Modeling of tautomerism of pyridine-2(1*H*)-thione from vapor to solution †

Víctor Martínez-Merino * **and María J. Gil**

Departamento de Quimica Aplicada, Universidad Publica de Navarra, E-31006 Pamplona, Spain. E-mail: merino@upna.es

Received (in Cambridge) 11th November 1998, Accepted 16th February 1999

Using the well-characterized pyridine-2(1*H*)-thione (**1**)/pyridine-2-thiol (**2**) system as a model of thiated nucleobases in DNA or RNA, *ab initio* (HF, MP2 and MP4) and density functional theory (B3LYP) methods using large basis sets were applied to reproduce structures and tautomerization free energies. Hydrogen bonds of the dimerized thione **1** in the solid state were well-reproduced by the B3LYP method. Electronic correlation similarly stabilized both **1** and **2** protomers, and so the gas-phase free energy of the **1/2** equilibrium was equally overestimated by 4 kJ mol^{-1} in favour of the **2** by HF, MP2(full) or MP4(SDTQ)/6-311+G(2d,p) *ab initio* methods. Otherwise, the B3LYP method systematically overestimated the relative stability of **1**, but using the 6-311+G(2d,p) basis set, the error was comparable to *ab initio* methods. Dunning's D95+(2d,p) and Aug-cc-pVDZ basis sets gave worse protomeric free energies than Pople's basis sets using the B3LYP method. The nonspecific solvent effect on **1/2** protomerization was accurately predicted (within 2 kJ mol^{-1}) by the SCIPCM self-consistent reaction field (SCRF) methodology at the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d) level of theory. The Onsager SCRF method or HF theoretical treatment of the solvent effects gave less good results.

Introduction

Nucleic acid (NA) bases and their thio analogs are of great interest in pharmacy, as antineoplastic, antiviral, antifungal and antiparasitic agents.[1] Their biological relevance has accelerated theoretical and experimental studies on these systems. Several teams have calculated the protomeric equilibria free energies in the gas phase of thio analogs of NA bases[2–4] and their model systems[4] using high level *ab initio* and DFT methodologies. *Ab initio* high order electronic correlation (MP4 or QCISD) seems to be necessary to approximate the free energies of these equilibria. The theoretical study of tautomerism in thio analogs of NA bases in solution has not been very successful using Tomasi's polarized continuum model +HF/6-31G*[2] nor Onsager's +MP2/6-31G(d,p)[3] self-consistent reaction field calculations. In spite of the above high level calculations, the experimental errors of many known free energies approximated to the difference in energy between closer tautomers,[4] and so conclusions about which level of theory is necessary to reproduce them remain unclear.

In the present work we used the experimentally well-characterized pyridine-2(1*H*)-thione (**1**)/pyridine-2-thiol (**2**) protomeric equilibrium in the gas phase[5] and solution[6] to check the validity of theoretical methods in the study of this tautomerism. The **1/2** system is usually recognized as a model of the thiol/thione rearrangement on thio analogs of NA bases.[4] First, an analysis of *ab initio* (HF, MP2 and MP4) and DFT (B3LYP, B3P86 and B3PW91) methods with several basis sets (D95, cc-pVDZ and from 6-31G to 6-311G) and augmented with diffuse and polarization functions was carried out on the **1/2** protomeric equilibrium in the gas phase. The **1/2** equilibrium in solution was modeled using the Onsager and SCIPCM self-consistent reaction fields for nonspecific solute–solvent interactions. We did not consider specific solute–solvent interactions because Beak *et al.*,[6] through a model based on a derivation of the Onsager–Kirkwood equations,[7] demonstrated that they were negligible (*ca.* 2 kJ mol^{-1}) in the free energy of solvation for the **1/2** system. The results may provide valuable information on the theoretical study of thiated NA bases, related heterocyclic systems and DNA.

Results and discussion

Geometrical considerations

Molecular geometries of 2-mercaptopyridine derivatives are known only for the crystallographic structures, while the comparison of experimentally calculated (g94) geometric parameters must be done in the gas phase or in solution (not available so far for the species under study). The compound pyridine-2(1*H*)-thione (**1**) is the preferred tautomer in the solid state whereas pyridine-2-thiol (**2**) is preferred in the gas phase.[5] Neutron diffraction spectroscopy demonstrated that crystalline pyridine-2(1*H*)-thione **1** was packed as dimers, where two molecules are linked to each other by two N–H · · · S hydrogen bonds through a crystallographic center of symmetry.[8] The hydrogen bond in the N–H · · · S group was bent (angle at H = 164°) and the nitrogen–sulfur distance was 3.289 Å.[8] Previous papers reported gas-phase geometrical optimizations of **1** as monomeric species at equal[9] or lower[5] than MP2 or B3LYP/6-31G(d,p) levels of theory. However, except for the hydrogen bond length, HF/3-21G(d) calculations agree well with experimental bond lengths and very well with angles of the dimeric molecule, as Fig. 1 shows. Calculated carbon–carbon bond lengths were slightly shorter than corresponding experimental double bonds and slightly longer than single bonds. Theoretical CX bond lengths were very exact, but the NS distance in the hydrogen bond was 0.120 Å longer. When electronic correlation was taken into account using the B3LYP/3-21G(d) method, the distances and angles of the hydrogen bonds were described very exactly (Fig. 1). At this level of

---

† Molecular structure data in *xyz* GAUSSIAN94 format are available as supplementary data (SUPPL. NO. 57503) from the British Library. For details of the Supplementary Publications Scheme, see 'Instructions for Authors', *J. Chem. Soc., Perkin Trans. 2*, available *via* the RSC web page (http://www.rsc.org/authors). For direct electronic access see http://www.rsc.org/suppdata/p2/1999/801/.

Table 1 Differences (experimentally calculated) in bond lengths (Δd) and angles (Δa) ± standard deviations (SD) between gas phase *ab initio* or DFT calculations and neutron diffraction of solid pyridine-2(1*H*)-thione **1**

Entry	Geometry optimization	Δd_{CC} ± SD[a]/Å	Δd_{CX} ± SD[b]/Å	Δd_{CH} ± SD[c]/Å	Δa ± SD[d] (°)
1	*HF/3-21G(d) dimer*[e]	*0.004 ± 0.015*	*0.000 ± 0.004*	*0.016 ± 0.003*	*0.0 ± 0.6*
2	HF/3-21G(d)	0.002 ± 0.023	0.001 ± 0.021	0.016 ± 0.003	0.2 ± 0.7
3	HF/6-311+G(2d,p)	0.004 ± 0.024	0.008 ± 0.016	0.013 ± 0.003	0.1 ± 0.6
4	MP2-full/6-311+G(2d,p)	−0.005 ± 0.007	0.000 ± 0.031	0.002 ± 0.003	0.1 ± 1.1
5	*B3LYP/3-21G(d) dimer*[e]	*−0.008 ± 0.007*	*−0.013 ± 0.014*	*0.003 ± 0.003*	*0.0 ± 0.8*
6	B3LYP/3-21G(d)	−0.010 ± 0.013	−0.013 ± 0.035	0.003 ± 0.003	0.3 ± 1.0
7	B3LYP/6-31G(d)	−0.009 ± 0.011	−0.009 ± 0.027	0.001 ± 0.002	0.2 ± 0.9
8	B3LYP/6-311+G(2d,p)	−0.004 ± 0.012	−0.004 ± 0.027	0.003 ± 0.003	0.2 ± 0.8

[a] Four C–C bond lengths. [b] Three bond lengths of the C6–N1–C2=S group. [c] Four C–H bond lengths. [d] Seven bond angles involving heavy atoms and four CCH angles. [e] Data from calculations on a group of two molecules bonded by two N–H · · · S interactions as in the solid state.

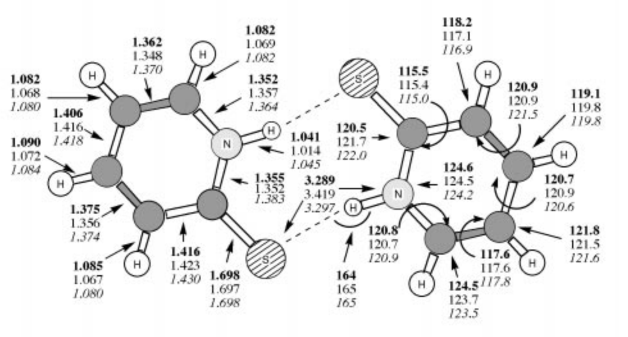

Fig. 1 Some selected distances (Å) and angles (°) of the dimer pyridine-2(1*H*)-thione from neutron diffraction of the solid (**bold face**), and from gas phase calculations at the RHF (normal) and B3LYP/3-21G(d) (*italic*) theoretical levels.

theory, all bond length and angle errors were *ca.* 1%, except for the N^1C^2 length (2%). The calculated CC bond lengths were slightly longer than experiment, whereas the N^1C^2 length was nearly 0.03 Å longer. On the other hand, the dipole moments of dimer species from the above calculations were zero.

Table 1 displays the geometric errors obtained when calculations of the gas phase monomer and dimer species of **1** were compared with neutron diffraction data of the solid dimer. Standard deviations (SD) of the monomer bond lengths increased with respect to the dimer (entries 1/2 and 5/6 of Table 1) because the thioamide group of the former does not participate in hydrogen bonding. Carbon–carbon bond lengths were better reproduced by methods which included electronic correlation, such as MP2 and B3LYP (entries 2–4, 6–8 of Table 1).

Theoretical C^2S^7 and NH bond lengths in the monomer decreased by nearly 0.03 Å as expected, but the MP2 and B3LYP length for N^1C^2 remained about 0.03 Å longer than that determined experimentally, even when large basis sets were used. Contrary to electron correlation methods, CC length errors of *ab initio* HF were not improved when the basis set was increased (entries 2 and 3 of Table 1). All C–H lengths were very exactly reproduced by either MP2 or B3LYP methods, but they were slightly shorter than those experimentally determined for HF (Δd_{CH} in Table 1). Bond angles were also accurately predicted by either method though correlation methods had greater SD as the calculated $N^1C^2C^3$ angle surpassed one degree of error. It is noteworthy that *ab initio* HF geometry was barely influenced by extending the basis set from 3-21G(d) to 6-311+G(2d,p), whereas the average bond length errors in the B3LYP/3-21G(d) calculation could be reduced to less than half at the B3LYP/6-311+G(2d,p) level.

The compound pyridine-2-thiol **2** does not exist in an isolable form and its *S*-methyl derivative is a liquid. Nevertheless, some attempts were made to compare the HF/3-21G(d) geometry of **2** with X-ray data of antimony(III) derivatives of **2**.[5]

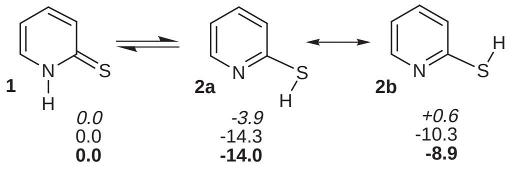

Fig. 2 Relative gas-phase free energies (ΔG_{298}) of tautomers and conformers using the 6-311+G(2d,p) basis set at the B3LYP (*italic*), HF (normal) and MP2-full (**bold face**) levels of theory for geometries and wavefunctions. ZPE + kinetic + entropic corrections were from B3LYP/6-311+G(2d,p)//B3LYP/6-311+G(2d,p) using a scaling factor of 0.98.

Free energies of tautomeric equilibria

The molecules had previously been studied by conformational analysis to find the most stable structures in the gas phase at the B3LYP/6-311+G(2d,p) level of theory.[10] This method has proved to be a very good choice for thermochemical calculations, being even more exact than the MP2/6-311+G(2d,p) method.[11] As Fig. 2 shows, we found two minima for the thiol form, one when the N^1–C^2–S^7–H dihedral angle equals zero (conformer **2a**) and another when it was 180° (conformer **2b**). The relative stability between conformers (**2a**/**2b**) calculated by *ab initio* methods without electronic correlation (HF) was essentially the same as by the MP2 or DFT method (B3LYP) (4–5 kJ mol^{-1} in favour of **2a**). On the contrary, the relative stability between the overall minimum of each tautomer in the protomeric equilibrium (**1** and **2a**) was remarkably different going from the *ab initio* [HF or MP2(full)] to the B3LYP methods. The experimental free energy of the **1**/**2** protomeric equilibrium in the gas phase (−9.9 ± 0.5 kJ mol^{-1} from a matrix isolation IR spectroscopic study at 350 K)[5] was 4 kJ mol^{-1} bigger than for HF or MP2(full)/6-311+G(2d,p) calculations and 6 kJ mol^{-1} smaller than the B3LYP estimation at the same level of theory (Fig. 2).

Other different theoretical methods and basis sets were checked to find a more accurate reproduction of the experimental free energy of tautomerization. Table 2 summarizes the obtained results in the calculation of the gas-phase electronic energy (ΔE_{elec}) for the **1**/**2a** thioamide–thioimidate protomeric equilibrium. This energy can be estimated from Gibbs free energy (ΔG), vibrational zero point energies (ΔZPE) and thermal (kinetic and entropic) corrections.[12] Thermochemical calculations carried out at the B3LYP/6-311+G(2d,p) level gave a difference in vibrational zero point energies of −10.7 kJ mol^{-1} (scaled by 0.98) between compounds **1** and **2a**. Similarly, ΔZPE (−10.8 kJ mol^{-1}) can be obtained from the HF/6-311+G(2d,p) method using a scaling factor of 0.89. The above scaling factors account for the difference between the harmonic vibrational calculation and the anharmonic oscillation of actual bonds. The rest of the thermal corrections (kinetic energy and entropic terms) to obtain Gibbs free energy, at 298 K and 1 atm pressure, were less important [−0.7 kJ mol^{-1} from scaled HF

Table 2 Influence of theoretical level in the variation of electronic energy (ΔE_{elec}) for the **1/2a** protomeric equilibrium in the gas phase and in the variation of dipole moments (μ) of tautomers

Entry	Method (energy//geometry)	ΔE_{elec} [a]	μ_1 [b]	μ_{2a} [b]
1	Estimated[c]	+1.5	—	—
2	B3LYP/Aug-cc-pVDZ//B3LYP/6-311+G(2d,p)	+11.8	5.623	2.029
3	B3LYP/D95+(2d,p)//B3LYP/6-311+G(2d,p)	+11.1	5.694	2.093
4	B3LYP/6-311+G(3df,2p)//B3LYP/6-311+G(2d,p)	+7.9	5.560	2.008
5	B3LYP/6-311++G(2d,p)//B3LYP/6-311+G(2d,p)	+7.4	5.614	1.987
6	B3LYP/6-311+G(2d,p)//B3LYP/6-311+G(2d,p)	+7.5[d]	5.61[d]	2.00[d]
7	B3LYP/6-311+G(d)//B3LYP/6-311+G(2d,p)	+14.0	5.740	2.073
8	B3LYP/6-311G(d)//B3LYP/6-311+G(2d,p)	+18.6	5.839	2.009
9	B3LYP/6-31+G(2d,p)//B3LYP/6-311+G(2d,p)	+9.3	5.677	2.029
10	B3LYP/6-31+G(d)//B3LYP/6-311+G(2d,p)	+10.9	5.796	2.096
11	B3LYP/6-31G(d)//B3LYP/6-311+G(2d,p)	+12.9	5.823	1.997
12	B3LYP/6-31G(d,p)//B3LYP/6-31G(d,p)	+14.1[e]	5.80[e]	2.01[e]
13	B3LYP/6-31G(d)//B3LYP/6-31G(d)	+12.9	5.829	2.019
14	B3P86/6-31G(d)//B3LYP/6-31G(d)	+15.3	5.838	2.041
15	B3PW91/6-31G(d)//B3LYP/6-31G(d)	+14.6	5.819	2.042
16	MP4(SDTQ)-fc/6-311+G(2d,p)//MP2-full/6-311+G(2d,p)	−3.8	—	—
17	MP4(SDQ)-fc/6-311+G(2d,p)//MP2-full/6-311+G(2d,p)	−2.2	—	—
18	MP2-fc/6-311+G(2d,p)//MP2-full/6-311+G(2d,p)	−3.7	—	—
19	MP2-full/6-311+G(2d,p)//MP2-full/6-311+G(2d,p)	−2.7[d]	5.52[d]	2.10[d]
20	HF/6-311+G(2d,p)//MP2-full/6-311+G(2d,p)	−4.1	—	—
21	MP2-full/6-31G(d)//MP2-full/6-311+G(2d,p)	+2.2	5.860	2.197
22	MP2-full/6-31G(d,p)//MP2-full/6-31G(d,p)	−2.0[e]	—	—
23	MP4(SDQ)-fc/6-31G(d,p)//HF/6-31G(d,p)	+1.6[e]	—	—
24	MP2-fc/6-31G(d,p)//HF/6-31G(d,p)	−3.0[e]	5.63[e]	2.14[e]
25	HF/6-311+G(2d,p)//HF/6-311+G(2d,p)	−2.9[d]	6.47[d]	2.21[d]
26	HF/6-31G(d)//HF/6-31G(d)	+3.6	6.766	2.341

[a] In kJ mol^{-1}. [b] In Debyes (1 D = 3.33564 × 10^{-30} C m). [c] Experimental Gibbs energy minus scaled (0.98) ZPE + kinetic + entropic corrections from B3LYP/6-311+G(2d,p)//B3LYP/6-311+G(2d,p) following ref. 12. [d] From ref. 10. [e] From ref. 9.

Table 3 Total energies (TE, in atomic units) and dipole moments (μ, in Debyes) of pyridine-2(1*H*)-thione **1** and pyridine-2-thiol **2a** optimized by Onsager and SCIPCM B3LYP/6-31G(d) methods

	Onsager					SCIPCM				
Medium	TE_1	μ_1	TE_{2a}	μ_{2a}	ΔTE/kJ mol^{-1}	TE_1	μ_1	TE_{2a}	μ_{2a}	ΔTE/kJ mol^{-1}
Wide	−646.4802799	5.83	−646.4753570	2.02	*12.9*	−646.4802799	5.83	−646.4753570	2.02	*12.9*
C_6H_6	−646.4829573	6.63	−646.4756895	2.24	*19.1*	−646.4862179	6.95	−646.4773129	2.29	*23.4*
$CHCl_3$	−646.4848007	7.19	−646.4759104	2.40	*23.3*	−646.4900066	7.66	−646.4785187	2.46	*30.1*
Pr^{i}OH	−646.4864861	7.69	−646.4761065	2.53	*27.2*	−646.4932940	8.26	−646.4795447	2.60	*36.1*
DMSO	−646.4869206	7.82	−646.4761560	2.56	*28.2*	−646.4941167	8.40	−646.4797993	2.63	*37.5*

or B3LYP/6-311+G(2d,p) methods]. So, the estimated ΔE_{elec} between **1/2a** protomers would be about +1.5 kJ mol^{-1} [−9.9 − (−10.7 − 0.7) where an additional thermal correction from 298 to 350 K was neglected] from experimental ΔG_{350} and calculated ΔZPE and thermal corrections. Because the estimated ΔE_{elec} was small, errors of ±2 kJ mol^{-1} led to the reversal of **1/2a** relative energy from *ab initio* to DFT methods.

Within hybrid DFT methods, we tested Becke's three parameter hybrid method using Lee, Yang and Parr (B3LYP),[13] Perdew 86 (B3P86),[14] and Perdew/Wang 91 (B3PW91)[15] correlation functionals. As entries 13–15 of Table 2 show, only minor changes were observed using the 6-31G(d) basis set, and the error of the B3LYP method was slightly smaller. However, the selection of basis sets was very important in the calculation of ΔE_{elec}. The successive addition of a diffuse function, a set of p polarization functions on hydrogens, and two sets of d polarization functions on heavy atoms, decreased the error of the B3LYP method significantly (entries 4–11 of Table 2). A second diffuse function had a very slight effect (entry 5) and the addition of more d and f polarization functions did not improve the results (entry 4). Dunning's D95+(2d,p)[16] and Aug-cc-pVDZ[17,18] basis sets gave bigger errors than Pople's 6-31+G(2d,p) or 6-311+G(2d,p) basis sets (entries 2–4 and 9 in Table 2). When diffuse and polarization functions were considered in second-order Møller–Plesset[19] post-HF calculations including all electrons in the correlation calculation (MP2-full), the thioimidate was stabilized by 4.8 kJ mol^{-1} more than the thioamide form when the basis set was extended from 6-31G(d) to 6-311+G(2d,p) (entries 19 and 21 in Table 2). Similar stabilization can be observed from the B3LYP method (entries 5 and 11), but in MP2(full) the major influence seems to come from the addition of p polarization functions on hydrogens (entries 21 and 22). The importance of hydrogen polarization functions in quantitative prediction of the protomeric tautomerisms by *ab initio* calculations has already been noted.[20] Using the 6-311+G(2d,p) basis set, the electronic correlation produced a net stabilization of the thioamide *vs.* thioimidate forms of only 1.4 kJ mol^{-1} by the MP2(full) method and of 0.4 kJ mol^{-1} when only valence electrons were included in the correlation (MP2 frozen core) (entries 18–20). The higher order of electron correlation also had a small effect (entries 16–18). Fourth-order Møller–Plesset including triple and quadruple excitations, MP4(SDTQ)/6-311+G(2d,p), recovered the ΔE_{elec} calculated by HF/6-311+G(2d,p) in this thioamide/thioimidate system (entries 16 and 20). A similar effect was recently reported for the amide/imidate system in NA bases.[21] The basis set also greatly influenced the MP4 calculations, as can be seen in Table 2 (entries 17 and 23). Finally, the B3LYP method was insensitive to the geometry optimization level whereas *ab initio* calculations seem to be more sensitive (entries 11/13 and 20/25).

Table 4 Free energies (ΔG)[a] and solvent effects (**$\Delta\Delta G$**)[b] in kJ mol^{-1} for **1/2a** protomeric equilibrium from corresponding SCRF B3LYP/6-31G(d) geometries

		Medium				
Entry	Wavefunction	Gas	C_6H_6	$CHCl_3$	Pr^{i}OH	DMSO
1	Experimental[c]	−9.9	−2.9	+2.9	+7.4	<+11.3
		0.0	***+7.0***	***+12.0***	***+17.3***	***<+21.2***
2	Onsager B3LYP/6-31G(d)	+1.6	+7.7	+12.0	+15.9	+16.9
		0.0	***+6.1***	***+10.4***	***+14.3***	***+15.3***
3	SCIPCM B3LYP/6-31G(d)	+1.6	+12.0	+18.8	+24.7	+26.2
		0.0	***+10.4***	***+17.2***	***+23.1***	***+24.6***
4	SCIPCM B3LYP/6-311+G(2d,p)	−3.9	+5.0	+10.8	+15.9	+17.1
		0.0	***+8.9***	***+14.7***	***+19.8***	***+21.0***
5	SCIPCM HF/6-311+G(2d,p)	−14.2	−3.2	+4.8	+11.7	+13.4
		0.0	***+11.0***	***+19.0***	***+25.9***	***+27.6***

[a] Theoretical free energies were computed as the sum of −11.4 kJ mol^{-1} [0.98 ΔZPE and thermal factors from B3LYP/6-311+G(2d,p) calculation in gas phase] and the relative energy (ΔTE) from indicated wavefunction and geometry. [b] Difference between gas phase ΔG and indicated ΔG in solution. [c] ΔG in gas phase from ref. 5, and ΔG in liquid solution from ref. 6.

The above calculations show that electronic correlation [MP4(SDTQ) *vs.* HF at the 6-311+G(2d,p) level] stabilizes similarly both thioamide/thioimidate tautomers in an heterocyclic system such as **1/2a**. The biggest effect in relative electronic energies comes from the basis set used in the calculations. *Ab initio* and DFT methods gave an additional relative stabilization of 5–6 kJ mol^{-1} for the thioimidate form when the basis set was extended with diffuse and polarization functions from 6-31G(d) to 6-311+G(2d,p). The geometry used in the energy calculations can vary the sign of ΔE_{elec} when tautomers close in energy were compared by *ab initio* methods. In spite of the high level calculations used, *ab initio* [MP4(SDTQ), MP2(full) or HF] methods overestimated by 4–6 kJ mol^{-1} the relative stabilization of the thioimidate tautomer **2a** in the gas phase whereas the DFT (B3LYP) method overestimated the thioamide tautomer **1** by the same amount (Table 2). Nevertheless, mainly because of high ΔZPE, all methods were in agreement with experiment and indicated thioimidate tautomer **2a** as the most stable in the gas phase. The experimental free energy favoured the thioimidate protomer by 9.9 kJ mol^{-1} whereas the MP4(SDTQ)/6-311+G(2d,p)//MP2(full)/6-311+G(2d,p), HF and B3LYP/6-311+G(2d,p) methods favoured **2a** by 15.2, 14.3 and 3.9 kJ mol^{-1} respectively (B3LYP vibrational and thermal corrections were used in all cases as indicated above). Contrary to common practice,[3,4,20] our results stress that caution is required for quantitative analysis of tautomeric populations in thiated nucleobases.

Solvent effects in tautomeric equilibria

For the most general case, the principal effects of solvation in free energy of tautomeric equilibria are modeled by five energy terms: differences in van der Waals radii which account for dispersion–repulsion plus volume effects, in cavitation, in restructuring the solvent, in hydrogen bonding with the solvent, and in electrostatic interaction with the solvent.[6] Self-consistent reaction field (SCRF) methods model the solvent as a continuum of uniform relative permittivity ε, the reaction field. The solute is placed into a cavity within the solvent. The Onsager SCRF model defines that cavity as a sphere of radius a_0 and the net stabilization of the solute is computed analytically from an induced dipole in the medium. The self-consistent isodensity polarized continuum (SCIPCM) SCRF model defines the cavity as an isodensity surface of the molecule and solves for the electron density which minimizes the energy taking into account that the isosurface and the electron density are fully coupled.[11] The SCRF methods do not consider the energy associated with restructuring the solvent nor specific solute–solvent interactions such as hydrogen bonds. Beak *et al.*,[6] using a multiple linear regression model, estimated for the **1/2** tautomeric equilibrium that the specific solute–solvent interactions favoured thioamide protomer **1** by only 0.7β (*ca.* 2 kJ mol^{-1} for the solvents studied here). It also seems reasonable that tautomeric differences in restructuring the solvent could be ignored in a first approximation. The above considerations should make the **1/2** system treatable by SCRF models. Table 3 shows total energies (TE) of the **1** and **2a** protomers after Onsager or SCIPCM geometric optimization at the B3LYP/6-31G(d) level of theory. Solvents stabilized **1** more intensely than **2a**, as expected from the greater polarizability and dipole of the thioamide with regard to the thioimidate protomer. Solvents also increased the polarity of the molecule, mainly in protomer **1** which had the greater dipole moment.

Comparing the Onsager model with SCIPCM, the former showed a smaller increase in the magnitude of solvation energy and the dipole moment for every solvent. Experimental free energies of the **1/2** protomerization have been determined by means of UV spectroscopy within a maximum error of ±3.8 kJ mol^{-1},[6] and they are displayed in Table 4 for the selected solvents (benzene, chloroform, propan-2-ol and dimethyl sulfoxide). Taking into account the similarities in the constitution of tautomers **1** and **2a**, one can assume that thermal energy and entropy will be similarly influenced by solvation and that both of the above terms from the equilibrium in the gas-phase could be used for liquid phases. Recently this approximation was suggested for amino/imino tautomerism.[22] Thus ΔG in Table 4 was calculated by adding ΔTE to −11.4 kJ mol^{-1} (−10.7 − 0.7, ΔZPE and thermal factors from the gas phase). The SCIPCM B3LYP/6-31G(d) method overestimated solvent effects ($\Delta\Delta G$ in Table 4) by 4.4 ± 1.1 kJ mol^{-1} for the benzene–DMSO series (entry 3 in Table 4). The Onsager B3LYP/6-31G(d) method underestimated the benzene solvent effect by 0.9 kJ mol^{-1} only, but underestimation increased with the relative permittivity of the solvent (−5.9 kJ mol^{-1} for DMSO). The SCIPCM error was mostly a consequence of the basis set used, and so the single point calculation at the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d) level reduced it to less than 1.7 ± 1.1 kJ mol^{-1} (entry 4 in Table 4). In spite of this, the relative stability (ΔG) of the protomers calculated by the last method varied from the experimental value because of the error of +6.0 kJ mol^{-1} obtained for the gas phase. However, *ab initio* SCIPCM HF/6-311+G(2d,p)//B3LYP/6-31G(d) calculations gave a qualitatively very good approximation to relative stability (ΔG), but the solvent effect ($\Delta\Delta G$) was overestimated by 6.5 ± 1.6 kJ mol^{-1} within this series (entry 5 in Table 4).

The relative abundance of conformers with different dipole moments should be modified by the solvent. The ΔTE of the **2a/2b** conformational equilibrium varied from 4.5 kJ mol^{-1} in the gas phase to 0.4 kJ mol^{-1} in DMSO solution at the

SCIPCM B3LYP/6-311+G(2d,p) level of theory. Thus in polar solvents, thioamide **1** will be the most abundant species, but the more polar thioimidate **2b** will exist in a very similar proportion to **2a**.

Conclusions

The B3LYP methodology has been shown to be a very efficient tool for the reproduction of geometries in the solid phase. Even at the B3LYP/3-21G(d) level of theory, hydrogen bonds in the dimerized thione **1** in the solid state were very exactly reproduced. Diffuse and polarization functions were necessary to match calculations and the experimental Gibbs free energy of the **1/2** protomeric equilibrium in the vapor phase and solution. The HF and B3LYP methods using Pople's 6-311+G(2d,p) basis set gave only reasonable free energies for the **1/2** equilibrium in the gas phase (±4–6 kJ mol^{-1}). The MP2 (frozen core or full) and MP4 (frozen core SDQ or SDTQ) calculations gave similar results to HF calculations at the 6-311+G(2d,p) level, indicating that electronic correlation effects similarly stabilized both tautomers for this kind of molecule. The results prevent the use of the above methods to develop quantitative population analysis of tautomers for thiated nucleobases, a hitherto common practice.

The Onsager B3LYP/6-31G(d) method underestimated, and the SCIPCM method overestimated, the stabilization effect on the thione form (**1**) *vs.* thiol form (**2**) of more polar solvents by about 4 kJ mol^{-1}. Single point calculations at the B3LYP/6-311+G(2d,p) level using the SCIPCM method approximated very well the solvent effect on free energy (overestimation of 2 kJ mol^{-1}). Methods which neglect electron correlation, such as SCIPCM HF/6-311+G(2d,p), gave the worst results and overestimated the solvent stabilization of **1** *vs.* **2** by 6 kJ mol^{-1}.

The results from the B3LYP/6-311+G(2d,p) method indicate that it is a very efficient tool for modeling geometries and solvent energies of thiated nucleobases.

Computational details and theoretical methods

All calculations were performed using the GAUSSIAN94 (g94)[23] suite of programs. Geometrical optimizations of neutral molecules in the gas phase were carried out at three different theoretical levels: restricted Hartree–Fock (RHF),[24] second-order Møller–Plesset[19] including all electrons in the correlation calculation (MP2-full) or keeping the core electrons frozen (MP2-fc), and DFT using the three-parameter hybrid functional developed by Becke (B3LYP).[13] The basis sets used were Pople's 6-311+G(2d,p) for H, C, N and S.[24,25] The analytical harmonic vibrational frequency calculations were performed to characterize the nature of stationary points on the potential energy surface and to estimate the zero-point vibrational energy (ZPE) and the Gibbs free energy (G_{298}) at the HF and B3LYP levels. The ZPE and the thermal energy corrections to G_{298} values were scaled by 0.89 at the HF level and by 0.98 at the B3LYP level to eliminate known systematic errors.[11,26]

To investigate further the effect of basis set on the structures and tautomerization energy of 2-mercaptopyridine, single-point energy calculations were carried out at the B3LYP level using Pople's[24,25] basis sets 6-31G(d) through 6-311+G(3df,2p), as well as Dunning–Huzinaga's[16] full double-zeta basis set D95+(2d,p) and Dunning's[17,18] correlation-consistent polarized valence double-zeta basis set Aug-cc-pVDZ for all atoms, based on the corresponding B3LYP/6-311+G(2d,p) geometries. Additionally, two other hybrid DFT methods (B3P86 and B3PW91) were checked using single point calculations on B3LYP/6-31G(d) geometries. The B3P86 and B3PW91 methods are Becke's three parameter functions with the non-local correlations provided by Perdew 86[14] and Perdew/Wang 91[15] expressions, respectively. Single-point calculations at the MP4/6-311+G(2d,p) level including single, double, triple and quadruple excitations (MP4-SDTQ)[27] have also been performed at the MP2(full)/6-311+G(2d,p) geometries.

Nonspecific solvent effects on the geometry and physicochemical properties of the molecules and relative stability at the tautomeric equilibrium were studied using the self-consistent reaction field (SCRF) with Onsager's model[28–30] and with the self-consistent isodensity polarized continuum model (SCIPCM)[11] developed from a reaction field based on the polarized continuum model proposed by Tomasi *et al.*[31] We considered four solvents (benzene, $\varepsilon = 2.28$; chloroform, $\varepsilon = 4.81$; propan-2-ol, $\varepsilon = 18.30$; dimethyl sulfoxide, $\varepsilon = 46.68$)[32] which span the relative permittivity (ε) of nonaqueous media. In the SCRF calculations, the solute is placed in a uniform electric field of solvent with a relative permittivity ε or reaction field.

In the SCRF Onsager model, the solute is assumed to occupy a spherical cavity of radius a_0 in the medium. The cavity radius for each conformer, a_0, was determined by performing a single-point calculation with the keyword VOLUME of g94 programs at the optimized geometry of the B3LYP/6-31G(d) level (gas phase). The resulting values a_0 were 4.24 Å for compound **1**, and 4.16 Å for its tautomer **2a** (dihedral angle N1–C2–S–H = 0°).

In the SCIPCM model, the cavity is defined as an isosurface of the molecule, and the coupling of the isosurface and the electron density are taken fully into account. This procedure solves for the electron density that minimizes the energy, including the solvation energy, which itself depends on the cavity which depends on the electron density. In this case, the effects of solvation are folded into the iterative SCF computation rather than comprising an extra step afterwards.

All SCRF geometry optimizations in solution were carried out at the B3LYP/6-31G(d) level of theory. Single-point energy calculations were also performed using SCRF HF or B3LYP/6-311+G(2d,p) wavefunctions and above SCRF B3LYP/6-31G(d) optimized structures.

Acknowledgements

We gratefully acknowledge financial support for this project by the Navarra Government (Department of Education) and by the UPNA Research Commission, Pamplona, Spain. We are also grateful to Dr J. I. Garcia for careful and constructive criticism.

References

1 J. B. Hobbs, *Comprehensive Medicinal Chemistry*, ed. C. Hansch, Pergamon Press, Oxford, 1990, vol. 2, ch. 7.3.
2 C. Alhambra, F. Luque, J. Estelrich and M. Orozco, *J. Org. Chem.*, 1995, **60**, 969.
3 M. J. Stewart, J. Leszczynsky, Y. V. Rubin and Y. P. Blagoi, *J. Phys. Chem.*, 1997, **101**, 4753.
4 M. J. Nowak, L. Lapinski, J. S. Kwiatkowski and J. Leszczynsky, *Computational Chemistry: Reviews of Current Trends*, ed. J. Leszczynsky, World Scientific Pub. Co. Pte. Ltd., Singapore, 1997, vol. 2, ch. 4, and references cited therein.
5 M. J. Nowak, L. Lapinski, H. Rostkowska, A. Les and L. Adamowcz, *J. Phys. Chem.*, 1990, **94**, 7406, and references cited therein.
6 P. Beak, J. B. Covington and J. M. White, *J. Org. Chem.*, 1980, **45**, 1347.
7 J. Powling and H. J. Bernstein, *J. Am. Chem. Soc.*, 1951, **73**, 4353.
8 U. Ohms, H. Guth, A. Kutoglu and C. Scheringer, *Acta Crystallogr., Sect. B*, 1982, **38**, 831.
9 J. S. Kwiatkowski and J. Leszczynsky, *J. Mol. Struct.*, 1996, **376**, 325.
10 V. Martinez-Merino and M. J. Gil, *J. Chem. Soc., Perkin Trans. 2*, 1999, 33.
11 J. B. Foresman and Æ. Frisch, *Exploring Chemistry with Electronic Structure Methods*, 2nd edn., Gaussian Inc., Pittsburgh, PA, 1996.
12 $G = H - TS = (E + RT) - TS = [(E_0 + E_{\mathrm{vib}} + E_{\mathrm{rot}} + E_{\mathrm{transl}}) + \mathrm{RT}] - TS$ where $E_0 = E_{\mathrm{elec}} + \mathrm{ZPE}$.
13 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
14 J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822.
15 J. P. Perdew and Y. Wang, *Phys. Rev. B*, 1992, **45**, 13244.

16 T. H. Dunning Jr. and P. J. Hay, *Modern Theoretical Chemistry*, ed. H. F. Schaefer, III, Plenum, New York, 1976, pp. 1–28.

17 D. E. Woon and T. H. Dunning Jr., *J. Chem. Phys.*, 1993, **98**, 1358.

18 R. A. Kendall, T. H. Dunning Jr. and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796.

19 (*a*) C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618; (*b*) M. Head-Gordon, J. A. Pople and M. J. Frisch, *Chem. Phys. Lett.*, 1988, **153**, 503; (*c*) M. J. Frisch, M. Head-Gordon and J. A. Pople, *Chem. Phys. Lett.*, 1990, **166**, 275; (*d*) M. J. Frisch, M. Head-Gordon and J. A. Pople, *Chem. Phys. Lett.*, 1990, **166**, 281.

20 A. Les and L. Adamowicz, *J. Am. Chem. Soc.*, 1990, **112**, 1504.

21 C. Colominas, F. J. Luque and M. Orozco, *J. Am. Chem. Soc.*, 1996, **118**, 6811.

22 J. Rak, P. Skurski, M. Gutowski, L. Jozwiak and J. Blazejowski, *J. Phys. Chem.*, 1997, **101**, 283.

23 M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. A. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Latham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andrés, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. J. P. Stewart, M. Head-Gordon, C. González and J. A. Pople, *GAUSSIAN94*, Revision C.3, Gaussian, Inc., Pittsburgh, PA, 1995.

24 W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, *Ab initio Molecular Orbital Theory*, Wiley, New York, 1986.

25 (*a*) M. J. Frisch, J. S. Binkley and J. A. Pople, *J. Chem. Phys.*, 1984, **80**, 3265; (*b*) A. D. McLean and G. S. Chandler, *J. Chem. Phys.*, 1980, **72**, 5639; (*c*) R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 650.

26 J. Urban, P. R. Schreiner, G. Vacek, P. v. R. Schleyer, J. Q. Huang and J. Leszczynsky, *Chem. Phys. Lett.*, 1997, **264**, 441.

27 (*a*) R. Krishnan and J. A. Pople, *Int. J. Quantum Chem.*, 1978, **14**, 91; (*b*) R. Krishnan, M. J. Frisch and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 4244.

28 M. W. Wong, M. J. Frisch and K. B. Wiberg, *J. Am. Chem. Soc.*, 1991, **113**, 4776.

29 M. W. Wong, K. B. Wiberg and M. J. Frisch, *J. Am. Chem. Soc.*, 1992, **114**, 523.

30 M. W. Wong, K. B. Wiberg and M. J. Frisch, *J. Am. Chem. Soc.*, 1992, **114**, 1645.

31 (*a*) S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117; (*b*) J. Tomasi, R. Bonaccorsi, R. Cammi and F. O. Valle, *J. Mol. Struct.*, 1991, **234**, 401; (*c*) J. Tomasi and M. Persico, *Chem. Rev.*, 1994, **94**, 2027.

32 *CRC Handbook of Chemistry and Physics*, ed. D. R. Lide and H. P. R. Frederikse, 75th edn., CRC Press, Boca Raton, FL, 1994.

Paper 8/08818G