The Prediction of Tautomer Equilibria in Hydrated 3-Hydroxypyrazole: A Challenge to Theory

Oswald G. Parchment,[†] Darren V. S. Green,[†,§] Peter J. Taylor,[‡] and Ian H. Hillier*,[†]

Contribution from the Chemistry Department, University of Manchester, Manchester, M13 9PL, U.K., and ICI Pharmaceuticals, Alderley Park, Macclesfield, Cheshire, SK10 4TG, U.K.
Received June 18, 1992

Abstract: Experimental and theoretical estimates of the tautomeric equilibria in aqueous 3-hydroxypyrazole are discussed. The use of pK_a data is shown to be critically dependent upon the inclusion of corrections reflecting the substitution pattern of the compounds employed. Ab initio calculations, including a high level of electron correlation, of the gas-phase species are combined with both molecular dynamics simulations and continuum methods, to model solvation effects. A combination of experimental data and gas-phase calculations yields estimates of the relative solvation energies which are compared to the results from the various theoretical models. It is concluded that solute polarization is important in predicting solvation energies and that none of the models considered are totally adequate in this respect.

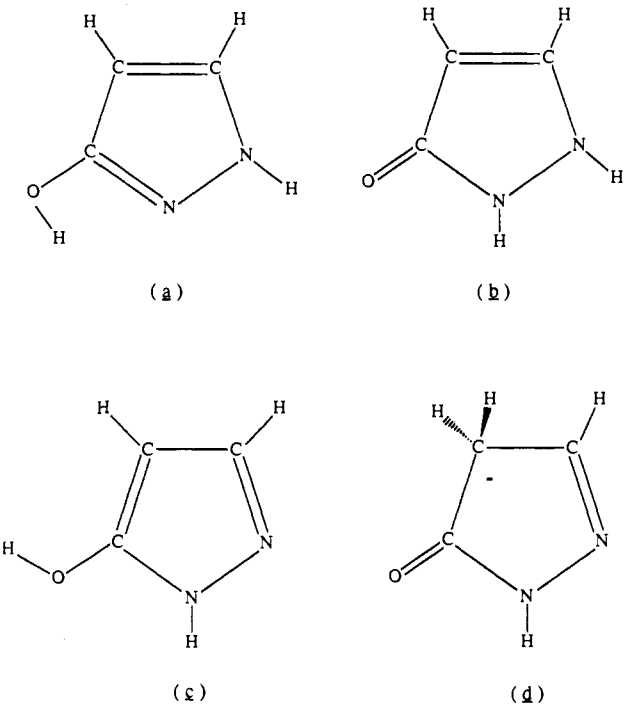
Introduction

Among the still remaining problems in heteroaromatic tautomerism, those that involve compounds containing contiguous heteroatoms are probably the most notable. So far as annular tautomerism is concerned, there is a strong presumption, due chiefly to Elguero,[1,2] that contiguous heteroatoms of the same hybridization type are strongly destabilizing. This will help to explain the strong preponderance of the 1H tautomer of 1,2,4-triazole[1] and the lesser, but still definite[1,3] preference of 1,2,3-triazole for the 2H form. Recent ab initio calculations by Elguero and co-workers[4,5] on both neutral species and cations as isolated molecules, and by us[6] on the neutral species both as isolated molecules and when solvated, have strongly confirmed this picture. Lone pair repulsion between contiguous imino nitrogens,[2,3] or dipolar repulsion between contiguous NH's,[2] is thought to account for this phenomenon.

In this paper we study the case of 3-hydroxypyrazole which can exist in four major tautomers according to solvent and the mode of substitution. We consider the two oxo forms (*b* and *d*) and the two hydroxy forms (*a* and *c*). On the basis of repulsion due to the adjacent)N–H groups, we might expect tautomer *b* to be destabilized in the gas phase. However, the effect of solvent on pyrazolone tautomerism is predictable in qualitative terms.[7] Water as the strongest common proton donor solvent[8] shows the highest proportion of this *b*-type oxo form, whereas the much less polar *d*-type oxo form is only detectable in nonpolar solvents, where it predominates. In amphiprotic solvents other than water the balance shifts toward the hydroxy form. It is generally believed[1] that it is the *a* rather than the *c* form.

We present here ab initio calculations of the tautomer preferences both in the gas phase and in aqueous solution. We use state-of-the-art methods to predict the gas-phase energetics and a number of approaches to model hydration energies. These are then compared both with the experimental results and with an approach, developed by one of us,[3] which applies correction factors to the conventional pK methodology[1] for aqueous solution that is discussed below, and appears capable of yielding a closer agreement with experiment than has hitherto been possible. This approach, if validated, shares with calculation the ability to predict tautomer ratios too large to be experimentally accessible, so it will first be described, followed by calculations of the tautomer energetics in the gas phase and in aqueous solution.

The Measurement of Tautomeric Ratio. The direct measurement of tautomer composition, while definitive in principle, is in practice limited by the signal-to-noise (or signal-to-signal) ratios of the (mostly spectroscopic) techniques that can be employed for this purpose,[1] and indirect methods have in practice to be used if tautomeric ratio very much exceeds one order of magnitude. Nevertheless it is equally difficult to validate these indirect methods, and one virtue of the present study is that both approaches, empirical and calculational, are applied here to a case in which the salient facts are quite accurately known.



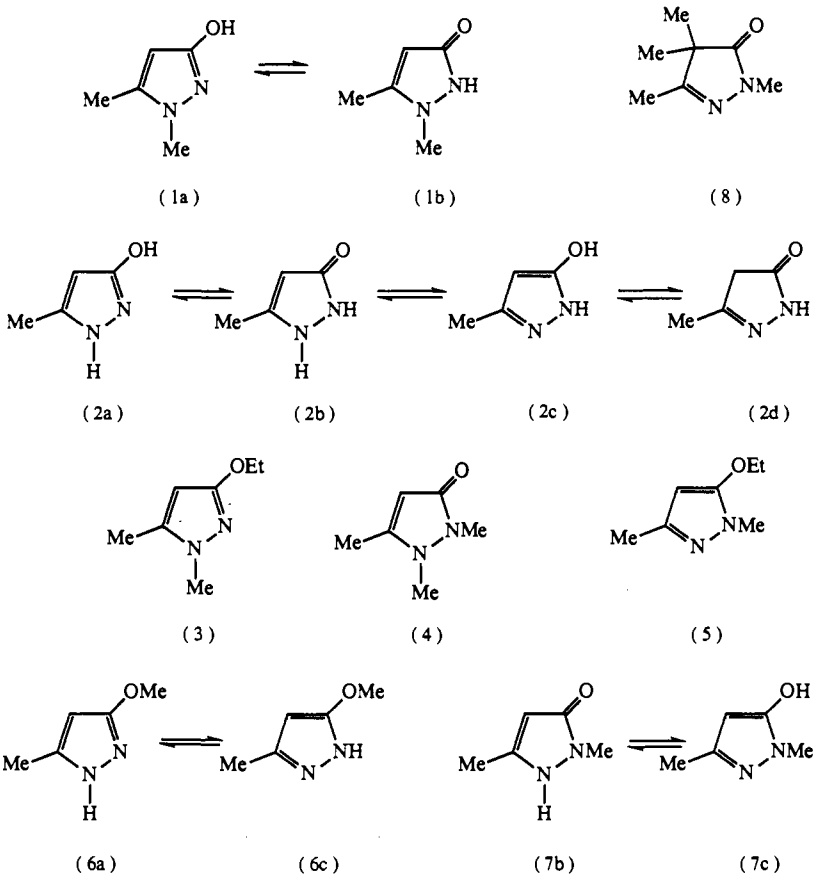
Chief among these indirect approaches is the basicity method.[1] This depends on the fact that many pairs of tautomers are related through a common cation, as e.g., **1a** and **1b** in Scheme I, where N-protonation and O-protonation respectively result in the same

[†] University of Manchester.
[‡] ICI Pharmaceuticals.
[§] Present address: Glaxo Group Research, Greenford Road, Greenford, Middlesex, UB6 0HE, U.K.

(1) Elguero, J.; Marzin, C.; Katritzky, A. R.; Linda, P. *The Tautomerism of Heterocycles*; Academic Press: New York, 1976.
(2) Taft, R. W.; Anvia, F.; Taagepera, M.; Catalán, J.; Elguero, J. *J. Am. Chem. Soc.* **1986**, *108*, 3237.
(3) Albert, A.; Taylor, P. J. *J. Chem. Soc., Perkin Trans. II* **1989**, 1903.
(4) Catalán, J.; de Paz, J. L. G.; Yanez, M.; Elguero, J. *Chem. Scr.* **1984**, *24*, 84.
(5) Catalán, J.; Sanchez-Cabezudo, M.; de Paz, J. L. G.; Elguero, J.; Taft, R. W.; Anvia, F. *J. Comput. Chem.* **1989**, *10*, 426.
(6) Cox, J. R.; Woodcock, S.; Hillier, I. H.; Vincent, M. A. *J. Phys. Chem.* **1990**, *94*, 5499.
(7) Elguero, J.; Jacquier, R.; Tarrago, G. *Bull. Soc. Chim. Fr.* **1967**, 3772, 3780.
(8) Abboud, J.-L.; Kamlet, M. J.; Taft, R. W. *Prog. Phys. Org. Chem.* **1981**, *13*, 485.

0002-7863/93/1515-2352$04.00/0

Scheme I



cationic species. Qualitatively, since the favored tautomer is that which most readily results from deprotonation of this common cation, it must necessarily be the weaker base. Quantitatively, if two or more tautomers share a common cation, then the contribution of each deprotonation route (i) to the overall proton loss equilibrium constant K_a, is expressed by eq 1.

$$K_a = \sum K_i \qquad (1)$$

At the same time, the proportion of each neutral species is given by K_i/K_a. Given the common situation that it is impossible to measure these proportions directly, recourse may be made to model compounds in which the position of the otherwise mobile proton(s) is fixed. Typically, O- of N-alkylation is employed for this purpose. Measurement of pK_a for these model compounds then has to assume either that basicity is unaffected by the alkylation process, or that it is affected to the same extent. We have defined this as the "naive basicity" hypothesis.[3]

There is however, considerable evidence that this hypothesis is misleading. Spinner and Yeoh[9] point out that both O- and N-methylation are expected to be base-weakening since alkylation of OH or NH leads to poorer solvation in the cation. Since solvation requirements are much greater for oxygen than for nitrogen cations,[10] a much greater correction is likely to be required for the former, and indeed Spinner and Yeoh[9] estimate this as a full pK unit. If tautomeric ratio is large enough, then qualitative conclusions are unlikely to be affected by such errors, as, for example, the log K_T ca. 3 estimated[1] for 2-pyridone in favor of the oxo form. However, they matter greatly when the result is a close-run thing; such calculations frequently give the wrong answer, sometimes reversing the observed stability order,[1] and there are even cases in which the calculation becomes incommensurable.[3,7]

Three correction factors will suffice for the pyrazolones considered here: NMe to NH when next to carbonyl, on protonation of the C = O; NMe to NH when contiguous to an imino nitrogen that becomes protonated; and OMe (or OEt) to OH as part of a ROC = N unit, on protonation of the nitrogen atom. For the first, we go to the case of 2-pyridone considered above. This and its *N*-methyl derivative possess pK_a 0.75 and 0.32, respectively,[11] hence a correction of ΔpK_a +0.43 appears to be required. We obtain the second from pyrazole (pK_a 2.52)[11] and its *N*-methyl derivative (pK_a 2.09),[11] which after statistical correction results in ΔpK_a +0.73. The absence of data for compound pairs of type HOC = N/ROC = N makes the third correction much more problematical. A rather involved comparison of ΔpK_a with the corresponding log K_α values,[12] which then has to be corrected for the different log K_α expected for an OH acid of similar pK_a,[12] leads to a value[13] indistinguishable from Spinner and Yeoh's estimate[9] of one. [The rationale of this treatment is that the proton donor ability of neutral species (known) should relate to that of cation (unknown), which in turn is the quantity responsible for ΔpK_a.]

There is sufficient model compound information[14] (Table I) to tackle two series of the 3(5)-methylpyrazol-5(3)-ones (Scheme I). The 1,5-dimethyl series **1** is the simpler, with only two possible tautomers. Here 70% of the oxo tautomer **1b** is found where the naive basicity hypothesis predicts 45%, i.e., the observed trend is reversed, as noted by Elguero.[7] Adjustment of the model compound pK values according to the proposed correction factors [column B of Table I] leads to a prediction of 75%. Even more notably, the observed overall (macro-) pK_a of 2.50 is closely predicted as 2.45, where naive basicity would predict 1.79; this is some evidence that the suggested factors are at least semiquantitatively correct.

(9) Spinner, E.; Yeoh, G. B. *J. Chem. Soc. B* **1971**, 279.
(10) Arnett, E. M.; Scorrano, G. *Adv. Phys. Org. Chem.* **1976**, *13*, 83.
(11) Perrin, D. D. *Dissociation Constants of Bases in Aqueous Solution: Supplement*; Butterworths: London, 1972.
(12) K_α is the equilibrium constant for hydrogen bond formation between *N*-methylpyrolidinone and a proton donor at 25 °C in 1,1,1-trichloroethane: Abraham, M. H.; Duce, P. P.; Prior, D. V.; Barratt, D. G.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans. II* **1989**, 1355.
(13) Taylor, P. J., unpublished results.
(14) Katritzky, A. R.; Maine, F. W. *Tetrahedron* **1964**, *20*, 299.

Table I. pK_a Values and Tautomer Composition in Water for Some 3(5)-Methylpyrazol-5(3)-ones

	A[a]	B[b]	C[c]
pK_i **1a** from **3**	2.05[d]	3.05	
pK_i **1b** from **4**	2.14[e]	2.57	
pK_a **1**	1.79	2.45	2.50[e]
x**1a**	0.55	0.25	0.3[e]
x**1b**	0.45	0.75	0.7[e]
pK_i **6a** from **3**	2.05[d]	2.78	
pK_i **6c** from **5**	3.51[d]	4.24	
pK_a **6**	2.03	2.76	2.62[e]
x**6a**	0.97	0.97	
x**6c**	0.03	0.03	*f*, *g*
pK_i **7b** from **4**	2.14[e]	>2.14	
pK_i **7c** from **5**	3.51[d]	4.51	
pK_a **7**	2.12	>2.14	2.31[e]
x**7b**	0.96	ca. 1.0	
x**7c**	0.04	ca. 0.0	*f*, *g*
pK_i **2a** from **6**	2.62[e]	3.62	
pK_i **2b** from **7**	2.31[e]	2.74	
pK_i **2c** from **5**	3.51[d]	5.24	
pK_a **2**	2.12	2.69	2.70[e]
x**2a**	0.32	0.12	0.1[e,g]
x**2b**	0.64	0.88	0.9[e]
x**2c**	0.04	2.9×10^{-3}	*f*, *g*
pK_a **8**	−3.79[a]		

[a] Model pK_a values and conclusions derived therefrom. [b] "Corrected" pK_a values (see text) and conclusions derived therefrom. [c] Observed. [d] Reference 14. [e] Reference 7. [f] Unobservable. [g] Reference 1. [h] Attributed to form (c) in ref 7 but corrected to form (a) in ref 1.

The N-unsubstituted compound **2** is capable in principle of existing in a number of tautomeric forms of which some are unknown (and probably impossible[7]) for any pyrazolone, but four remain viable (Scheme I). Of these, **2d** does not share a common cation with the other three but fortunately can be eliminated since it is known to exist only in nonpolar solvents[1,7] [Given pK_a −3.79 for **8**, this mode of protonation for **1b** can also be eliminated]. As models for **2a–c**, we have compounds **6**, **7**, and **5**, respectively. The first two are unlikely to exist in the heavily disfavored c-form,[1] but as a precaution we have analyzed them via the fixed tautomers **3**, **4**, and **5** as appears in Table I. Even without correction, the pK values of **4** and **5** as models for **7b** and **7c** suggest the former to predominate, and its dominance becomes still greater if correction is applied. For **3** and **5** as models for **6a** and **6c** the corrections cancel, but again the difference in stability is large; here the pK_a match for **6**, while satisfactory, is not quite so good as elsewhere. Hence, the final analysis for **2** corrects the pK values of **5–7** to give a result which, both for pK_a and for tautomer composition, corresponds with remarkable precision to what is actually found (Table I). In view of the calculations that follow, it may particularly be noted that the estimated proportion of the c-form is reduced from 4% on the naive basicity hypothesis[3] to less than a tenth of this value.

If final tautomer composition is assumed, it is possible alternatively to calculate sub-species pK_a values, and this has been tried.[15] However, back-calculation of macro pK_a's is not then so accurate, and this approach is clearly less rigorous than the above.

Computational Methods

Free Molecule Calculations. In studies[16] of a number of tautomeric equilibria using ab initio methods, it has been found that geometry optimization using at least a double-ζ basis, followed by energy calculations using a large basis with polarization functions and including electron correlation, is needed to obtain tautomer energy differences accurate to ~4 kJ mol^{-1}. Thus, the structures of the four tautomers studied here were optimized at the 3-21G[17] level and characterized as minima by calculations of their harmonic frequencies. Extended basis set (6-31G**)[18] calculations were then carried out at these geometries (6-31G**//3-21G). For tautomers *d* and *b* geometry optimization was also carried out at the 6-31G**//6-31G** level. Subsequently, the effect of electron correlation was studied at the 6-31G**//3-21G level, using Moller–Plesset perturbation theory[19] to second, third, and full fourth order (MP2, MP3, MP4SDTQ). These calculations were carried out using the programs GAMESS[20] and GAUSSIAN90.[21]

Modeling of Solvent–Solute Interactions. Two strategies have been employed to estimate solute–solvent interaction energies.

The first considered such interactions explicitly using a free energy perturbation (FEP) method[22] in molecular dynamics simulations to obtain directly the contribution of solvation ($\Delta\Delta G_{solv}$) to the total free energy difference ($\Delta\Delta G_{tot}$) between pairs of tautomers. In this method, mutation of one tautomer A into another B is accomplished by means of a coupling parameter λ. At any point of this mutation, the molecular mechanics Hamiltonian ($\mathcal{H}$) which describes the system is a mixture of the Hamiltonians for A and B, $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, such that

$$\mathcal{H} = \lambda\mathcal{H}_A + (1 - \lambda)\mathcal{H}_B$$

These calculations were carried out using the program AMBER[23] with atomic partial charges obtained using the method of Singh and Kollman[24] (using a 6-31G* basis) and van der Waals parameters from Weiner et al.[25] The molecular dynamics simulations were carried out at $T = 300$ K and 1 atm of pressure in a water bath containing 580 TIP3P water molecules. The perturbation calculations were performed with the window growth method, using a series of 21 "windows" with the coupling parameter (λ) differing by 0.05 between each window. For each value of λ, 500 steps of equilibration and data collection were performed with a time step of 0.001 and 0.002 ps, respectively, using periodic boundary conditions. We have thus followed the strategy of Cieplak et al.[26] who successfully described the tautomeric equilibria for 2-oxopyridine, 2-oxopyrimidine, and cytosine tautomers in aqueous solution. In particular, the use of a larger number of steps (1000) for data collection and equilibration was investigated and produced no significant change in the solvation free energy difference for the mutation $b \rightarrow d$.

The second strategy employed two continuum models, in which the solute is modeled in a cavity surrounded by solvent characterized by a relative permitivity (ϵ). We first use the self-consistent reaction field (SCRF) model developed by Tapia and Goscinski[27] and by Rivail and co-workers.[28] This model follows the work of Kirkwood[29] who derived the expression for the free energy of interaction of a charge distribution immersed in a cavity surrounded by a dielectric continuum. The solute charge distribution is represented by a single center expansion in multipole moments and appropriate modification of the one-electron terms in standard SCF programs allows for minimization of the sum of the molecular electronic energy and solvation free energy. The formalism may be extended from the spherical cavity model, to the case of an ellipsoidal cavity, which allows for the more realistic modeling of the shape of the solute molecule. In this work we first use a spherical cavity of radius 2.5 Å and truncate the multipole expansion at the dipole level. We have also carried out calculations of the relative energies of the solvated species using the SCRF method and the semiempirical AM1 Hamiltonian as implemented in the package MOPAC.[30] The ab initio calculations were extended by the use of ellipsoidal cavities whose dimensions are determined by the atomic van der Waals radii as suggested by Rivail et al. Here we extend the multipole expansion up to $\ell = 6$. These calculations were carried out using the computer code of Rivail

(15) Katritzky, A. R.; Maine, F. W.; Golding, S. *Tetrahedron* **1965**, *21*, 1693.
(16) Moreno, M.; Miller, W. H. *Chem. Phys. Lett.* **1990**, *171*, 475. Gould, I. R.; Hillier, I. H. *Chem. Phys. Lett.* **1989**, *161*, 185.
(17) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.
(18) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(19) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(20) Guest, M. F.; Kendrick, J. GAMESS User Manual CCP1/86/1, Daresbury Laboratory, 1986.
(21) Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Robb, M.; Binkley, J. S.; Gonzalez, C.; DeFrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R. L.; Kahn, L. R.; Stewart, J. J. P.; Topiol, S.; Pople, J. A. GAUSSIAN90, Revision J, Gaussian Inc., Pittsburgh, PA, 1990.
(22) Jorgenson, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.
(23) Singh, U. C.; Weiner, P. K.; Caldwell, J. W.; Kollman, P. A. AMBER (UCSF) Version 3.0, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, 1986.
(24) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(25) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
(26) Cieplak, P.; Bash, P.; Chandra Singh, U.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 6283.
(27) Tapia, O.; Goscinski, O. *Mol. Phys.* **1975**, *29*, 1653.
(28) Rivail, J. L.; Terryn, B. *J. Chim. Phys.* **1982**, *79*, 2. Rinaldi, D. *Comput. Chem.* **1982**, *6*, 155. Rinaldi, D.; Rivail, J. L.; Rguini, N. *J. Comput. Chem.* **1992**, *13*, 675.
(29) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351.
(30) Karelson, M.; Tamm, T.; Katritzky, A. R.; Cato, S. J.; Zerner, M. C. *Tetrahedron Comput. Methodol.* **1989**, *2*, 295.

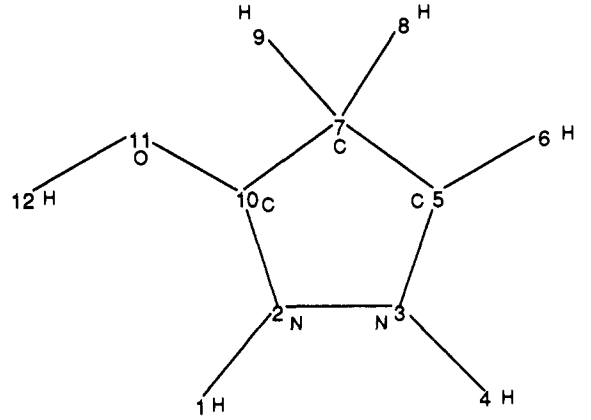


Figure 1. Numbering scheme for Table II.

et al. implemented within the GAUSSIAN90 program.

A further way of improving upon the use of the dipole approximation within a spherical cavity is to model a more realistic shape for the solvent cavity and a more accurate representation of the solute charge distribution as incorporated in a polarizable continuum model (PCM) developed by Tomasi and co-workers.[31] This method involves the generation of a cavity from spheres centered at each atom in the molecule and the calculation of virtual charges on the cavity surface to represent the polarization 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 wave function. The individual sphere radii of the PCM treatment were calculated in terms of the Mulliken atomic charges and basis set used, from the parameters of Aguilar and del Valle.[32]

Calculations of hydration energies using the SCRF and PCM treatments were carried out at the SCF (6-31G**//3-21G) level, using geometries optimized for the gas-phase molecules. The predictions of gas-phase energy differences between the different tautomers obtained at the highest level of theory (6-31G**(MP4SDTQ)) were then used together with these estimates of solvation energies to predict free energy differences between the tautomers in water.

Computational Results

The optimized geometries of the four tautomers at the 3-21G level are shown in Table II. All are predicted to be planar apart from the oxo form *b*, where ring puckering occurs, associated with out-of-plane hydrogen atoms, giving a H–N–N–H dihedral of 77.9°. This effect is increased when optimization is carried out at the 6-31G** level (H–N–N–H = 89.8°) although the relative energies of tautomers *d* and *b* are little altered, being 29.2 and 26.7 kJ mol^{-1}, respectively, at the 6-31G**//3-21G and 6-31G**//6-31G** levels. The predicted structure of tautomer *b* is in satisfactory agreement with the experimental solid-state structure for the 5-methyl derivative,[33] although the latter is found to have ring planarity. The nonplanarity of form *b* was confirmed by optimizing this form with the restriction of planarity imposed. At the 6-31G**(MP4SDTQ)//3-21G level the planar form is calculated to be 28.4 kJ mol^{-1} higher in energy than the nonplanar form, and calculation of the harmonic frequencies at the 3-21G level showed it to be a transition state.

The predicted energies of the isolated tautomers are shown in Table III. At all levels of theory, except 6-31G**(MP2)//3-21G, the 4H-oxo form (*d*) is predicted to be the most stable. No gas-phase data are available for this prediction to be tested. However, it is in agreement with the occurrence of form *d* in nonpolar solvents. The use of a basis having polarization functions on all atoms (6-31G**) preferentially stabilizes the forms *a* and *d*. The ordering of the tautomers predicted after the inclusion of correlation effects is dependent upon the level of correlation considered. However, we note that at the higher levels (MP3 and MP4) the order is the same, although the relative energies differ by up to 10 kJ mol^{-1}. This order is the same as that predicted at the SCF (6-31G**//3-21G) level. At this level, and when correlation effects are included, the highest energy tautomer is form *b*, which is in line with simple ideas previously discussed. We have not included the effect of zero-point energies in this discussion, since the values for all four tautomers are predicted to be within 2 kJ mol^{-1} at the 3-21G level. Statistical corrections to yield gas-phase free energies have been calculated

Table II. Optimized Geometries[a] at the 3-21G Level[b]

	d	*b*	*c*	*a*
		Bond Lengths (Å)		
1–2	0.9932	0.9965	0.9928	
2–3	1.4239	1.4032 (1.365)	1.3960	1.3922
3–4		1.0009		0.9916
3–5	1.2684	1.3863 (1.331)	1.3099	1.3517
5–6	1.0657	1.0701	1.0646	1.0648
5–7	1.5055	1.3288 (1.358)	1.4198	1.3619
7–8	1.0831	1.0676	1.0633	1.0617
7–9	1.0831			
7–10	1.5312	1.4670 (1.410)	1.3660	1.4179
10–11	1.2082	1.1950 (1.284)	1.3530	1.3549
10–2	1.3706	1.3905 (1.337)	1.3423	1.3031
11–12			0.9640	0.9665
		Bond Angles (deg)		
1–2–3	119.11	114.64	121.00	
2–3–4		111.00		119.25
2–3–5	107.14	104.83 (108.0)	103.89	111.09
3–2–10	113.67	109.51 (109.7)	111.30	104.06
3–5–6	120.91	118.48	120.31	122.26
3–5–7	113.66	112.32	112.61	107.75
5–7–8	112.94	128.39	127.70	128.78
5–7–9	112.93			
5–7–10	101.30	106.73 (107.2)	104.18	104.28
7–10–11	129.09	130.87 (131.9)	133.84	124.83
7–10–2	104.23	104.82 (106.1)	108.02	112.83
10–11–12			113.36	111.33

[a] See Figure 1 for atom numbering. [b] Experimental results[31] in parentheses.

but are not quoted since (at 298 K) these differ by less than 1 kJ mol^{-1} for the four tautomers studied herein.

We now turn to the prediction of the hydration energies obtained from the FEP method and the PCM and SCRF treatments summarized in Tables IV–VII. In the FEP calculations (Table IV), mutations between four pairs of tautomers have been carried out to judge the consistency of the simulations. The results show such consistency within the statistical bounds of the results. Furthermore, each forward and backward simulation was carried out twice and yielded similar consistency. The predicted differential solvation energies are in line with the calculated molecular dipole moments (Table VI) with form *b* being more strongly solvated than any of the three other tautomers. However, these predicted solvation energies do not differ sufficiently to change the relative energies from those predicted in the gas phase, $d < a < c < b$.

In Table V we show the electrostatic contribution to the solvation free energies for the four tautomers calculated by the PCM treatment, together with total free energy differences when these are combined with the gas-phase values. The differential solvation free energies in general follow the trend shown by the MD simulations, although the relative solvation energy of *b* is somewhat increased due to electron polarization effects. This results in an inversion of the ordering of *b* and *c* compared to the MD results, but, apart from this, the gas-phase ordering is preserved. This situation is different for the SCRF model using a spherical cavity and the dipole approximation (Table VI). Here tautomer *b* is predicted to be the most stable in aqueous solution. It can be seen that this result arises from the large degree of polarization occurring upon solvation, with the dipole moment increasing from 5.2 to 11.2 D. In this model, the solvation energy is proportional to the square of this quantity, resulting in the predicted change in the relative energies of the tautomers upon solvation. A similar effect is seen for the AM1 results, where *b* is again predicted to be the most stable form in aqueous solution. In addition, the differential solvation energies have now inverted the ordering of forms *c* and *a* from that found in the gas phase, so that the predicted order in water is $b < d < c < a$. At the AM1 level, the predicted ordering is $b < d < a < c$. However, when the ab initio SCRF

(31) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
(32) Aguilar, M. A.; Olivares del Valle, F. J. *Chem. Phys.* **1989**, *129*, 439.
(33) de Camp, W. H.; Stewart, J. M. *Acta Crystallogr.* **1971**, *27B*, 1227.

Table III. Total Energies (au) and Relative Energies (kJ mol^{-1}) of Tautomers of 3-Hydroxypyrazole

tautomer	3-21G//3-21G	6-31G*//3-21G	6-31G**//3-21G[a]	6-31G**(MP2)//3-21G	6-31G**(MP3)//3-21G	6-31G**(MP4SDTQ)//3-21G
d	−297.9745 (0)	−299.6616 (0)	−299.6708 (0) (−299.6734)	−300.5676 (0)	−300.5846 (0)	−300.6362 (0)
b	−297.9688 (15.2)	−299.6483 (34.9)	−299.6597 (29.2) (−299.6632)	−300.5549 (33.4)	−300.5707 (36.7)	−300.6219 (37.6)
c	−297.9672 (19.3)	−299.6455 (42.2)	−299.6602 (28.0)	−300.5644 (8.3)	−300.5788 (15.3)	−300.6269 (24.7)
a	−297.9711 (8.8)	−299.6510 (27.7)	−299.6656 (13.8)	−300.5702 (−6.8)	−300.5841 (1.3)	−300.6325 (9.9)

[a] In parentheses are 6-31G**//6-31G** values.

Table IV. Free Energy Differences (kJ mol^{-1}) of Tautomers in Water Obtained from MD Simulations

tautomer pair	$\Delta\Delta G_{solv}$	$\Delta\Delta G_{tot}$[a]
$d \rightleftharpoons b$	$d \rightarrow b$	
	−12.8	25.1 ± 0.4
	$b \rightarrow$ d	
	12.1	
$b \rightleftharpoons c$	$b \rightarrow c$	
	11.1	−2.7 ± 1.0
	$c \rightarrow b$	
	−9.2	
$c \rightleftharpoons a$	$c \rightarrow a$	
	10.0	−4.9 ± 0.1
	$a \rightarrow c$	
	−9.8	
$a \rightleftharpoons d$	$a \rightarrow d$	
	−6.4	−16.6 ± 0.3
	$d \rightarrow a$	
	7.0	

[a] Evaluated using 6-31G**(MP4SDTQ)//3-21G energies (Table III).

Table V. Solvation Free Energy (ΔG_{solv})[a] and Total Relative Free Energies (ΔG_{tot})[b] in Water, Obtained by PCM Treatment[c]

tautomer	ΔG_{solv}	ΔG_{tot}	μ (D)
d	−59.3	0	4.4
b	−68.3	28.4	7.0
c	−55.4	28.7	3.9
a	−54.6	14.7	3.3

[a] Evaluated at 6-31G**//3-21G level. [b] Evaluated using gas-phase values at 6-31G**(MP4SDTQ)//3-21G level. [c] All energies are in kJ mol^{-1}.

Table VI. Tautomer Energy Differences (kJ mol^{-1}) in Water Obtained from the Self-Consistent Reaction Field (SCRF) Method

tautomer	AM1	ab initio[a]	μ(6-31G**//3-21G) (D)[d]
d	0	0	3.0 (5.5)
b	−7.0	−55.9	5.2 (11.2)
c	54.5	20.8	3.1 (5.7)
a	38.6	21.9	2.4 (4.4)

[a] Evaluated at 6-31G**/3-21G level with gas-phase values at 6-31G**(MP4SDTQ)//3-21G level. [b] The values are for the gas phase, and in parentheses, for the hydrated species.

model is extended to use ellipsoidal cavities, and the multipole expansion up to $\ell = 6$ is used, the predictions are drastically altered. The relative solvation free energies (Table VII) are now very similar to those given by the PCM treatment. In particular the solvation energy of tautomer *b* is drastically reduced, so that the relative energies of the four species in aqueous solution, $d < a < b < c$, is now the same as that predicted by the PCM treatment. A comparison of the solvation free energies predicted by the various models studied (Table VII) shows the similarity between the relative values predicted by all models, with the exception of the SCRF model using a spherical cavity and the dipole approximation.

Comparison with Experiment. The experimental estimates of tautomer ratios in aqueous solution, discussed previously provide a sensitive test of the theoretical predictions. The experimental energy ordering is seen to be $b < a < c < d$. The predominance of the *b* over the *a* form in the ratio of <10:1 leads to a free energy difference between these forms of less than 6 kJ mol^{-1}, the corresponding estimate for the pair *b*, *c*, being ~17 kJ mol^{-1} with *b* being of lower energy.

Table VII. Comparison of Relative Solvation Free Energies (kJ mol^{-1}) Predicted by Continuum and Molecular Dynamics Models

tautomer	SCRF($\ell = 6$)[a]	PCM	MD
d	0 (0)	0	0
b	−8.7 (−93.5)	−9.0	−12.5
c	4.5 (−3.9)	3.9	−2.3
a	7.4 (12.0)	4.7	7.6

[a] The values are for an ellipsoidal cavity. In parentheses are values at the dipole level for a spherical cavity as in Table VI.

It is clear that none of the solvation models considered yield results of an accuracy comparable to that achieved from our analysis of the experimental data. For all continuum models, the predicted solvation energy is critically dependent on the cavity size. For a spherical cavity of radius *a*, a simple reaction field model yields a solvation energy that varies as a^{-3}. Similarly, with the PCM model, solvation energies are critically dependent on chosen atomic radii as exemplified by Aguilar and del Valle.[32] However, it is encouraging that both the PCM and SCRF ($\ell = 6$, ellipsoidal cavity) models yield relative solvation energies that agree to better than 5 kJ mol^{-1}. Such an agreement is expected since the PCM treatment should model the SCRF treatment for a sufficiently large value of ℓ. For the MD simulations, the results naturally depend upon the molecular mechanics parameters used, in particular the formal atomic charges. We have followed the strategy of Cieplak et al.[26] which was successful in modeling the tautomeric equilibria in 2-hydroxypyridine and cytosine and use it without further justification. It is pleasing that the relative solvation energies are close to those predicted by the continuum models, particularly the PCM model, where the agreement is better than 6 kJ mol^{-1}. As a result, all three models predict essentially the same tautomer ordering, $d < a < b \approx c$. This is in disagreement with the experimental ordering, $b < a < c < d$. A possible origin of this discrepancy comes from consideration of the results from the SCRF ($\ell = 1$, spherical cavity) model (Table VI). Here, the predicted ordering, $b < d < c \approx a$, arises from the large solvation energy of *b* (93.5 kJ mol^{-1}), about half of which arises from electron polarization, resulting in a remarkable increase in the dipole moment upon solvation (Table VI). Such polarization is reduced in both the PCM and SCRF ($\ell = 6$) models, which are expected to be more accurate. However, even for these models the polarization of *b* is greater than for the other tautomers, and it is possible that a more accurate treatment of solute polarization (not considered in our MD simulations) would lead to more close agreement with experiment. It is of interest to note that it is the SCRF treatment, used in conjunction with the AM1 Hamiltonian, which yields the correct ordering for the tautomers *b*, *c*, and *a* but not for form *d*. However, in view of our ab initio results for $\ell = 6$, we consider such an agreement to be fortuitous.

Finally we note that we may couple our experimental estimates of the equilibrium constants, with our predicted gas-phase energies (Table III) (which are probably not greatly in error) to yield estimates of the *relative* stabilization due to hydration, for tautomers *a*, *b*, and *c*. The respective values are *a*, 0; *b*, 34; *c*, 4 kJ mol^{-1}. These values are in line with the calculated gas-phase dipole moments (Table VI), the relative values for *c* and *a* being in quite good agreement with the results of all models. However, comparison with the predictions of Table VII suggest additional differential polarization of *b* of 15–20 kJ mol^{-1}, above that predicted by the models used here. Clearly further theoretical studies of this system are needed.

Acknowledgment. We thank SERC (U.K.) for support of this research and Professor J. L. Rivail for use of his computer codes.