A Theoretical Study, using *ab initio* Methods, of Tautomerism in 3-Amino-1,2,4-Triazole in the Gas Phase and in Aqueous Solution

Oswald G. Parchment,[a] Ian H. Hillier,[a] Darren V. S. Green,†[a] Neil A. Burton,[b] John O. Morley[c] and Henry F. Schaefer III[d]

[a] *Chemistry Department, University of Manchester, Manchester, M13 9PL, UK*
[b] *Manchester Computing Centre, University of Manchester, Manchester, M13 9PL, UK*
[c] *ICI Specialities, PO Box 42, Hexagon House, Blackley, Manchester, M9 3DA, UK*
[d] *Centre for Computational Quantum Chemistry, University of Georgia, Athens, Georgia 30602, USA*

Ab initio molecular orbital methods have been used to predict the relative energies of the three tautomers of 3-amino-1,2,4-triazole, in the gas phase and in aqueous solution. At the highest level of theory employed [6-31G**(CCSD)//6-31G**(HF)] the 1*H* and 2*H* tautomers are essentially isoenergetic in the gas phase, with the 4*H* tautomer 7 kcal mol^{-1} to higher energy. Both the self-consistent reaction field and polarisable continuum models predict differential stabilisation of the 4*H* tautomer upon hydration. However, neither model is in accord with the relative energies of the hydrated tautomers suggested by ^{15}N NMR spectroscopic data.

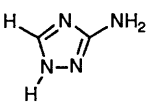
Quantum mechanical calculations of varying degrees of sophistication and accuracy have been widely employed to study the molecular structure and energetics of heterocyclic tautomers in the gas phase.[1] However, for the majority of tautomeric equilibria little gas phase data exist, and most measured equilibrium constants refer to the condensed phase. In addition, the interest in such equilibria is largely directed towards intermolecular interactions, particularly of biological relevance, involving the possible tautomers.[2] In these molecular modelling investigations, the effect of the environment on the tautomeric equilibria is naturally important.

The effect of the solvent or other molecular environment can be modelled *via* molecular dynamics or Monte Carlo methods, where the molecular structure of the environment is considered explicitly. Alternatively, a model can be employed in which the solvent is treated as a dielectric continuum. A range of continuum models have been implemented within both semi-empirical and *ab initio* molecular orbital (MO) methods. A number of such methods incorporate the Onsager reaction field model into MO calculations following Tapia and Goscinski.[3] Using a dipole approximation for the solute charge distribution, this method has been incorporated into the semi-empirical MO package MOPAC[4] and has subsequently been used to study the effect of solvent upon a variety of tautomeric equilibria. We have incorporated this method into the *ab initio* code GAMESS[5] and used it to study the solvation of hydroxypyridine[6] and cytosine tautomers.[7] This self-consistent reaction field (SCRF) method models the interaction of the solute dipole in a spherical cavity, with the polarisable solvent characterised by the relative permittivity (ε). A more realistic shape for the solvent cavity and a more accurate representation of the solute charge distribution are incorporated in a Polarisable Continuum Model (PCM) developed by Tomasi and coworkers.[8] This method involves the generation of a cavity from spheres centred at each atom in the molecule and the calculation of virtual charges on the cavity surface to represent the polarisation of the solvent. The magnitude of these charges is proportional to the derivative of the solute electrostatic potential at each point directly calculated from the electronic wavefunction. The PCM treatment has also been implemented by us in the program GAMESS.

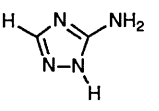
† *Present address:* Glaxo Group Research, Greenford Road, Greenford, Middlesex UB6 OHE, UK.

The only parameters in the SCRF and PCM models, in addition to those present in the MO treatments, are those related to the cavity geometry. The radius of the spherical cavity in the SCRF model can be chosen either from simple geometric considerations or from the molar volume. For the PCM treatment, the individual sphere radii must naturally depend upon atom type, but should also vary with formal atomic charge, and are also expected to be basis set dependent. Appropriate parameters have been developed[9] to allow the atomic radii to be calculated in terms of Mulliken charges and basis set.

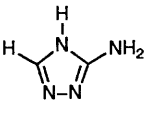
In order to predict the energetics of tautomeric equilibria in solution, the most accurate calculations of the energy differences of the gas phase tautomers must be combined with accurate estimates of the solvation free energies. In this paper we apply such a strategy to the study of the tautomeric equilibria of 3-amino-1,2,4-triazole, which can be written in three possible tautomeric forms, 1*H*(**I**), 2*H*(**II**) and 4*H*(**III**).



1*H* I



2*H* II



4*H* III

Gas Phase Calculations.—Unconstrained geometry optimisation of the three tautomers was carried out at the Hartree Fock (HF) level using both a 3-21G and 6-31G** basis. The structures obtained were characterised as minima by calculation of the second derivatives of the energy. The influence of electron correlation on the relative energies of the tautomers was estimated by two approaches at the 6-31G**//6-31G**(HF) level.

Firstly, second-order Møller–Plesset perturbation theory (MP2) was employed, considering all filled and virtual orbitals (including core orbitals). These calculations were carried out using the programs GAMESS and CADPAC.[10] To include higher orders of electron excitation in the calculation of correlation energy in a size consistent manner, the coupled cluster method[11] was employed using the program PSI.[12] We have used the coupled cluster, singles and doubles (CCSD) method,[13] which, as an infinite order method, provides a more balanced inclusion of the higher order correlation effects, neglected by MP2. For computational feasibility the six core (and six virtual)

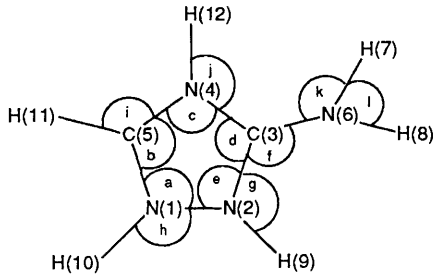


Fig. 1 Atom labelling and angle definition of 3-amino-1,2,4-triazole

orbitals were omitted from the expansion to give 862641 configurations. We denote these calculations 6-31G**(CCSD)//6-31G**.

In view of the success of semi-empirical methods in predicting tautomeric equilibria in heterocyclic molecules, we have also carried out calculations of the structures and energetics of the three tautomers of 3-amino-1,2,4-triazole using AM1 and MNDO-PM3 Hamiltonians as implemented in the program MOPAC.[14] (See Fig. 1 for atom labelling and angle definition of 3-amino-1,2,4-triazole).

The structures optimised at the 3-21G and 6-31G** level are summarised in Table 1, where comparison with the observed solid state structure of the 2*H* tautomer is made.[15] Both the 3-21G and 6-31G** basis sets predict structures for this tautomer in good agreement with experiment, with average heavy atom bond length deviations of ~0.02 Å and mean bond angle deviations of less than 1°. For the structures predicted by AM1 and MNDO-PM3 methods, the corresponding deviations are 0.06 and 0.04 Å, and less than 1.5° and 2.5°, respectively. It is in the structure of the amino group where the different levels of theory yield different predictions. At the 3-21G level the amino groups are essentially coplanar with the aromatic ring, the maximum out of plane angle for the amino hydrogen atoms being 4°. However, the inclusion of polarisation functions (6-31G**) renders the amino group non-planar in all three tautomers with corresponding out of plane angles of up to 58°. A similar effect was found in our studies of the tautomers of cytosine.[16] This lack of planarity is reflected in the C(3)–N(6) bond length. Thus, all other heavy atom bond lengths decrease on the addition of polarisation functions. The corresponding increase in the C(3)–N(6) bond length may be attributed to the decrease in N(6) lone pair conjugation for the non-planar amino group.

As far as the relative energies of the tautomers are concerned (Table 2), all levels of theory, both *ab initio* and semi-empirical, predict the energies of the 1*H* and 2*H* tautomers to differ by up to 1 kcal mol^{-1}. Such a small energy spread is expected, since the two structures are identical in the absence of the amino group. We note that increase in basis set size, and the inclusion of correlation effects preferentially stabilise the 1*H* tautomer. The most accurate calculation that we have carried out, [6-31G**(CCSD)//6-31G**] predicts the 1*H* tautomer to be more stable than the 2*H* tautomer by only 0.4 kcal mol^{-1}. The calculated zero point energy is essentially the same for both these tautomers so at the high level of theory that we have employed, both tautomers are essentially isoenergetic. For all the *ab initio* calculations that we have performed, the 4*H* tautomer has a significantly higher energy than the 1*H* or 2*H* tautomers. At the highest level of theory, this energy difference is 7.4 kcal mol^{-1}, a value close to the corresponding value of 7.0 kcal mol^{-1} for the unsubstituted 1,2,4-triazoles.[17] Both the PM3 and more particularly the AM1 methods seriously overestimate the relative stability of the 4*H* tautomer so that for the AM1 method it is predicted to be the most stable of the three tautomers. A similar failure of semi-empirical methods has been noted for the unsubstituted triazole, and has been attributed to underestimation of the lone-pair repulsion between adjacent nitrogen atoms.[18] A correction of ~12 kcal mol^{-1} has been suggested in order to compensate for this deficiency in the semi-empirical Hamiltonians. Our calculations indicate that a somewhat smaller correction is appropriate for the PM3 Hamiltonian.

Solvation Calculations.—The electrostatic contribution to the hydration energy of the three tautomers was estimated using both the SCRF and PCM methods discussed earlier. Here we have used a 6-31G** basis and structures optimised at this level. For the SCRF calculations, a cavity radius of 3.0 Å was employed, whilst for the PCM calculations, atomic radii were adjusted using the parameters of Aguilar and del Valle.[9] The results are summarised in Table 3. For the SCRF model used here, the solvation energy is essentially proportional to the square of the molecular dipole moment. The calculated solvation energies are in accord with this consideration. When combined with the energies of the free species, the SCRF model predicts the 2*H* tautomer to be most stable, followed by the 4*H* and 1*H* species. It is the large dipole of the 4*H* tautomer that results in the predicted change in the relative order of the tautomers upon hydration. The PCM treatment also predicts a greater solvation energy for the 4*H*, compared to the 1*H* and 2*H* tautomers. However, in contrast to the SCRF model, the 1*H* hydration energy is predicted to be somewhat larger than that for the 2*H* tautomer and although the hydration energy of the 4*H* tautomer is greater than that of the other two tautomers the differences are smaller than those predicted by the SCRF model. The result of these differences is that at the PCM level, the 1*H* tautomer is predicted to be the most stable in water, followed by the 2*H* species, with the 4*H* tautomer being at significantly higher energy.

We turn now to a comparison of our predictions with experimental solution data for the tautomeric equilibria. A study of 3-amino-1,2,4-triazole in aqueous solution using ^{15}N NMR spectroscopy concludes that the 4*H* tautomer is not present to any appreciable extent and that the 2*H* tautomer dominates, with the 2*H*:1*H* ratio being *ca.* 2:1.[19] Neither of the two models studied are in accord with these findings. Although the SCRF model does predict that the 2*H* tautomer dominates in aqueous solution, the 4*H* tautomer is predicted to be more stable than the 1*H* species. The PCM treatment, whilst failing to predict the most stable species in aqueous solution, succeeds in the prediction that the 4*H* tautomer is the least stable, and is unlikely to be observed.

Conclusions

We have presented calculations of the relative energies of the three tautomers of 3-amino-1,2,4-triazole at high levels of theory. In the gas phase, we predict that the 1*H* and 2*H* tautomers have essentially the same energy, with the 4*H* tautomer being of considerably higher energy. There are presently no data available to test these predictions. Low temperature matrix isolation studies of the infra-red spectra of this compound may be a valuable source of such data.

In aqueous solution it is the differential solvation energies of the 1*H* and 2*H* tautomers that determines their populations. The SCRF model predicts the 2*H* tautomer to be dominant in agreement with experiment, a result arising from the greater dipole moment of this species. However, if our gas phase calculations are accurate, then it is suggested that this model overestimates the relative solvation energy of the 4*H* tautomer. On the other hand, the PCM treatment does correctly predict that the 4*H* tautomer should not be observed in aqueous

Table 1 Calculated bond lengths/Å and angles/° of tautomers of 3-amino-1,2,4-triazole. Experimental values [15] are in parentheses

	1*H*		2*H*		4*H*	
	6-31G**	3-21G	6-31G**	3-21G	6-31G**	3-21G
Bond[a]						
N(1)–N(2)	1.358	1.418	1.359 (1.378)	1.423	1.373	1.443
N(2)–C(3)	1.298	1.311	1.333 (1.340)	1.348	1.282	1.299
N(2)–H(9)	—	—	0.991 (0.862)	0.992	—	—
N(1)–C(5)	1.321	1.336	1.288 (1.331)	1.300	1.274	1.281
N(1)–H(10)	0.991	0.992	—	—	—	—
C(5)–H(11)	—	—	1.070 (1.030)	1.062	1.070	1.062
C(5)–N(4)	1.300	1.309	1.358 (1.362)	1.375	1.367	1.390
N(4)–H(12)	—	—	—	—	0.992	0.994
N(4)–C(3)	1.358	1.383	1.300 (1.340)	1.318	1.355	1.353
C(3)–N(6)	1.372	1.349	1.371 (1.342)	1.347	1.382	1.371
N(6)–H(7)	0.995	0.994	0.996 (0.888)	0.996	0.999	0.993
N(6)–H(8)	0.995	0.994	0.996 (0.918)	0.996	0.999	0.995
Angle						
a	109.7	109.1	102.0 (102.6)	101.6	107.8	107.3
b	110.8	110.7	115.7 (114.8)	115.5	110.3	110.7
c	—	—	102.5 (103.2)	104.0	103.9	104.6
d	115.1	114.6	110.3 (109.5)	110.0	110.7	111.2
e	101.9	101.7	109.6 (109.7)	108.9	107.3	106.1
f	123.5	124.4	123.6 (125.1)	124.7	126.9	125.7
g	—	—	129.5 (130.4)	131.1	—	—
h	120.7	119.8	—	—	—	—
i	125.8	125.4	122.5 (128.2)	122.4	123.5	122.8
j	—	—	—	—	127.2	127.9
k	113.7	119.2	111.9 (126.3)	118.0	113.7	122.7
l	114.1	120.3	112.4 (107.7)	119.2	110.6	119.1

[a] See Fig. 1 for atom labelling.

Table 2 Total energies (a.u.) and relative energies[a] (kcal mol^{-1}) of 3-amino-1,2,4-triazole

	Tautomer		
Method	1*H*	2*H*	4*H*
3-21G//3-21G	−294.1638 (0.8)	−294.1651 (0)	−294.1539 (7.1)
6-31G**//3-21G	−295.8498 (0.5)	−295.8505 (0)	−295.8348 (9.8)
6-31G**//6-31G**	−295.8554 (0.5)	−295.8562 (0)	−295.8424 (8.7)
6-31G**(MP2)//6-31G**	−296.7967 (−1.0)	−296.7950 (0)	−296.7841 (6.8)
6-31G**(CCSD)//6-31G**	−296.7769 (−0.4)	−296.7764 (0)	−296.7646 (7.4)
AM1	(0.8)	(0)	(−4.1)
PM3	(−0.4)	(0)	(0.3)
ZPE 6-31G**//6-31G**	0.0758	0.0757	0.0752

[a] Relative energies are in parentheses.

Table 3 Hydration free energies (ΔG_{solv}) and total relative free energies in water (ΔG_{tot}) (kcal mol^{-1})

	Tautomer		
Model	1*H*	2*H*	4*H*
SCRF			
ΔG_{solv}	−2.0 (2.4)[a]	−4.9 (3.7)	−10.9 (5.7)
ΔG_{tot}	2.5	0	1.4
PCM			
ΔG_{solv}	−15.4	−14.0	−18.2
ΔG_{tot}	−1.8	0	3.2

[a] The calculated dipole moment (D) is given in parentheses.

solution, but predicts that the 1*H* tautomer should dominate. This result is in conflict with ^{15}N NMR data. Clearly this system provides a sensitive test of solvation models and warrants further study.

Acknowledgements

We thank SERC (UK) for support of this research. H. E. S. was supported by the US National Science Foundation.

References

1 (*a*) J. S. Kwiatowski, T. J. Zielinski and R. Rein, *Adv. Quantum Chem.*, 1986, **18**, 85; (*b*) A. R. Katritzky and M. Karelson, *J. Am. Chem. Soc.*, 1991, **113**, 1561; (*c*) I. R. Gould and I. H. Hillier, *J. Chem. Soc., Perkin Trans. 2*, 1990, 329.
2 R. Rein and M. Shibata, in *Nucleic Acid; the Vectors of Life*, eds. B. Pullman and J. Jortner, Reidel, Dordrecht, 1986, p. 479.
3 O. Tapia and O. Goscinski, *Mol. Phys.*, 1975, **29**, 1653.
4 M. Karelson, T. Tamm, A. R. Katritzky, S. J. Cato and M. C. Zerner, *Tetrahedron Comput. Methodol.*, 1989, **2**, 295.
5 M. F. Guest and J. Kendrick, GAMESS User Manual CCP1/86/1, Daresbury Laboratory, 1986.
6 O. G. Parchment, I. H. Hillier and D. V. S. Green, *J. Chem. Soc., Perkin Trans. 2*, 1991, 799.
7 I. R. Gould, D. V. S. Green, P. Young and I. H. Hillier, *J. Org. Chem.*, in the press.
8 S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.
9 M. A. Aguilar and F. J. Olivares del Valle, *Chem. Phys.*, 1989, **129**, 439.
10 R. D. Amos and J. E. Rice, CADPAC, The Cambridge Analytic Derivative Package, issue 4.0, Cambridge, 1987.
11 R. J. Bartlett, *J. Phys. Chem.*, 1989, **93**, 1697.
12 PSI 1990, Psitech Inc., Watkinsville, Georgia, USA.
13 G. E. Scuseria, C. L. Janssen and H. F. Schaefer, *J. Chem. Phys.*, 1988, **89**, 7382.
14 J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 209.
15 G. L. Starova, O. V. Frank-Kamenetskaya, V. V. Makarskii and V. A. Lopirev, *Kristallografiya*, 1978, **23**, 849.
16 I. R. Gould and I. H. Hillier, *Chem. Phys. Lett.*, 1989, **161**, 185.
17 J. R. Cox, S. Woodcock, I. H. Hillier and M. A. Vincent, *J. Phys. Chem.*, 1990, **94**, 5499.
18 (*a*) E. Fos, J. Vilarrasa and J. Fernandez, *J. Org. Chem.*, 1985, **50**, 4894; (*b*) J. P. Ritchie, *J. Org. Chem.*, 1989, **54**, 3553.
19 H. Fritz, *Bull. Soc. Chim. Belg.*, 1984, **93**, 559.

Paper 2/02597C
Received 19*th May* 1992
Accepted 16*th June* 1992