l = 4) a_o = 4.08 Å .
[c] Relative energies in kcal mol^{-1}.
[d] $\Delta G°_s = \Delta E$(solution - gas).
[e] $G°_{soln} = \Delta G_{gas} + \Delta G°_s$.

believe that IPCM reproduces the solvent effect better than SCRF ($l = 4$), although in both methods just the electrostatic interactions are taken into account; in IPCM the solute charge distribution expansion is implicitly taken to infinite order, whereas in the SCRF method the series is truncated at the hexadecapole term. It is worth nothing that the effective cavity radius in SCRF ($l = 4$) is *ca* 33% smaller than that used in the IPCM method. The smaller cavity radius would produce a stronger reaction field, although this is not observed in the ΔG°_{soln} values, which are much smaller than the values calculated by the IPCM method. Anyway, both methods predict that 8*H*- and 9*H*-azaade must be present in similar concentrations in aqueous solution and hence both must be taken into account for further work on pairing with pyrimidine bases. The SCRF ($l = 4$) stability order 9*H*-> 8*H*-> 7*H*-azaade differs from that found for azapurine, where 7*H*-azapurine is more stable than 8*H*-azapurine(8).[8] It is worth noting that for the tautomeric species of azapurine, the multipolar expansion was truncated at the dipole ($l = 1$) term. These results indicate that truncation of multipolar expansion at an arbitrary level neglects important contributions to the electrostatic effect. 7*H*-Azaade is slightly the better solvated species, although the solvation free energy of 8*H*-and 9*H*-azaade are *ca* 1 and 2 kcal mol^{-1} smaller, respectively. The lowest stability of 7*H*-azaade comes from its intrinsic instability in the gas phase.

CONCLUSIONS

9*H*-and 8*H*-azaade tautomers are the most important species and probably are the only ones in aqueous solution. The inclusion of a solvent stabilizes different species, depending on the method used to simulate the solvent. Since the solvation energies are similar for the two most stable tautomers (less than 2 kcal mol^{-1}), the stability in solution is governed by the intrinsic stability in the gas phase.

Both the SCRF at $l = 4$ and the IPCM solvation methods, produce slightly different results, although the goodness of the former may be the result of fortuitous cancellation of different errors, as pointed out by Foresman *et al.*[16] In fact, the smaller solute cavity calculated for SCRF ($l = 4$) should produce a larger reaction field than the IPCM method. Since this is not the case, it is probable that this effect is cancelled by the truncation of the multipolar expansion.

For future work on the calculation of the energetics of azaadenine pairing with some pyrimidine base analogs, the 8*H*-and 9*H*-azaade tautomers will be the only ones to be considered.

Acknowledgement

This work was supported by operating grants from the Universidad de Concepcion (Nos 96.021.003-1, 97.21.006-1.2 and 97.023.015-11D).

REFERENCES

1. D. C. Ward and E. Reich. *Annu. Rep. Med. Chem.* 272 (1969).
2. G. Brule, S. J. Eckhardt, T. C. Hall and A. Winkler. *Drug Therapy of Cancer*. World Health Organization, Geneva (1973).
3. F. M. Schabel, Jr. *Chemotherapy* **13,** 321 (1968).
4. J. D. Watson and F. H. C. Crick. *Nature (London)* **171,** 737 (1953).
5. J. A. Piccirilli, T. Krauch, S. E. Moroney and S. A. Benner. *Nature (London)* **343,** 33 (1990).
6. A. R. Leach and P. A. Kollman. *J. Am. Chem. Soc.* **114,** 3675 (1992).
7. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart. *J. Am. Chem. Soc.* **107,** 3902 (1985).
8. J. G. Contreras, S. T. Madariaga and J. B. Alderete. *J. Mol. Struct. (THEOCHEM)* **365,** 63 (1996).
9. J. G. Contreras and J. B. Alderete. *J. Phys. Org. Chem.* **8,** 395 (1995).
10. S. Woodcock, D. V. S. Green, M. A. Vincent, I. H. Hilliers, M. F. Guess and P. J. Sherwood. *J. Chem. Soc., Perkin Trans.* **2,** 2151 (1992).
11. J. G. Contreras and J. B. Alderete. *J. Mol. Struct. (THEOCHEM)* **283,** 283 (1993); J. G. Contreras, S. T. Madariaga and J. B. Alderete. *Bol. Soc. Chil. Quim.* **42,** 17 (1997).
12. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Peterson, J. A. Montgomery, K. Raghvachari, M. A. Al-laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, G. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Frees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez and J. A. Pople. *Gaussian 94, Revisions B.2 and B.3.* Gaussian, Pittsburgh, PA (1995).
13. J. J. P. Stewart. *QCPE Bull.* **7**(4), (1987).
14. E. L. Stewart, C. K. Folet, N. L. Allinger and J. P. Bowen. *J. Am. Chem. Soc.* **116,** 7282 (1994).
15. M. W. Wong, K. B. Wiberg and M. J. Frisch. *J. Am. Chem. Soc.* **114,** 1645 (1992).
16. J. B. Foresman, T. A. Keith, K. B. Wiberg, J. Snoolan and M. J. Frisch. *J. Phys. Chem.* in press.
17. (a) D. Rinaldi, M. F. Ruiz-Lopez and J. L. Rivail. *J. Chem. Phys.* **78,** 834 (1983). (b) D. Rinaldi, J. L. Rivail and N. Rguini. *J. Comput. Chem.* **13,** 675 (1992).
18. (a) S. Miertius, B. Scrocco and J. Tomasi. *Chem. Phys.* **55,** 117 (1981); (b) J. Tomasi, R. Bonaccorsi, R. Cammi and F. J. O. del Valle. *J. Mol. Struct. (THEOCHEM)* **234,** 401 (1991); (c) J. Tomasi and M. Persico. *Chem. Rev.* **94,** 2027 (1994).
19. (a) K. B. Wiberg, P. R. Rablen, D. J. Rush and T. A. Keith. *J. Am. Chem. Soc.* **117,** 4261 (1995); (b) K. B. Wiberg, T. A. Keith, M. J. Frisch and M. Murcko. *J. Phys. Chem.* **99,** 9072 (1995). (c) K. B. Wiberg, H. Castejon and T. A. Keith. *J. Comput. Chem.* **17,** 185 (1996).
20. P. E. Young and I. H. Hilliers. *Chem. Phys. Lett.* **215,** 405 (1993).
21. A. L. Shoemaker and D. J. Hodgson. *J. Am. Chem. Soc.* **99,** 4119 (1977).
22. J. R. Cox, S. Woodcock, I. H. Hilliers and M. A. Vincent. *J. Phys. Chem.* **94,** 5499 (1990).
23. F. Thomas, J. Catalan, P. Perez and J. Elguero. *J. Org. Chem.* **59,** 2799 (1994).

