Theoretical Study of the Tautomerism and Protonation of 7-Aminopyrazolopyrimidine in the Gas Phase and in Aqueous Solution

Modesto Orozco*,† and F. J. Luque*,‡

Contribution from the Departament de Bioquímica, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain, and Departament de Farmàcia, Unitat Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal sn, Barcelona 08028, Spain

Received August 10, 1994®

Abstract: The tautomerism and protonation of 7-aminopyrazolopyrimidine in the gas phase are studied by means of *ab initio* methods. The effect of the solvent on the tautomerization and protonation processes is accounted for by using several high-level techniques including molecular dynamics−free energy perturbation (MD−FEP), an optimized *ab initio* self-consistent reaction field (SCRF) method, and two different semiempirical SCRF methods. The results not only provide a complete picture of the tautomerism and protonation of 7-aminopyrazolopyrimidine but also allow us to compare different "state-of-the-art" techniques to represent solvation effects. The biochemical and pharmacological importance of the tautomerism and protonation of 7-aminopyrazolopyrimidine is discussed.

Introduction

7-Aminopyrazolo[4,3-*d*]pyrimidine is an aromatic structure present in a series of *C*-nucleosides (see Figure 1), among which formycin A (7-amino-3β-D-ribofuranosylpyrazolo[4,3-*d*]-pyrimidine) is of particular interest due to its biological and pharmacological properties. Formycin A was first isolated from cultures of *Norcardia interforma*,[1] and since then, a large amount of research effort has been devoted to the study of its properties. This nucleoside, whose structure closely resembles that of adenosine, has a potential therapeutic interest in a wide range of areas, which include antiviral,[2] antitumoral,[3] antibiotic,[1,4] immunossupressant,[3a] and antimetabolic activities.[2] Recent studies have demonstrated that the monophosphorylated derivative of formycin can be incorporated into a growing RNA as substitution of adenosine monophosphate.[5] This finding makes possible the expansion of the genetic code *via* incorporation of new hydrogen-bonding patterns.[5,6]

Owing to its similarity with adenosine, formycin can interact with physiological targets of adenosine like adenosine kinase,[3b] purine nucleoside phosphorylase,[4] T7 RNA polymerase,[5] and adenosine deaminase.[7] The interaction with adenosine deaminase (ADA, EC 3.5.4.4) is of particular relevance from biological and pharmacological points of view.[8] ADA, which plays a key role in purine metabolism, catalyzes the hydrolytic conversion of adenosine to inosine.[9] The enzyme has a broad specificity for substrates structurally related to adenosine.[9] The action of ADA on formycin A (Figure 2) leads to the formation of formycin B, whose pharmacological properties are generally less important.[10]

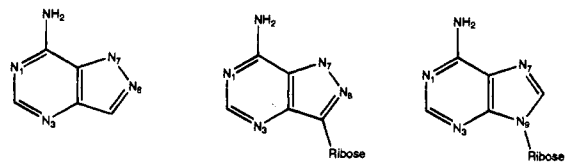


Figure 1. Schematic representation of 7-aminopyrazolopyrimidine, formycin A, and adenosine. The regular IUPAC name for the 7-aminopyrazolopyrimidine is used. However, atom numbering has been indicated according to the purine nomenclature system. This is preferable to using the specific IUPAC nomenclature for pyrazolopyrimidine compounds in order to clarify the discussion.

The interaction mechanism of a nucleoside with ADA implies the formation of hydrogen bonds between hydrogen bond donor groups of the enzyme and the lone pairs at N3 and N7 of the nucleoside.[11] The presence of a hydrogen atom at N3 or N7

* Address correspondence to either author.
† Departament de Bioquímica.
‡ Departament de Farmàcia.
® Abstract published in *Advance ACS Abstracts,* January 1, 1995.

(1) Hori, M.; Ito, E.; Takita, T.; Koyama, T.; Takeuchi, T.; Umezawa, A. *J. Antibiot., Ser. A* **1964**, *17*, 96.
(2) (a) Ishida, N.; Honiura, K.; Ksmegui, Y.; Schimizu, Y.; Matsumoto, S.; Izawa, A. *J. Antibiot., Ser. A* **1967**, *20*, 49. (b) Takeuchi, T.; Iwanaga, J.; Aogalad, T.; Umezawa, H. *J. Antibiot., Ser. A* **1966**, *19*, 286. (c) Giziewicz, J.; De Clerq, M.; Luczak, M.; Shugar, D. *Biochem. Pharmacol.* **1975**, *24*, 1813.
(3) (a) Odaka, T.; Takizawa, K.; Yamamura, K.; Yamamoto, T. *Jpn. J. Exp. Med.* **1975**, *39*, 327. (b) Suhaldonik, R. J. In *Nucloesides as Biological Probes;* Wiley: New York, 1979; Chapter 5.
(4) Bzowska, A.; Kulikowska, E.; Shugar, D. *Biochim. Biophys. Acta* **1992**, *1120*, 239.
(5) Piccirilli, J. A.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1991**, *30*, 10350.
(6) (a) Piccirilli, J. A.; Krauch, T.; Moroeny, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33. (b) Benner, S. A.; Ellington, A. D.; Tauer, A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7054.
(7) (a) Secrist, J. A.; Shortnacy, A. I.; Montgomery, J. A. *J. Med. Chem.* **1985**, *28*, 1740. (b) Wierzchowski, J.; Shugar, D. *Z. Naturforsch.* **1983**, *38*, 67. (c) Zemlicka, J. *J. Am. Chem. Soc.* **1975**, *97*, 5896.
(8) (a) Centelles, J.; Franco, R.; Bozal, J. *J. Neurosci. Res.* **1988**, *19*, 258. (b) Birnbaum, G. I.; Shugar, D. In *Topics in Nucleic Acids Structure*; Macmillan Press: London, 1987; Part III, Chapter 1. (c) Morisaki, T.; Fujii, H.; Miwa, S. *Am. J. Hematol.* **1985**, *19*, 37. (d) Murray, J. L.; Loftin, K. C.; Munn, C. G.; Reuber, J. M.; Mansell, P. W. A.; Hersh, E. M. *Blood* **1985**, *65*, 1318. (e) Giblett, E. R.; Anderson, J. E.; Cohen, F.; Pollara, B.; Meuwissen, H. J. *Lancet* **1972**, No. 2, 1067.
(9) (a) Wolfenden, R. *Biochemistry* **1969**, *8*, 2409. (b) Wolfenden, R.; Frick, L. *J. Protein Chem.* **1986**, *5*, 147. (c) Leonard, N. J.; Sprecker, M. A.; Morrice, A. G. *J. Am. Chem. Soc.* **1976**, *98*, 3987. (d) Simon, L. N.; Bauer, R.; Tolman, R. L.; Robins, R. K. *Biochemistry* **1970**, *9*, 573. (e) Mikhailopulo, I. A.; Wiedner, H.; Cramer, F. *Biochem. Pharmacol.* **1987**, *30*, 1001. (f) Hampton, A.; Harper, P. J.; Sasaki, T. *Biochemistry* **1972**, *11*, 4736. (g) Cory, J. G.; Suhadolnik, R. J. *Biochemistry* **1965**, *4*, 1729.
(10) McKenna, R.; Neidle, S.; Serafinowski, P. *Acta Crystallogr.* **1987**, *C43*, 2358.

0002-7863/95/1517-1378$09.00/0

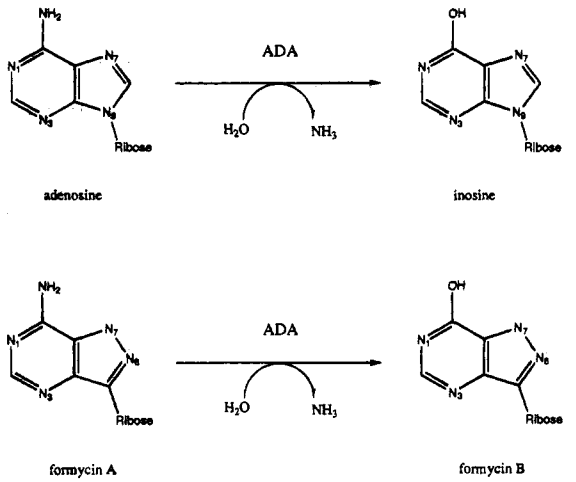


Figure 2. Deamination reaction of adenosine and formycin A.

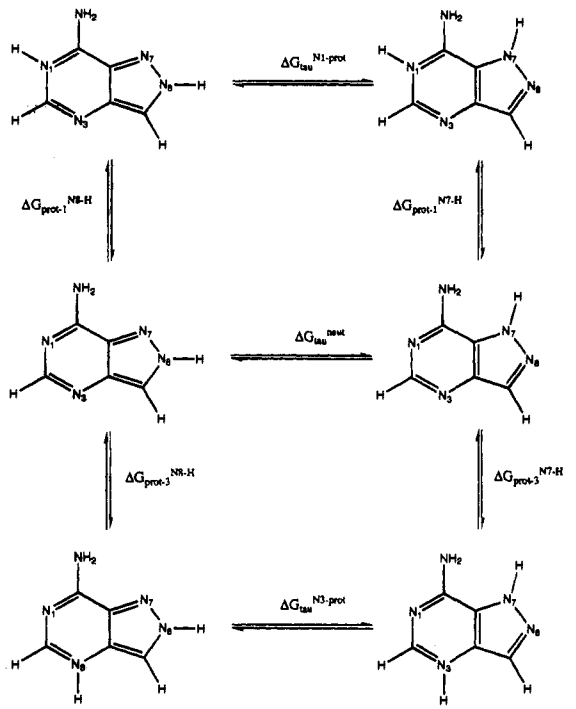


Figure 3. Schematic representation of the N7H → N8H tautomerism and N1 and N3 protonations of 7-aminopyrazolopyrimidine.

leads to the inhibition of the enzyme,[11a,12] as shown by the enzymatic behavior of 3- and 7-deazaadenosine.[12] This finding, in conjunction with the experimental evidence that formycin A is mainly in the N7H tautomeric form (Figure 3),[13] gives rise to a paradox: If formycin exists in the N7H tautomeric form, which is unable to interact with the enzyme, why is formycin recognized and deaminated very efficiently by ADA?

A plausible explanation for the susceptibility of formycin to ADA is based on the change of the tautomeric preference from the N7H to the N8H tautomer when it anchors to the active site of the enzyme.[11a,14] Previous studies within the quantum mechanical SCF framework in the gas phase suggested a possible relationship between the ionization state of the nucleoside and its tautomeric preference.[15] Thus, data derived from *ab initio* SCF calculations indicated that the protonation of formycin leads to a change in its tautomeric preference to the N8H tautomer, which can be recognized by the enzyme. Unfortunately, since at neutral pH the nucleoside is likely to be un-ionized, these *ab initio* results do not provide a fully satisfactory explanation to the paradox of the recognition of formycin by ADA under physiological conditions.

In this study, we explore the tautomerism of the aromatic moiety of formycin A, 7-aminopyrazolopyrimidine. *Ab initio* methods at the SCF, MP2, MP3, and MP4 levels are used to examine the tautomeric preference in the gas phase (Figure 3). Furthermore, *ab initio* self-consistent reaction field (SCRF), semiempirical SCRF, and molecular dynamics–free energy perturbation (MD–FEP) techniques are used to account for the influence of water on the tautomeric equilibrium in aqueous solution. In addition, due to the relevance of the ionization state on the tautomerism of 7-aminopyrazolopyrimidine,[13,15] and on the reaction mechanism of ADA,[9,11b,14,16] attention has also been paid to the protonation at N1 and N3 of the N7H and N8H tautomers of 7-aminopyrazolopyrimidine. The results provide a detailed picture of the tautomerism of this molecule and suggest interesting perspectives in the design of drugs related structurally to formycin A but resistant to the action of ADA.

Methods

Gas Phase Calculations. Equilibrium geometries for N7H and N8H tautomers of formycin in both neutral and protonated (at N1 or N3) forms were optimized at the SCF-RHF level using the 3-21G[17] and the 6-31G*[18] basis sets. Force constant analyses were carried out to verify the minimum state nature of the final geometries. Single-point calculations at the MP2, MP3, and MP4(SDTQ) levels using the 6-31G* basis set were performed to incorporate electron correlation effects. The energy obtained from SCF and Moller–Plesset calculations was corrected to obtain the Gibbs free energy (at 298 K) by the addition of thermal and entropic effects using standard procedures in Gaussian 90.[19] The free energy differences between the tautomers or the protonation states were computed from the thermodynamic cycle shown in Figure 3.

Solvation Calculations. Free energy calculations in aqueous solution were performed following the thermodynamic cycle shown in Figure 4, where it is implicitly assumed that the free energy change during a reaction can be computed as the addition of two independent terms: (i) the change in internal free energy (ΔG_{B-A}^{gas}), which is determined at the quantum mechanical level, and (ii) the change in the solute–solvent interaction ($\Delta\Delta G_{B-A}^{hyd}$), which is estimated from either SCRF or MD–FEP computations.

Self-Consistent Reaction Field Calculations. SCRF calculations based on semiempirical AM1[20] and *ab initio* 6-31G* wave functions

(11) (a) Orozco, M.; Canela, E. I.; Franco, R. *Eur. J. Biochem.* **1990**, *188*, 155. (b) Wilson, D. K.; Rudolp, F. B.; Quiocho, F. A. *Science* **1991**, *252*, 1278. (c) Sharff, A. J.; Wilson, D. K.; Chang, Z.; Quiocho, F. *J. Mol. Biol.* **1992**, *226*, 917.

(12) Ikehara, M.; Fukui, T. *Biochim. Biophys. Acta* **1974**, *338*, 512.

(13) (a) Wierzhowski, J.; Shugar, D. *Photochem. Photobiol.* **1982**, *35*, 445. (b) Krugh, T. R. *J. Am. Chem. Soc.* **1973**, *95*, 4736. (c) Chenon, M. T.; Panzica, R. P.; Smith, J. C.; Pugmire, R. J.; Grant, D. M.; Townsend, L. B. *J. Am. Chem. Soc.* **1976**, *98*, 4736. (d) Cho, B. P.; McGregor, M. A. *Nucleosides Nucleotides* **1994**, *13*, 481. (e) Dodin, G.; Bensaude, O.; Dubois, J. *J. Am. Chem. Soc.* **1980**, *102*, 3897. (f) Townsend, L. B.; Long, R. A.; McGrwa, J. P.; Miles, D. W.; Robins, R. K.; Eyring, H. *J. Org. Chem.* **1974**, *39*, 2023. (g) Prusiner, P.; Brennan, T.; Sundaralingam, M. *Biochemistry* **1973**, *12*, 1196.

(14) (a) Orozco, M.; Canela, E. I.; Franco, R. *Mol. Pharmacol.* **1989**, *35*, 257. (b) Orozco, M.; Lluis, C.; Mallol, J.; Canela, E. I.; Franco, R. *J. Pharm. Sci.* **1990**, *79*, 133.

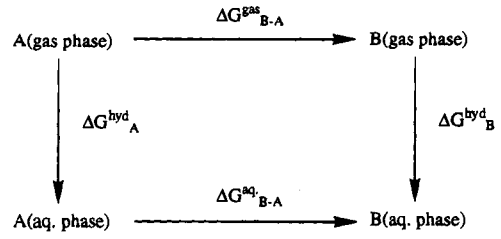
(15) Orozco, M.; Lluis, C.; Mallol, J.; Canela, E. I.; Franco, R. *J. Org. Chem.* **1990**, *55*, 753.

(16) (a) Weiss, P.; Cook, P. F.; Hermes, J. D.; Cleland, W. W. *Biochemistry* **1987**, *26*, 7378. (b) Wilson, D. K.; Quiocho, F. A. *Biochemistry* **1993**, *32*, 1689. (c) Kurz, L. C.; Molx, L.; Riley, M. C.; Frieden, C. *Biochemistry* **1992**, *31*, 39.

(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) 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. *Gaussian 90*, Revision F; Gaussian, Inc.: Pittsburgh, PA, 1990.



$$\Delta G^{aq}_{B-A} = \Delta G^{gas}_{B-A} + \Delta G^{hyd}_{B} - \Delta G^{hyd}_{A} = \Delta G^{gas}_{B-A} + \Delta\Delta G^{hyd}_{B-A}$$

Figure 4. Thermodynamic cycle used to compute free energy differences in aqueous solution.

were used. Semiempirical calculations were carried out following two procedures: (i) the Generalized Born Model (GBM[21]) as implemented in AMSOL using the AM1/SM2 parametrization[22] and (ii) the SCRF model formulated by Miertus, Scrocco, and Tomasi (MST[23]), which has recently been adapted to semiempirical Hamiltonians in our laboratory (AM1/MST).[24] *Ab initio* SCRF calculations were performed using a 6-31G*-optimized version[25] of the MST method.[22]

Both AMSOL and MST methods compute the hydration free energy as the addition of an electrostatic component and a steric contribution. The electrostatic term represents the interaction between the charge distribution of the solute and the reaction field generated by the solute in the solvent. The steric contribution includes several energy terms like cavitation, dispersion, repulsion, and volume effects. The two methods use a cavity adapted to the molecular shape to represent the solute−solvent interface. AMSOL determines all the non-electrostatic interactions by means of a simple relationship with the molecular surface area.[22] In contrast, the MST method computes separately the cavitation component following Pierotti's scaled particle theory,[26] while the remaining contributions are represented in our version of the method by means of a relationship with the molecular surface area.[24,25]

Besides the difference in the procedure used to evaluate the steric term, the main difference between AMSOL and MST lies in the procedure used for the computation of the electrostatic contribution to the free energy of hydration. AMSOL is based on the GBM model[21] and on the representation of the molecular charge distribution from Mulliken charges.[27] The MST method computes the electrostatic contribution by means of a set of imaginary charges, which are determined by solving the Laplace equation at the solute−solvent interface. In this case, the solute charge distribution is rigorously represented at the SCF level by means of the molecular electrostatic potential (MEP).[28] These brief considerations suggest that the MST formalism must provide a more accurate evaluation of the electrostatic contribution to the free energy of hydration than GBM-based methods (see refs 21−25 for a detailed description of the two methods). Nevertheless, this formal superiority of MST methods might be compensated by the empirical parametrization of AMSOL (six parameters fitted for each atom, as opposed to the limited set of parameters fitted in MST calculations).

AMSOL-AM1/SM2 calculations were performed following the standard protocol and using the optimized parameters.[22] The molecular geometries optimized both in the gas phase and in aqueous solution were considered in the calculations. AM1/MST calculations were carried out according to the standard protocol,[24] and the orthogonal procedure was used to calculate the semiempirical MEP.[29] Computations were performed using the gas phase-optimized geometry, as well as the AMSOL-optimized geometry in aqueous solution. *Ab initio* SCRF calculations were performed using a 6-31G* optimized version of the MST model.[23,25] Since the change of the molecular geometry upon solvation has a negligible effect on the thermodynamic parameters (see below), only the gas phase-optimized geometry was used. Following our previous findings,[30] the cavity was uniformly reduced by a factor of 92% in all the MST calculations of the protonated forms of 7-aminopyrazolopyrimidine.

Free Energy Perturbation Simulations. FEP techniques[31] compute $\Delta\Delta G_{hyd}^{B-A}$ directly (in contrast to SCRF methods, which determine this magnitude as the difference of the hydration free energy for each tautomer). This is done by the calculation of the work necessary to mutate tautomer A into tautomer B along a reversible pathway. In practice, this is achieved by changing the parameters defining the solute−solvent interactions (charges and Lennard-Jones parameters) from the values of tautomer A into those of tautomer B, as shown in eq 1, where λ is the coupling parameter varied from 0 to 1 during the mutation.

$$V_\lambda = \lambda V_{\lambda=1} + (1-\lambda)V_{\lambda=0} \tag{1}$$

Two mutation procedures were used in this study: (i) the standard windowing scheme and (ii) a finite difference thermodynamic integration (FDTI) method[32] based on the Gaussian quadrature technique.[33] In the first case, the mutation A → B was performed in a large number of small steps (windows), and the total free energy was determined as the addition of the free energy in each window (eq 2).

$$\Delta\Delta G_{hyd}^{B-A} = \sum_{\lambda=0}^{\lambda=1-\Delta\lambda} -RT \ln \left\langle e^{\frac{-(V_{\lambda+\Delta\lambda} - V_\lambda)}{RT}} \right\rangle_\lambda \tag{2}$$

FDTI calculations were done following eq 3. The numerical derivative of the free energy with regard to λ was evaluated by using eq 2, with a very small $\Delta\lambda$, and the definition of the potential functions were those shown in eq 1. The points (λ) selected to evaluate the free energy derivative were determined according to a Gaussian quadrature procedure.

$$\Delta\Delta G_{hyd}^{\beta-A} = \int_0^1 \frac{\Delta G(\lambda,\Delta\lambda)}{\Delta\lambda}\, d\lambda \tag{3}$$

The simulation system was a cube of ~18.000 Å^3 containing 1 solute plus around 550 TIP3P[34] water molecules. Periodic boundary conditions were used in conjunction with a 10 Å cutoff for nonbonded interactions. Simulations were performed at the isothermic−isobaric (NPT) ensemble (1 atm; 298 K) using MD algorithms.[35a] All the bonds were fixed at their equilibrium values using SHAKE,[35b] which allows the use of a 2 fs integration step. The system was initially optimized for 500 cycles of the steepest descent and then heated and equilibrated for 20 ps. Simulations were typically performed for 200−210 ps, but in some cases, 420 ps simulations were carried out. The perturbation was usually done starting from the N7H tautomer. However, in other cases, the N8H tautomer was used as the starting structure in order to assess whether the FEP calculation provided a zero value for the futile process A → B → A.

Simulations were carried out using 10 ps windows, which were divided into two parts: 4 ps of equilibration plus 6 ps of averaging. The 420 ps perturbation was done starting from equilibrated samplings

(20) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(21) Still, W. C.; Tempczak, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127 and references therein.
(22) Cramer, C. J.; Truhlar, D. *Science* **1992**, *256*, 213.
(23) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117. (b) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239.
(24) (a) Negre, M.; Orozco, M.; Luque, F. J. *Chem. Phys.* **1992**, *196*, 27. (b) Luque, F. J.; Negre, M.; Orozco, M. *J. Phys. Chem.* **1993**, *97*, 4386. (c) Luque, F. J.; Bachs, M.; Orozco, M. *J. Comput. Chem.* **1994**, *15*, 847. (d) Orozco, M.; Bachs, M.; Luque, F. J. *J. Comput. Chem.*, in press, 1994.
(25) Bachs, M.; Luque, F. J.; Orozco, M. *J. Comput. Chem.* **1994**, *15*, 446.
(26) Pierotti, R. A. *Chem. Rev.* **1976**, *76*, 717.
(27) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
(28) Scrocco, E.; Tomasi, J. *Top. Curr. Chem.* **1973**, *42*, 95.
(29) Alhambra, C.; Luque, F. J.; Orozco, M. *J. Comput. Chem.* **1994**, *15*, 12.
(30) Orozco, M.; Luque, F. J. *Chem. Phys.* **1994**, *182*, 237.
(31) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
(32) Mezei, M. *J. Chem. Phys.* **1985**, *86*, 7084.
(33) Press, W. H.; Flannery, B. P.; Tenkolsky, S. A.; Vetterling, W. T. *Numerical Recipes;* Cambridge University Press. New York, 1992; Chapter 4.
(34) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 296.
(35) Berendsen, H. J. C.; Postma, P. P. M.; van Gunsteren, V. W.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

of the N8H tautomer of the neutral 7-aminopyrazolopyrimidine. The simulation was divided into two parts. First, charges for the N8H tautomer were mutated in 210 ps to those corresponding to the N7H tautomer. Then, a hydrogen atom was annihilated at N8, and simultaneously generated at N7, in the remaining 210 ps. In both simulations, 21 windows were considered (10 ps window). The 210 ps simulation was done starting from the N7H tautomer, which was obtained from the previous FEP plus 20 ps of additional MD for equilibration. In this simulation, charges and van der Waals parameters were changed simultaneously considering 21 windows (10 ps each window).

FDTI simulations were done by dividing the total perturbation into two parts: $\lambda = 1-0.5$ and $\lambda = 0.5-0.0$. This was done to achieve a better representation of the extremes ($\lambda = 0$ and $\lambda = 1$) and of the maximum in the free energy profile, which was found to be around $\lambda = 0.5$ from windowing simulations. A 10-point Gaussian quadrature method was used to select representative points and weighting values for integration in the two parts of the perturbation (abcises and weights are available upon request), which yields a total of 20 points. Numerical derivatives ($\Delta G/\Delta\lambda$) were determined using eq 2, with $\Delta\lambda = 0.001$. The total length of the perturbation was 200 ps. Evaluation of $\Delta G/\Delta\lambda$ at every λ was performed for 5 ps, after 5 ps of equilibration. The double-wide sampling protocol was used in both windowing and FDTI calculations.

Atomic charges were determined (charges are available upon request) by fitting SCF and Coulombic MEPs in four Connolly's layers placed at 1.4, 1.6, 1.8, and 2.0 times the van der Waals radii of atoms. A density of 10 points/Å^2 was used to guarantee the quality of the fitting.[36] It is worth noting that reasonable estimates of hydration free energies for neutral molecules are obtained when 6-31G* electrostatic charges are used in MD−FEP and Monte Carlo−FEP techniques.[37] The rest of the force field parameters were taken from the AMBER all atom force field.[38]

Ab initio calculations in the gas phase were performed using Gaussian 90.[19] *Ab initio* SCRF calculations were carried out with a modified version of MonsterGauss,[39] which includes the MST method. AM1/MST calculations were performed with MOPAC93, Revision 2,[40a] which incorporates the optimized AM1 version of the MST model. AMSOL-AM1/SM2 calculations were done with the standard code developed by Cramer and Thrular.[22] MEP calculations and charge fittings were carried out using the MOPETE/MOPFIT computer programs.[40b] Finally, MD−FEP simulations were performed with the AMBER 4.0 computer program.[38] All the calculations were done on the Cray Y-MP of the Centre de Supercomputació de Catalunya (CESCA) and on workstations in our laboratory.

RESULTS

Gas Phase Simulations. The energy, enthalpy, and free energy differences between the N7H and N8H tautomers of the neutral and N1- and N3-protonated species of 7-aminopyrazolopyrimidine calculated at different levels of theory are shown in Table 1. Comparison of the results allows us to examine the dependence of the tautomeric equilibrium on the quality of the quantum mechanical calculation.

Table 1. Energy, Enthalpy, and Free Energy Differences (in kcal/mol) between the N7H and N8H Tautomers of 7-Aminopyrazolopyrimidine[a]

method	ΔE	ΔH	$-T\Delta S$	ΔG
		Neutral Form		
3-21G/3-21G	−2.1	−2.6	−0.2	−2.8
6-31G*/3-21G	−0.4	−0.9	−0.2	−1.1
6-31G*/6-31G*	−1.1	−1.4	0.1	−1.3
MP2/6-31G*	1.6	1.3	0.1	1.4
MP4/6-31G*	1.2	0.9	0.1	1.0
		N1-Protonated		
3-21G/3-21G	8.3	7.8	−0.2	7.6
6-31G*/3-21G	8.1	7.6	−0.2	7.4
6-31G*/6-31G*	8.2	7.5	−0.4	7.1
MP2/6-31G*	9.3	8.6	−0.4	8.2
MP4/6-31G*	8.8	8.1	−0.4	7.7
		N3-Protonated		
3-21G/3-21G	4.8	4.4	−0.1	4.3
6-31G*/3-21G	5.3	4.9	−0.1	4.8
6-31G*/6-31G*	5.2	4.7	−0.2	4.5
MP2/6-31G*	5.6	5.1	−0.2	4.9
MP4/6-31G*	5.4	4.9	−0.2	4.7

[a] Thermal and entropic corrections to the energy were determined at 298 K. Values are computed as [N7H] − [N8H].

For the neutral species, all the SCF results suggest that the N7H tautomer is the preferred form in the gas phase both in terms of energy and free energy. The difference shrinks slightly as the basis set is enlarged from 3-21G to 6-31G*, but even at the 6-31G* level, the N7H tautomer is favored by around 1 kcal/mol. Geometry reoptimization at the 6-31G* level leads to negligible changes in the final results with respect to the values obtained using the geometry optimized at the 3-21G level. In contrast, electron correlation effects (computed using 6-31G* wave functions and 3-21G-optimized geometries) have a dramatic influence on the tautomeric preference, since the N8H tautomer is stabilized by 1.2 kcal/mol in energy at the highest level of theory, which is reduced to 1.0 kcal/mol when thermal and entropic corrections are considered. This result, which agrees with previous findings on the tautomerism of similar compounds,[41,42] demonstrates the need to account for the correlation effects to obtain reliable representations of tautomeric processes.

Protonation at N1 or N3 leads to a great stabilization of the N8H tautomer in the gas phase. The effect is larger for the N1-protonated form, where all the calculations estimate the free energy difference of tautomerization to be around 7.1−8.2 kcal/mol (7.7 kcal/mol at the highest level). Such a difference amounts to 4.3−4.9 kcal/mol (4.7 kcal/mol at MP4/6-31G*) for the N3-protonated species. As noted before, geometry reoptimization has a negligible effect on the results. Furthermore, the introduction of correlation effects, which are essential for the neutral form, has a very small effect on the protonated species.

The ability of the N7H and N8H tautomers of 7-aminopyrazolopyrimidine to be protonated in the gas phase at N1 and N3 can be examined from data in Table 2. All the methods suggest that the N3-protonated form is clearly favored in the gas phase. The preference for the protonation at N3 is around 3 kcal/mol smaller for the N8H tautomer with respect to the N7H tautomer, but even in this case, the N3-protonated species is preferred by more than 4 kcal/mol.

SCRF Calculations in Water. SCRF values of the hydration free energy difference ($\Delta\Delta G_{hyd}$) between the N7H and N8H

(36) (a) Connolly, M. *QCPE Bull.* **1981**, *1*, 175. (b) Orozco, M.; Luque, F. J. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 411.

(37) (a) Orozco, M.; Jorgensen, W. L.; Luque, F. J. *J. Comput. Chem.* **1993**, *14*, 1498. (b) Carlson, H.; Nguyen, T. B.; Orozco, M.; Jorgensen, W. L. *J. Comput. Chem.* **1993**, *14*, 1240.

(38) (a) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. *J. Comput. Chem.* **1986**, *7*, 230. (b) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. *AMBER 4.0*, Revision A; University of California: San Francisco, CA, 1991.

(39) Peterson, M.; Poirier, R. *MonsterGauss Computer Program*; Department Chemistry, University of Toronto: Toronto, Canada. Modified by R. Cammi, R. Bonaccorsi, and J. Tomasi, University of Pisa, Pisa, Italy, 1987. Modified by F. J. Luque and M. Orozco, University of Barcelona, Barcelona, Spain, 1994.

(40) (a) Stewart, J. J. P. *MOPAC93*, Revision 2; Stewart Comp. Chem., 1994. This program contains the newest versions of the CMS routines to perform MST calculations developed by F. J. Luque and M. Orozco. (b) Luque, F. J.; Orozco, M. *MOPETE/MOPFIT*; University of Barcelona: Barcelona, Spain, 1994.

(41) Cieplak, P.; Bash, P.; Singh, U. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 683.

(42) Tomas, F.; Catalán, J.; Pérez, P.; Elguero, J. *J. Org. Chem.* **1994**, *59*, 2799.

Table 2. Energy, Enthalpy, and Free Energy Differences (in kcal/mol) between the N1- and N3-Protonated Species of the N7H and N8H Tautomers of 7-Aminopyrazolopyrimidine[a]

method	ΔE	ΔH	$-T\Delta S$	ΔG
	N7H Tautomer			
3-21G/3-21G	9.3	8.8	−0.1	8.7
3-21G/6-31G*	9.2	8.7	−0.1	8.6
6-31G*/6-31G*	9.1	8.7	−0.2	8.5
6-31G*-MP2	8.6	8.2	−0.2	8.0
6-31G*-MP4	8.5	8.1	−0.2	7.9
	N8H Tautomer			
3-21G/3-21G	5.4	5.1	0.0	5.1
3-21G/6-31G*	6.4	6.1	0.0	6.1
6-31G*/6-31G*	6.1	5.9	0.0	5.9
6-31G*-MP2	4.9	4.7	0.0	4.7
6-31G*-MP4	5.0	4.8	0.0	4.8

[a] Thermal and entropic corrections to the energy were determined at 298 K. Values are computed as [N1-protonated] − [N3-protonated].

Table 3. SCRF Values of the Hydration Free Energy Differences ($\Delta\Delta G_{hyd}$) between the N7H and N8H Tautomers of 7-Aminopyrazolopyrimidine[a]

method	geometry	$\Delta\Delta G_{hyd}$	ΔG^{taut}
	Neutral Form		
AMSOL/SM2[b]	gas	3.6	4.6
AMSOL/SM2	solution	3.7	4.7
AM1/MST[b]	gas	0.4	1.4
AM1/MST	solution	0.3	1.3
6-31G*/MST	gas	−0.3	0.7
	N1-Protonated		
AMSOL/SM2	gas	−2.5	5.2
AMSOL/SM2	solution	−3.1	4.6
AM1/MST	gas	−6.0	1.7
AM1/MST	solution	−6.6	1.1
6-31G*/MST	gas	−6.4	1.3
	N3-Protonated		
AMSOL/SM2	gas	−1.0	3.7
AMSOL/SM2	solution	−1.0	3.7
AM1/MST	gas	−4.2	0.5
AM1/MST	solution	−4.6	0.1
6-31G*/MST	gas	−4.1	0.6

[a] The free energy of tautomerization (ΔG^{taut}) in aqueous solution was determined using the MP4/6-31G* results in gas phase (see text). Values (in kcal/mol) are computed as [N7H] − [N8H]. [b] See ref 43.

tautomers of neutral and protonated 7-aminopyrazolopyrimidine are shown in Table 3, which also gives the free energy difference of tautomerization (ΔG^{taut}) at 298 K in aqueous solution (Figure 3).

The results in Table 3 are evidence for the reduced effect of the molecular geometrical changes induced by water on the $\Delta\Delta G_{hyd}$ between the tautomers of the neutral form. Thus, both AMSOL and mixed AM1/MST//AMSOL calculations[43] suggest that the changes in $\Delta\Delta G_{hyd}$ due to the geometry relaxation upon solvation are very small (±0.1 kcal/mol). For the N3-protonated molecule, the solvent-induced geometry relaxation also has a negligible influence on the ratio between tautomers, and for the N1-protonated form, such an effect is around 0.6 kcal/mol, which is a small fraction compared to the total hydration free energy of the protonated tautomers (around −70 kcal/mol). The changes in geometry induced by water are slightly more important when the relative stability of the N1- and N3-protonated forms is discussed (see below and Table 4), but even in this case, such a contribution is only around 1 kcal/mol.

(43) The molecule is forced to be planar in the AM1 gas phase optimization to avoid artefactual deformations of the aromatic ring. For comparison, the $\Delta\Delta G_{hyd}$ values obtained for the real AM1 minima are +5.8 kcal/mol (AMSOL) and +1.6 kcal/mol (AM1/MST). When the 6-31G*-optimized geometries are used, the values are +4.4 kcal/mol (AMSOL) and +1.2 kcal/mol (AM1/MST).

Table 4. SCRF Values of the Hydration Free Energy Differences ($\Delta\Delta G_{hyd}$) between the N1- and N3-Protonated Species of 7-Aminopyrazolopyrimidine[a]

method	geometry	$\Delta\Delta G_{hyd}$	ΔG^{prot}
	N7H Tautomer		
AMSOL/SM2	gas	−3.6	4.2
AMSOL/SM2	solution	−5.1	2.7
AM1/MST	gas	−6.3	1.5
AM1/MST	solution	−7.3	0.5
6-31G*/MST	gas	−6.0	1.8
	N8H Tautomer		
AMSOL/SM2	gas	−2.0	2.8
AMSOL/SM2	solution	−3.0	1.8
AM1/MST	gas	−4.4	0.4
AM1/MST	solution	−5.4	−0.6
6-31G*/MST	gas	−3.7	1.1

[a] The free energy of protonation (ΔG^{prot}) in aqueous solution was determined using the MP4/6-31G* results in gas phase (see text). Values (in kcal/mol) are computed as [N1-protonated] − [N3-protonated].

Therefore, keeping in mind the expensiveness of the geometry optimization in aqueous solution, the fact that MST methods were parametrized to fit experimental data considering gas phase geometries, and the reduced effect of the geometry relaxation by water to determine the $\Delta\Delta G_{hyd}$ (within the expected limits of accuracy of the models), the use of gas phase-optimized geometries in *ab initio* MST calculations seems fully justified.

Inspection of the results in Table 3 shows that in general the AM1 and 6-31G* versions of the MST method are in good agreement (differences in $\Delta\Delta G_{hyd}$ in the range 0.1−0.7 kcal/mol). In contrast, there are notable discrepancies with AMSOL, which gives a larger $\Delta\Delta G_{hyd}$ for the neutral species and a smaller value for the protonated forms with regard to the 6-31G*/MST results. All the SCRF methods predict a preferential hydration of the N7H tautomer for protonated forms of the molecule. Inspection of Table 3 shows that the $\Delta\Delta G_{hyd}$ values between the N7H and N8H tautomers are 1.5−2.3 kcal/mol larger for the N1-protonated species than for the N3-protonated species. Discrepancies exist about the hydration of the two tautomers of the neutral molecule since AMSOL predicts the hydration of the N8H tautomer to be 3.6−3.7 kcal/mol more favored, AM1/MST reduces such a difference to only 0.4 kcal/mol, and the 6-31G*/MST method predicts that the N7H tautomer is slightly better hydrated than the N8H tautomer.

Combination of the MP4/6-31G* gas phase free energy difference with the $\Delta\Delta G_{hyd}$ values allowed us to compute the N7H ↔ N8H tautomerization free energies of the neutral and protonated forms. For the neutral species, AMSOL predicts the N8H tautomer to be favored in water by more than 4 kcal/mol with respect to the N7H tautomer. This difference decreases to around 1.3 kcal/mol at the AM1/MST level, while the 6-31G*/MST method suggests a difference of only 0.7 kcal/mol favoring the N8H tautomer. All the SCRF calculations point out a reduction in the free energy difference between the N7H and N8H tautomers for the protonated 7-aminopyrazolopyrimidine upon solvation, but the N8H form is still the preferred tautomer in aqueous solution by 1−2 kcal/mol (N1-protonated) and around 0.5 kcal/mol (N3-protonated), this preference being larger (around 4 kcal/mol for the N3- and N1-protonated forms, respectively) according to AMSOL values.

The effect of water on the protonation of 7-aminopyrazolopyrimidine is shown in Table 4. All the methods predict that the N1-protonated form is better hydrated for both tautomers. Inspection of the free energy difference between the N1- and N3-protonated species shows no qualitative change in the preferential protonation of the N7H tautomer upon solvation,

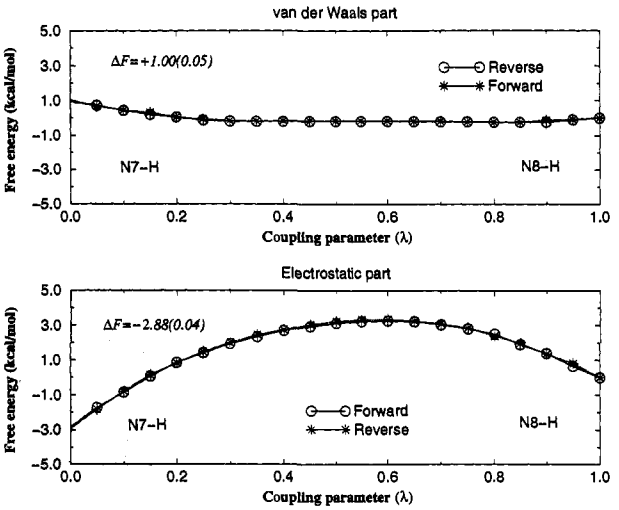


Figure 5. Electrostatic and van der Waals free energy profiles for the mutation N8H($\lambda = 1$) ↔ N7H($\lambda = 0$). Simulation was done using the windowing scheme. Simulation time was 420 ps.

for which around 95% of the molecules are protonated at N3 in water, according to the 6-31G*/MST value. In contrast, the results derived from MST calculations suggest that the protonation at N1 can be important for the N8H tautomer. Thus, for this tautomer, the population of the N1-protonated species is around 15% of the N3-protonated form according to the 6-31G*/MST results. The AM1/MST value suggests similar stability for the two protonated species, while AMSOL calculations clearly favor the N3-protonated form.

From the results in Tables 3 and 4, the MST method provides the following ordering (determined from values computed using gas phase geometries) of stability in water for the protonated 7-aminopyrazolopyrimidine species: N3-protonated (N8H) > N3-protonated (N7H) ≥ N1-protonated (N8H) > N1-protonated (N7H). The AMSOL-predicted ordering is N3-protonated (N8H) > N1-protonated (N8H) > N3-protonated (N7H) > N1-protonated (N7H). The use of AMSOL geometry in AM1/MST calculations gives a different ordering, since the N1-protonated (N8H) form is suggested to be the most stable protonated form.

Free Energy Perturbation Calculations. MD−FEP techniques were used to gain further insight into the solvent effect on the tautomerism and protonation of 7-aminopyrazolopyrimidine. For this purpose, N7H ↔ N8H mutations were done for the neutral and protonated forms of the molecule. In addition, for the two tautomers (N7H and N8H), the N1-protonated species was mutated into the N3-protonated ones to examine the protonation process.

Prior to the systematic computation of the free energy differences between tautomers and protonated species, an analysis of the optimum mutation protocol was performed. For this purpose, the N8H ↔ N7H mutation of the neutral molecule was studied by using different simulation protocols. First, a windowing scheme with electrostatic decoupling was used in a 420 ps simulation, where the starting samplings were those of the N8H tautomer (see Methods). The free energy profiles (see Figure 5) were smooth, and no discontinuity was observed. Indeed, hysteresis was negligible, as noted in the agreement between "forward" and "backward" values obtained using the double-wide sampling protocol. The total ΔG for the N8H → N7H perturbation was −1.88(±0.06) kcal/mol, which indicates that the N7H tautomer is better solvated than the N8H tautomer.

With this 420 ps MD−FEP simulation as a reference, a 210 ps simulation was performed within the windowing scheme, but without electrostatic decoupling. In this case, the starting

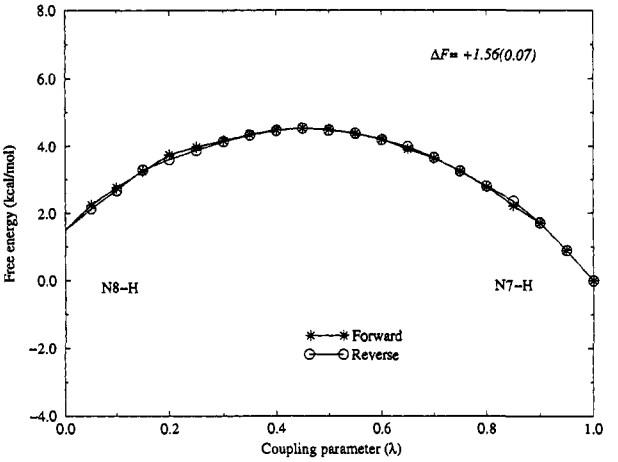


Figure 6. Free energy profile for the mutation N7H($\lambda = 1$) ↔ N8H($\lambda = 0$). Simulation was done using the windowing scheme. Simulation time was 210 ps.

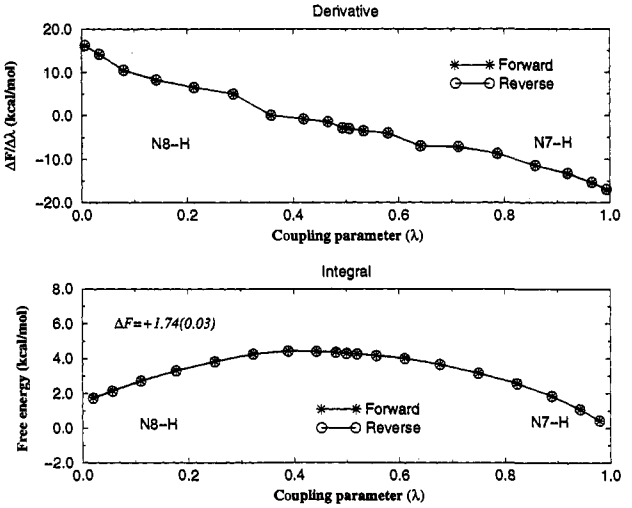


Figure 7. $\Delta G/\Delta\lambda$ and free energy profile for the mutation N7H($\lambda = 1$) ↔ N8H($\lambda = 0$). Simulation was done using the FDTI procedure. Simulation time was 200 ps.

samplings were those of the neutral N7H tautomer. The free energy profile for the N7H ↔ N8H mutation (see Figure 6) was also smooth, without major discontinuities, and showed negligible hysteresis. This demonstrates that the electrostatic decoupling was unnecessary in this case. The value of the N7H → N8H mutation was +1.56(±0.07) kcal/mol, and accordingly, the cycle N8H → N7H → N8H was closed with a moderate error of ±0.3 kcal/mol.

FDTI calculations on the neutral form (see Methods) were done during 200 ps without electrostatic decoupling. The starting samplings were those of the N7H tautomer. Owing to the existence of a maximum in the middle of the perturbation profile (see Figures 5 and 6), the perturbation was divided into two parts ($\lambda = 1-0.5$; $\lambda = 0.5-0.0$) to increase the statistical significance of the λ values selected by the Gaussian quadrature method. The $\Delta G/\Delta\lambda$ profile and the integrated free energy profile are shown in Figure 7. The total ΔG for the N7H → N8H mutation was +1.74(±0.03) kcal/mol. This value yields to a minimum error in the cycle N8H → N7H → N8H of only ±0.14 kcal/mol and improves the free energy difference estimated from the standard 210 ps windowing protocol. Indeed, the FDTI simulations have additional advantages: (i) the values of λ with an inadequate sampling can be detected easily in FDTI computations, (ii) the Gaussian quadrature method guarantees that the computational effort is mainly

Table 5. FEP Values of the Hydration Free Energy Differences ($\Delta\Delta G_{hyd}$) between the N7H and N8H Tautomers of Neutral and Protonated Species of 7-Aminopyrazolopyrimidine[a]

form	$\Delta\Delta G_{hyd}$	ΔG^{taut}
neutral	−1.7	−0.7
N1-protonated	−5.0	2.7
N3-protonated	−4.5	0.2

[a] The free energy of tautomerization (ΔG^{taut}) in aqueous solution was determined using the MP4/6-31G* results in gas phase (see text). Values (in kcal/mol) are computed as [N7H] − [N8H].

Table 6. FEP Values of the Hydration Free Energy Differences ($\Delta\Delta G_{hyd}$) between the N1- and N3-Protonated Species of 7-Aminopyrazolopyrimidine[a]

form	$\Delta\Delta G_{hyd}$	ΔG^{prot}
N7H tautomer	−4.6	3.2
N8H tautomer	−3.7	1.1

[a] The free energy of protonation (ΔG^{prot}) in aqueous solution was determined using MP4/6-31G* results in gas phase (see text). Values (in kcal/mol) are computed as [N1-protonated] − [N3-protonated].

focused on the key part of the simulation, and (iii) calculations at a given λ are independent with respect to calculations at previous values of λ, which avoids problems due to the error accumulation arising from incomplete equilibrations in successive windows. Therefore, considering the apparent superiority of the FDTI protocol with respect to the standard windowing scheme, all the remaining FEP simulations were performed using the FDTI protocol.

The results for the N7H ↔ N8H mutations are summarized in Table 5 (for the neutral form, the $\Delta\Delta G_{hyd}$ value in the table is the average of the three simulations noted above). FEP simulations suggest that the N7H tautomer is better solvated than the N8H tautomer when the neutral species is considered, in qualitative agreement with the 6-31G*/MST result and in disagreement with AM1/MST and especially with AMSOL calculations. The MD−FEP $\Delta\Delta G_{hyd}$ estimate for the neutral form, which is larger (in absolute value) than the 6-31G*/MST result, causes a change in the predicted tautomeric preference of 7-aminopyrazolopyrimidine. Thus, combination of the MP4/6-31G* gas phase free energy difference and the MD−FEP $\Delta\Delta G_{hyd}$ value points out that for the neutral species, the N7H tautomer is around 0.7 kcal/mol more stable than the N8H tautomer in aqueous solution (0.9 kcal/mol more stable if the 420 ps perturbation is considered).

FEP simulations agree well with SCRF calculations in the description of the stabilizing effect of water on the N7H tautomer of both N1-protonated and N3-protonated forms. The $\Delta\Delta G_{hyd}$ values are −5.0 and −4.5 kcal/mol for the N1- and N3-protonated forms, respectively, which compare well with the 6-31G*/MST results (−6.4 and −4.1 kcal/mol, respectively) and with the AM1/MST values (−6.0 and −4.2 kcal/mol, respectively). Combination of gas phase MP4/6-31G* tautomerization free energies with MD−FEP values shows that the N8H tautomer is the most stable tautomer for the N1-protonated species in water. In contrast, for the N3-protonated molecule, the N7H tautomer is only 0.2 kcal/mol less stable than the N8H tautomer when the effect of hydration is considered.

The effect of solvation on the protonation of 7-aminopyrazolopyrimidine was explored from mutations between the N1- and N3-protonated forms for both N7H and N8H tautomers (Figure 3). The results (see Table 6) confirm the SCRF-based suggestion that the N1-protonated species are better hydrated than the N3-protonated ones. The solvent effect is around 1 kcal/mol larger for the N7H tautomer than for the N8H tautomer,

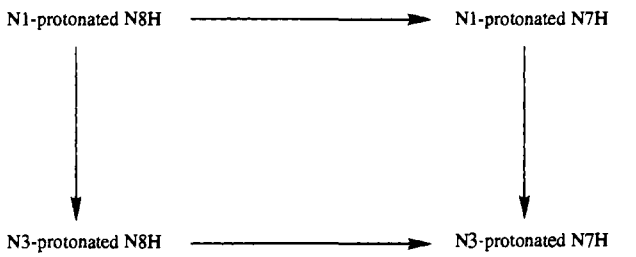


Figure 8. Thermodynamic cycle used to determine possible errors in FEP calculations of differences in hydration free energies (see text).

in reasonable agreement with SCRF results (Table 4). FEP simulations do not predict a change in the preference for the protonation at N3 of the N7H and N8H tautomers. However, it should be noted that solvation leads to a notable reduction in the difference of stability between the N1- and N3-protonated forms. The FEP estimate for ΔG^{prot} (1.1 kcal/mol) is in agreement with the 6-31G*/MST result.

Further information on the tautomerism and protonation of 7-aminopyrazolopyrimidine can be derived from combination of data in Tables 5 and 6. First, an estimate of the errors in FEP results due to the simulation protocol can be obtained from the cycle shown in Figure 8. Following this cycle, $\Delta\Delta G_{hyd}$([N7H] − [N8H] for the N1-protonated form) − $\Delta\Delta G_{hyd}$([N7H] − [N8H] for the N3-protonated form) should be equal to $\Delta\Delta G_{hyd}$([N1-protonated] − [N3-protonated] for the N7H tautomer) − $\Delta\Delta G_{hyd}$([N1-protonated] − [N3-protonated] for the N8H tautomer). According to the results in Tables 5 and 6, an error of (−5.0 + 4.5 + 4.6 − 3.7) = 0.4 kcal/mol is found. This error, which only amounts to 2% of the total free energy changes in the cycle (17.8 kcal/mol), confirms the goodness of the FEP simulation protocol. In addition, the use of this thermodynamic cycle allows us to determine the stability ordering for the protonated forms in water, which is as follows according to FEP calculations: N3-protonated (N8H) > N3-protonated (N7H) ≥ N1-protonated (N8H) > N1-protonated (N7H). This ordering agrees with the results derived from the 6-31G*/MST calculations.

Discussion

The study of the tautomerism and protonation of formycin A is essential for gaining insight into the biological activity of this molecule. Since the ribosyl group at position 9 (position 3, following the IUPAC nomenclature for pyrazole structures) could influence the tautomeric and protonation preferences of the 7-aminopyrazolopyrimidine moiety, extrapolation of results to formycin A should be done with caution. Fortunately, experimental data by Chenon *et al.*[13c] strongly suggest that substituents at position 9 have an almost negligible influence on the tautomeric equilibrium of 7-aminopyrazolopyrimidine derivatives, which allows us to extend with some confidence the discussion of present results to formycin A and other 9-ribosyl derivatives. On the other hand, inspection of the pK_a values of purines and N9-riboderivatives indicates[44] that the ribose at position 9 promotes a moderate decrease in the basicity of the purine ring. This finding could be extrapolated to 7-aminopyrazolo and formycin A, but in this case the decrease of the basicity at N3 due to the ribosyl group at position 9 might be a little bit larger if formycin exists in the *syn* conformation (with an O5′−N3 hydrogen bond). Therefore, our results are expected to enhance slightly the relative basicity of N3 with regard to N1 when they are extrapolated to formycin A.

(44) Izatt, R. M.; Christensen, J. J.; Rytting, J. H. *Chem. Rev.* **1971**, *71*, 439.

Gas Phase. *Ab initio* MP4/6-31G* calculations suggest that the most stable tautomer of 7-aminopyrazolopyrimidine in the gas phase is the N8H form, but an amount (around 15%) of the N7H tautomer exists as well. Comparison of the results shows a moderate effect of the basis set on the determination of the ratio between tautomers and a very small influence of the geometry. In contrast, the addition of correlation effects is very important for the proper representation of the tautomerism of the neutral 7-aminopyrazolopyrimidine derivatives. In this respect, comparison between 3-21G and 6-31G* data, and between SCF, MP2, MP3 (data not shown, but available upon request), and MP4 results, shows a progressive convergence of the results as the quality of the calculation is increased, which suggests that dramatic changes in the results cannot be expected. However, and considering all the potential source of errors (geometry, basis set, and level of accuracy in the representation of correlation effects), caution is necessary since the range of error of the calculation can be larger than the value of ΔG_{taut} found at the highest level of computation.

Protonation of 7-aminopyrazolopyrimidine derivatives takes place at the pyrimidine ring rather than at the pyrazolo ring.[13e] The results presented here indicate that protonation largely raises the population of the N8H tautomer in the gas phase. Protonation at N3 is clearly favored, especially for the N7H tautomer. Accordingly, the N8H tautomer of the N3-protonated species is the most stable species of the protonated 7-aminopyrazolopyrimidine in the gas phase, followed by the N3-protonated N7H tautomer and the N1-protonated N8H tautomer. The protonation energies at the MP4/6-31G* level (−228 kcal/mol (N7H) and −236 kcal/mol (N8H) for N1 protonation, −237 kcal/mol (N7H) and −241 kcal/mol (N8H) for the N3 protonation) suggest that a notable increase in the basicity of the nitrogens occurs (around 4−8 kcal/mol), especially at N1 (8 kcal/mol), when the neutral molecule is forced to be in the N8H tautomeric form. Unfortunately, to our knowledge, there are no experimental data on the tautomerism of 7-aminopyrazolopyrimidine derivatives in the gas phase available for comparison. Previous *ab initio* SCF results in the gas phase[15] calculated using mixed basis sets[45] agree well, in general, with the present results, but they were unable to describe the tautomerism for the neutral form due to the neglect of correlation effects.

Solvent Effect. Due to the relevance of solvent effects on the tautomeric equilibrium, the analysis of the results determined for 7-aminopyrazolopyrimidine merits a detailed discussion. In this respect, reasonable agreement between the different methods used to describe solvation effects is found for protonated forms of the molecule. However, a somewhat disappointing discrepancy is found for the neutral molecule. AMSOL clearly favors the N8H tautomer. The AM1/MST favors slightly the N8H tautomer, but the 6-31G*/MST and FEP results indicate that the N7H tautomer is better hydrated. All the experimental evidence derived from luminescence spectra,[13a] NMR,[13b-d] temperature jump studies,[13e] and magnetic circular dichroism[13f] demonstrates that for the neutral molecule, the N7H form is the major tautomer in water, although some population of the N8H tautomer has also been detected. Only the FEP calculations are able to reproduce the preference for the N7H tautomer in water. To our knowledge, only two quantitative measures of the tautomerization free energy (ΔG^{taut}) of neutral 7-aminopyrazolopyrimidine have been reported: −1.0 kcal/mol from temperature jump experiments in water[13e] and −1.6 kcal/mol from NMR data in D_2O.[13d] Considering the source of errors in both MD−FEP and gas phase calculations, the agreement between MP4/6-31G* and MD−FEP (−0.7 kcal/mol; −0.9 kcal/mol from the 420 ps simulation) and the experimental data is quite reasonable. This agreement supports the goodness of these calculations and also gives confidence to the suitability of the 6-31G*/MST calculations, which provide $\Delta\Delta G_{hyd}$ for the neutral species close to the FEP values.

The protonation of 7-aminopyrazolopyrimidine is likely to occur at N3, but some fraction (around 15% of the N3-protonated molecules from FEP results) is expected to exist as the N1-protonated form in aqueous solution. In any case, it seems that protonation favors the change from the N7H tautomer to the N8H tautomer, both in the gas phase and in aqueous solution. Extrapolation of these results to formycin, where the N3 can be hindered sterically, must be done with some caution. Experimental data on the relationship between tautomerization and protonation, which would be extremely useful, are scarce and have only qualitative value. Very recent NMR measurements in D_2O[13d] are consistent with our suggestion that N3 is the most basic point of the molecule. Temperature jump experiments[13e] agree with the larger basicity of the N8H tautomer predicted in our calculations. Furthermore, the change of tautomeric preference from N7H to N8H, which is suggested to occur upon protonation, seems to be consistent with recent NMR data in aqueous solution.[13d]

At this point, some consideration about the FEP and SCRF techniques is appropriate. The FEP method is a suitable technique for the determination of $\Delta\Delta G_{hyd}$ between tautomers, since (i) $\Delta\Delta G_{hyd}$ is computed directly, which largely reduces errors, and (ii) it is easy to define a reversible pathway for the interchange of tautomers. In contrast, the determination of $\Delta\Delta G_{hyd}$ is a challenging task for SCRF methods since $\Delta\Delta G_{hyd}$ is computed as the difference between the total free energies of hydration of the two tautomers; *i.e.* a difference of 4−5 kcal/mol is to be obtained from the difference between ΔG_{hyd} values in the range of 60−70 kcal/mol. Therefore, in general, for the study of tautomerism, SCRF methods are more susceptible to technical problems than FEP techniques, and it is not surprising that the best agreement with the scarce available experimental data is obtained by using this latter method.

The agreement between FEP- and MST-based SCRF calculations is good, especially when the 6-31G*/MST values are dealt with. Thus, for the five processes studied, the RMS difference between MST and FEP results is 1.1 (6-31G*), 1.3 (AM1/MST with gas phase geometry), and 1.8 (AM1/MST with AMSOL geometry) kcal/mol. Larger RMS deviations are found for AMSOL (3.2 kcal/mol with gas phase geometry; 3.0 kcal/mol with AMSOL-optimized geometry). It is interesting to note that the agreement between AM1/MST and FEP values slightly decreases when the AMSOL-optimized geometry is used, while the goodness of AMSOL results slightly increases when the geometry is fully optimized in water. This is not surprising, because AMSOL was parametrized considering the optimized geometry in solution, while gas phase geometries were used to parametrize non-electrostatic interactions in the MST method.

The results in Tables 3 and 5 show that the main failure of the SCRF methods lies in the determination of solvent effects on the N7H ↔ N8H tautomerism of the neutral molecule. In this respect, the 6-31G*/MST method gives the correct sign for the $\Delta\Delta G_{hyd}$, which indicates that the shortcoming of the AM1-based SCRF methods cannot be fully attributed to intrinsic deficiencies of the SCRF method. However, comparison of AM1/MST and AMSOL values suggests that this latter method has some intrinsic limitations for the representation of these systems.

Biological Relevance. The understanding of the tautomeric equilibrium of 7-aminopyrazolopyrimidine is essential for

(45) Orozco, M.; Luque, F. J. *Chem. Phys. Lett.* **1989**, *160*, 305.

gaining insight into the biological and pharmaceutical properties of formycin A and its derivatives. For instance, the N7H ↔ N8H tautomerism of formycin derivatives can be very important in the determination of the mechanism that allows formycin A to inhibit the bacterial purine nucleoside phosphorylase, while it is not an inhibitor of the mammalian enzyme.[4] Another case where a detailed knowledge of the tautomerism of formycin is necessary corresponds to the use of formycin as a surrogate of adenosine in the nucleotidic sequence of nucleic acids.[5,6] The manipulation of the tautomeric ratio of formycin might change the hydrogen bond pattern in the major groove while preserving identical Watson−Crick pairing with a thymine. It is suggested that in the polar medium surrounding the nucleic acid, formycin (or 2′-deoxyformycin) might exist as the N7H tautomer, which would alter the hydrogen bond pattern along the major groove with respect to that of adenosine. However, changes of the environment polarity, like those originated from the interaction of DNA with a protein, could modify the tautomeric preference. In this context, the identification of the factors modulating the tautomerism of formycin in water are of the greatest importance. Moreover, the study of the tautomeric preferences of enzymes like RNA polymerases can be also relevant in understanding the mechanism of action of formycin. Experiments of incorporation of 7-deazaadenosine and ribosyl derivatives of 4-aminopirazolopyrimidine in the nucleic acids would be very helpful.

The interaction mechanism of formycin with ADA is another example that emphasizes the relevance of the N7H ↔ N8H tautomerism. Theoretical (this work) and experimental studies[13] suggested that neutral formycin should exist in the N7H tautomeric form. Present results indicate that in aqueous solution, the N8H tautomer will be a minor species, in agreement with the available experimental data. However, gas phase calculations support the hypothesis that the N8H form can be a stable tautomer for neutral formycin in the gas phase (perhaps the most stable one). Inspection of the tridimensional structure of ADA[11b,c] shows that the active pocket is absolutely buried at the bottom of a β-barrel, surrounded by an environment in which the solvent is excluded. These conditions are not very different from the vacuum and, accordingly, can lead to the a relative stabilization of the N8H tautomer. As noted previously, the existence of formycin in the N8H tautomeric form leads to a notable increase in the basicity of the N1, which will make the protonation at this atom by a nonionized aspartic acid easier,[11b,c] the resulting species being markedly stable. All these findings might justify the surprisingly large V^{max} of deamination of formycin (10 times greater than that of adenosine[7]), especially if the protonation at N1 is involved in the rate-limiting step of the enzymatic reaction, as previously suggested.[11a,14,46] It can be expected that a modification of the tautomeric preference of formycin A in the gas phase toward an increased population of the N7H tautomer may yield a new drug with most of the properties of formycin A, but resistant to the action of ADA.

Acknowledgment. This work has been supported by the Centre de Supercomputació de Catalunya (CESCA, Mol. Recog. Project) and by the Spanish Government (DGICYT project PB93-0779).

JA9426453

(46) Orozco, M.; Velasco, D.; Canela, E. I.; Franco, R. *J. Am. Chem. Soc.* **1990**, *112*, 8221.