where the reaction partitioned to two distinct, but symmetry equivalent, products.[38]

This latter secondary orbital interaction is extremely important in NC (see Figure 5B). The overlap population between $C_1$ and $C_6$ is only 0.018, which is substantially smaller than the so-called "secondary" overlap between $C_3$ and O. Indeed these latter two atoms are closer together (2.588 Å) than $C_1$ and $C_6$ (2.960 Å) between which a bond is formed experimentally. The structure NC might lead to the formation of vinyldihydropyran. However, the potential surface is extremely flat. Structure NCA (Figure 6) lies on the potential ridge separating reactants from products. It has a shorter $C_1$-$C_6$ distance (2.728 Å) and a longer $C_3$-O distance (2.820 Å) than NC, but it is only 0.05 kcal/mol higher in energy at the 3-21G level. Indeed, RHF/6-31G* single-point calculations place NCA 0.2 kcal/mol below NC. Except for the slight twisting of the plane of the butadiene relative to the acrolein which leads to the difference in bond lengths, there are no further major differences between NC and NCA. Indeed we found a number of similar structures on the potential ridge separating reactants and products, all of which had similar energies, and similar Coulombic attractions.

There is much greater diastereoselectivity in the catalyzed Diels-Alder reactions of chiral acrylates as compared to the uncatalyzed reactions. The facial selectivity is presumably due to steric effects from the chiral auxiliary. One would expect that these would be mitigated by both the longer $C_1$-$C_6$ bond distance and the softer torsional potential. However, the energies of the competing transition structures must also be considered. For the uncatalyzed case,[12] NC and NT are only 0.6 kcal/mol apart, and the reaction will proceed to some extent through both pathways. For chiral reactants, and assuming steric factors give 100% diastereoselectivity in each transition structure, this still gives two diastereomers as determined by the relative barrier heights. In contrast, for the catalyzed case, NT is 1.5 kcal/mol higher than NC, resulting in greater selectivity. Additionally, the strong Coulombic interactions may favor closer approach of some atoms not directly involved in bond formation, which would then permit a greater steric differentiation of diastereomeric transition structures than in the uncatalyzed case.

(38) Stanton, R. E.; Mclver, J. W., Jr. *J. Am. Chem. Soc.* **1975**, *97*, 3632-3646. See Figure 2.

## Conclusion

The concerted, but very asynchronous, catalyzed Diels-Alder reaction proceeds via a transition structure with one strong and one weak forming bond. The latter imparts considerable flexibility to the structures. FMO interactions as well as the resulting charge separation and dipolar interactions are important factors in the reaction. Together these factors lead to the lowering of the activation energy and to the enhancement of the endo selectivity and regioselectivity commonly observed upon catalysis.

**Acknowledgment.** We are grateful to the National Science Foundation for financial support of this research and for generous grants of computer time at the Pittsburgh Supercomputing Center. Jeffrey D. Evanseck provided valuable assistance on computations and graphics.

**Registry No.** $BH_3$, 13283-31-3; butadiene, 106-99-0; acrolein, 107-02-8.

**Supplementary Material Available:** Full geometries and energies of all the transition structures (12 pages). Ordering information is given on any current masthead page.

# Calculations of p*K* Differences between Structurally Similar Compounds

**Alexander A. Rashin,*,† James R. Rabinowitz,‡ and Jason R. Banfelder†**

*Contribution from the Department of Physiology and Biophysics, Mount Sinai School of Medicine of the City University of New York, New York, New York 10029, and Environmental Protection Agency, Research Triangle Park, North Carolina 27711.* *Received May 18, 1989*

**Abstract:** An attempt is made to find out whether accurate calculations of p*K* differences in a series of biologically active compounds are feasible. The computational method used employs a combination of quantum mechanical calculations of the vacuum proton affinities and a new method for calculation of hydration energies based on a continuum representation of the solvent. We demonstrate that our method leads to good agreement with experimental results and results of the free energy perturbation method in calculations of tautomeric equilibria and some computational "mutation" simulations. Application of our method to the calculations of p*K* differences between congeners of imidazolium shows that experimental values of such differences can be computationally predicted within one p*K* unit for pairs of congeners differing in the hydrogen to methyl substitution or in the position of the same substituent in the molecule. However, calculated values deviate from the experimental ones by more than four p*K* units for the pairs of congeners differing in the substitution of the hydrogen or methyl group to the chloro or nitro group. The deviations from experimental values found for these substitutions may be attributed to errors in the calculated proton affinities in vacuum that do not cancel for substitutions involving groups with very different chemical properties. These findings suggest a practical restriction on the quantitative calculations of p*K* differences to narrow classes of compounds. Computations of p*K* differences performed for three congeners of the neurotransmitter histamine from such a class lead to the results that agree with experimental physiological data.

Many chemical and biochemical phenomena in solution are governed by relatively small free energy differences. Such phenomena include conformational and tautomeric equilibria, proton transfer between two groups, differential ligand binding, p*K* differences between closely related compounds, etc. The commonality and importance of these phenomena made them a popular object of experimental and theoretical studies.[1-7] Until

† Mount Sinai School of Medicine.
‡ Environmental Protection Agency.

(1) Aue, D. H.; Webb, H. M.; Bowers, M. T. *J. Am. Chem. Soc.* **1976**, *98*, 318.
(2) Pullman, A.; Pullman, B. *Quart. Rev. Biophys.* **1975**, *7*, 505.

0002-7863/90/1512-4133$02.50/0 

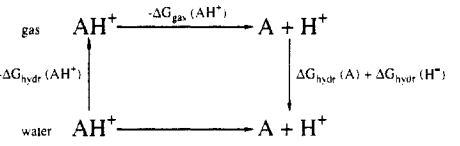


**Figure 1.** The thermodynamic cycle relating the aqueous p$K_a$ of $AH^+$ to free energies of hydration and the gas-phase proton affinity.

recently, however, most theoretical efforts were limited to the in vacuo studies due to the lack of reliable quantitative methods for the computational treatment of solution energetics. Attempts to include a few key water molecules in such calculations, while neglecting effects of the bulk solvent, did not alter their basically in vacuo character.[2]

Significant progress toward the quantitative treatment of these phenomena has been achieved during the last few years with the development and wide application of the free energy perturbation techniques[8] based on molecular representations of the solvent. Recently we have introduced a new method for the computational treatment of the energetics of solvation effects[9] based on the continuum representation of the solvent. This method is computationally simpler and more efficient than methods based on molecular descriptions of the solvent and leads to excellent agreement between the calculated and experimental hydration enthalpies for a large number of molecules.[9-11]

In this communication we address the feasibility of reliable predictions of p$K$ differences between structurally similar compounds. In addressing this problem we have chosen to calculate p$K$ differences between congeners of the neurotransmitter histamine and between congeners of imidazolium, which is related to histamine. The choice of histamine and its analogues was dictated by the finding that only the monocation form of histamine is recognized at the histamine-$H_2$ receptor, and thus the p$K_a$ of the ring nitrogens of this neurotransmitter and of its analogues directly bears on their physiological activity.[12]

To further test the reliability of our results we also check whether our method can reproduce the results obtained with the free energy perturbation technique. In carrying out such a test we focused on the results of the free energy perturbation studies for which all parameters (i.e., geometries and charges) were published, experimental values were available, and parameters were not fitted to reproduce these experimental values. The published parameters were used as input in computations with our method based on the continuum representation of the solvent, which at this stage is limited to evaluation of enthalpies alone. The results of our computations were then compared with the corresponding results from the free energy perturbation studies.

We demonstrate that our method for computation of enthalpy differences based on the continuum description of the solvent leads to the same results as molecular free energy perturbation methods in the applications considered. We also show that accurate quantitative prediction of p$K$ differences between structurally similar compounds is possible at present only if these compounds do not differ in groups with very different chemical properties (e.g., the difference is only in the position of the same group in the molecule) and may be quite inaccurate otherwise mainly due to errors in quantum mechanically calculated proton affinities.

## Methods

**1. Contributions to p$K$ Differences.** Consideration of the thermodynamic cycle in Figure 1 shows that for a cation $AH^+$

$$pK_a(AH^+) = 1/2.3RT[\Delta G_{hyd}(A) + \Delta G_{hyd}(H^+) - \Delta G_{gas}(AH^+) - \Delta G_{hyd}(AH^+)] \quad (1)$$

where $\Delta G_{gas}(AH^+)$ is the proton affinity in the gas phase, and the subscript hyd denotes free energies of transfer of different species from the gas phase to water.

Application of expression 1 to the difference between p$K$s of two structurally similar compounds leads to

$$\Delta pK = pK_a(BH^+) - pK_a(AH^+) = 1/2.3RT[\Delta G_{hyd}(B) - \Delta G_{hyd}(A) - \Delta G_{hyd}(BH^+) + \Delta G_{hyd}(AH^+) + \Delta G_{gas}(AH^+) - \Delta G_{gas}(BH^+)] \quad (2)$$

It is clear from expression 2 that the p$K$ differences are determined by the free energy differences. The entropic contributions to individual p$K$s can be significant.[1] However, these contributions can be expected to be similar for structurally similar compounds so that their contribution to p$K$ differences between such compounds can be small. Such expectation is confirmed by the extensive compilation of the entropic and enthalpic contributins to p$K$s of the series of compounds given in ref 1. The compilation shows that entropic contributions to p$K$ differences between congeners of methyl-, ethyl-, and butylamines differing by one methyl group do not exceed ~1.5 kcal/mol.[1] This allows for the substitution of the free energies of hydration in expression 2 for hydration enthalpies without introducing large errors, as methods for evaluation of full hydration entropies are not presently available.[11] In fact, the error introduced by such a substitution is acceptably close to the ideal of the "chemical accuracy" (1 kcal/mol) which is rarely achieved in the free energy calculations (see Discussion below).

**2. Calculation of Hydration Enthalpies.** Hydration enthalpy of a polar or charged molecule can be represented as a sum of two terms[9]

$$\Delta H_{hyd} = \Delta H_{el} + \Delta H_{np} \quad (3)$$

where $\Delta H_{el}$ is the electrostatic contribution to the hydration enthalpy, and $\Delta H_{np}$ is the nonpolar contribution.

**A.** The electrostatic contribution equals the energy of interaction of the charge distribution of the solute with the reaction field due to the polarization of the solvent by this charge distribution, corrected for small (see below) entropic contributions due to the temperature dependence of the dielectric constant and of the cavity size. The reaction field can be obtained for molecules of arbitrary shape as a numeric solution of the Poisson equation for the charge distribution of the solute in the cavity formed by the solute in a continuum dielectric representing the solvent. The numeric solution of the Poisson equation can be represented as the potential of polarization charges induced on the surface of the cavity by the charge distribution of the solute.[9] The surface density of these polarization charges can be obtained with the Boundary Element Method as a solution of the system of linear algebraic equations[9]

$$\{[\mathbf{I}] - (D_i - D_o)/(2\pi(D_i + D_o))\,[\mathbf{K}]\}\,[\sigma] = [E_f]\,(D_i - D_o)/(2\pi(D_i + D_o)) \quad (4)$$

where the surface of the cavity is divided in $N$ boundary elements with a constant polarization charge density $\sigma_i$ on $i$th element ($i = 1, N$), $[\mathbf{I}]$ is a unit matrix, $[\mathbf{K}]$ is a $N \times N$ matrix of coefficients determined by the geometry of the boundary elements and by their mutual orientation, $[\sigma]$ is a column of $\sigma_i$'s, $[E_f]$ is a column of the mean values of the normal component of the electric field of the charge distribution of the solute on each element, and $D_i$ = 1 and $D_o$ = 78 are the dielectric constants inside and outside the

(3) Osman, R.; Mazurek, A. P. *J. Phys. Chem.* **1985**, *89*, 460.
(4) Weinstein, H.; Mazurek, A. P.; Osman, R.; Topiol, S. *Mol. Pharmacol.* **1986**, *29*, 28.
(5) Gersten, J. I.; Sapse, A. M. *J. Phys. Chem.* **1981**, *85*, 3407.
(6) (a) Rossky, P.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101*, 1913. (b) Beveridge, D. L.; Ravishanker, G.; Mezei, M.; Gedulin, B. In *Biomolecular Structure and Dynamics III*; Ramaswami, H. S., Sarma, M. H., Eds.; Adenine Press: Albany, NY, 1985. (c) Warshel, A. *J. Phys. Chem.* **1979**, *83*, 7640.
(7) (a) Mazut, R. A.; Kushick, J. N. *J. Comput. Chem.* **1985**, *6*, 148. (b) Lifson, S.; Felder, C. E.; Shanzer, A. J. *Biomol. Struct. Interact.* **1984**, *2*, 641. (c) Lybrand, T. P.; McCammon, J. A.; Wipff, G. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 833.
(8) (a) Mezei, M.; Beveridge, D. L. In Computer Simulations of Chemical and Biomolecular Systems. *Annals NY Acad. Sci.* **1986**, *482*, 1. (b) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: 1987.
(9) Rashin, A. A.; Namboodiri, K. *J. Phys. Chem.* **1987**, *91*, 6003.
(10) Rashin, A. A. *Int. J. Quant. Chem.: Quant. Biol. Symposium* **1988**, *15*, 103.
(11) Rashin, A. A. Hydration Phenomena, Classical Electrostatics, and the Boundary Element Method. Submitted to *J. Phys. Chem.*
(12) Ganellin, C. R.; Durant, G. J. In *Burger's Medicinal Chemistry*; Wolff, M. E., Ed.; Wiley-Interscience: New York, 1981; p 487.

cavity. (A justification for the use of $D_0 = 78$ starting right outside the cavity surface has been given recently on the basis of Monte Carlo simulations[13]). A detailed derivation of this system of equations, a method of its numeric solution, and a discussion of the applicability of the continuum representation of the solvent are given in ref 9, where it is demonstrated that the method leads to accurate predictions of hydration enthalpies of polar and charged molecules. An extension of this formalism that accounts for polarization of the solute by the solvent, which is neglected in this work as well as in most free energy perturbation calculations, is presented in refs 10 and 11. This extended formalism has been applied successfully only to a few molecules and has not been sufficiently tested to be used with confidence in the present studies.

It should be noted that we have chosen to calculate hydration enthalpies and not free energies[9] for the following reasons. Both thermodynamic quantities have nonelectrostatic components. For enthalpy this component can be rather accurately evaluated from successful theories of hydration of nonpolar solutes.[14] This cannot be presently done for the nonelectrostatic contribution to hydration entropy as entropies of hydration of ions and of nonpolar molecules change in a completely different way with the size of the solute.[11] Entropic contribution to the free energy of hydration of ions is ~10% (or ~10 kcal/mol), and ~80% of this is nonelectrostatic.[11] On the other hand, electrostatic part of hydration enthalpy can be obtained from the electrostatic work by two small corrections. The first one due to the temperature dependence of the solvent's dielectric constant is ~1.7%, and the second one due to the temperature dependence of the cavity size should be obtained from the thermal expansion coefficients which can be reliably interpreted only for pure substances or very high concentrations of the solute. It is rather obvious that the cavity expansion should decrease with the strength of the solute–solvent interactions and thus should be smaller for ions than for dipolar solutes. This second correction is ~2% for water, and about an order of magnitude less for ions, as estimated from the thermal expansion coefficients of water and ionic crystals correspondingly.[11] This second correction was neglected here, as its contribution to p$K$ differences is negligible and as in most cases it is difficult to evaluate accurately. Thus the full enthalpy of hydration can be calculated and verified against experimental data. The latter cannot be done for the electrostatic component of the free energy of hydration. We prefer to deal with a quantity better related to experiment at the price of a negligible few percent of the calculated values.

**B.** The cavity radius for each atom in a molecule is defined as the radius of a sphere around its nucleus at which the polarizable electron density of the solvent rises sharply and which contains inside a negligible electron density of the solvent.[9] Such radii are self-consistently determined from crystallographic and NMR data and are shown to be transferable.[9] For polyatomic molecules the cavity should contain parts of the sharp corners between the overlapping spheres of individual atoms, as such corners cannot be completely filled by the solvent molecules having a shape and a finite size. Therefore the boundary between the molecular cavity and the solvent represented as a dielectric can be defined as a "molecular surface".[15] This surface is formed by rolling a spherical probe on a surface formed by the overlapping spheres of atomic cavities.[9]

According to our definition the probe radius describes a sphere delimited by the electron density of the solvent molecule high enough to determine the dielectric boundary of the molecular cavity ("high electron density core of the solvent molecule")[9] and therefore can be smaller than the van der Waals radius of the solvent.[9,14,16]

The determination of the cavity radii consistent with our definition based on the electron density distribution of the solvent around the solute comes from the analysis of electron density distributions in ionic crystals.[9,16] Such analysis shows that if ionic radii are accepted as the cavity radii for anions, then it follows from our definition that the covalent radii should be accepted as the cavity radii for cations (see refs 9 and 16 for details). These radii are then used to obtain the radius of the high electron density core of the solvent molecule from experimentally known internuclear distances between ions and solvent molecules in the first solvation shell.[9,16] A substitution of experimental values for these radii in the Born formula leads to a systematic 7% difference between the calculated and experimental hydration enthalpies for alkali halide salts.[9,16] Introduction of a 7% increase of the table values of ionic and covalent radii used in the Born formula has led to agreement within a few percent between the calculated and experimental values of hydration enthalpies for over 30 spherical ions.[16]

The cavity radii of nonionic solutes are obtained from experimentally known internuclear solute–solvent distances and the radius of the high electron density core of the solvent molecule (see ref 9 for details). The 7% increase in the cavity radii obtained from crystallographic data is the only empirical fitting constant present in our method. This constant has been introduced on the basis of experimental thermodynamic data for alkali halide salts alone and transferred to all cavity radii, leading to excellent agreement between the calculated and experimental values of hydration enthalpies for a large number of spherical ions and polar polyatomic solutes.[9-11,16]

**C.** The scaled particle theory[14] shows that starting with some size of the nonpolar solute larger than the size of the solvent molecule the solvation enthalpy becomes proportional to the surface area of the "hard core" cavity formed by the solute in the solvent. These hard core radii are different from the cavity radii used in the calculations of the electrostatic contribution to the hydration enthalpy because they are related to a different physical property.[9] The values of the "hard core" radii can be found in publications on the scaled particle theory.[14] In our method we assume that nonelectrostatic contributions to solvation enthalpy of a polar or charged molecule are equal to the enthalpy of hydration of a nonpolar molecular forming the same hard core cavity in the solvent. Enthalpies of hydration of nonpolar molecules are obtained as a function of the surface area of their hard core cavity surface by interpolation of either the results of the scaled particle theory or of the experimental values (see ref 9).

**3. Quantum Calculations of the Gas-Phase Proton Affinities.** All quantum calculations were done with the GAUSSIAN 82 and GAUSSIAN 86 packages.[17] Geometry optimizations were performed with the 6-31G basis set. Energy calculations were performed either with the 6-31G basis set or with the 6-31G** basis set.

**4. Representation of the Charge Distribution of the Solute.** Two methods for representing the charge distribution of the solute were used. One of them employs the traditional Mulliken charges centered on the nuclei.[18] The shortcomings of such representation of the electron density are well-known.[18] The alternative representation of the electron density from the 6-31G basis set was obtained from a multipolar expansion up to quadrupoles with the method described in ref 19. In this representation partial point charges, point dipoles, and point quadrupoles are placed on nuclei and on lines connecting them. Their number is of the order of the number of the basis functions used, i.e., it is larger than the number of the nuclei (compare ref 20). It has been shown that this representation reproduces very accurately electrostatic potentials and molecular dipole moments calculated from the molecular wave functions.[19]

**5. Comparison with the Free Energy Perturbation Results.** For this comparison we selected two groups of results obtained with the free energy perturbation method. The first group consists of

(13) Jayaram, B.; Fine, R.; Sharp, K.; Honig, B. *J. Phys. Chem.* **1989**, 93, 4320.
(14) Pierotti, R. A. *J. Phys. Chem.* **1965**, *69*, 281.
(15) Connolly, M. L. *Science* **1983**, *221*, 709.
(16) Rashin, A. A.; Honig, B. *J. Phys. Chem.* **1985**, *89*, 5588.
(17) Binkley, J. S.; Frosch, M. J.; DeFrees, D. J.; Raghavachari, K.; Whiteside, R. A.; Schlegel, H. B.; Fluder, E. M.; Pople, J. A. GAUSSIAN 82; Carnegie-Mellon University: Pittsburgh, PA, 1983.
(18) Rein, R.; Rabinowitz, J. R.; Swissler, T. J. *J. Theor. Biol.* **1972**, *34*, 215.
(19) Rabinowitz, J. R.; Namboodiri, K.; Weinstein, H. *Int. J. Quant. Chem.* **1986**, *29*, 1697.

Table I. Comparison of the Enthalpy Differences Obtained with the Macroscopic Representation of the Solvent with the Free Energy Differences Obtained with the Microscopic Representation of the Solvent

| | kcal/mol | | | |
|---|---|---|---|---|
| | macroscopic $\Delta H^a$ | | microscopic | |
| process | electrostatic | total[d] | $\Delta G^b$ | expt[c] |
| 2-hydroxopyridine → 2-oxopyridine | −3.9 | −3.3 | −4.7 ± 0.4 | −4.1 |
| 2-hydroxopyrimidine → 2-oxopyrimidine | −3.8 | −1.3 | −2.9 ± 0.2 | <−1.6 |
| iminocytosine → aminocytosine | −4.72 | −4.75 | −4.2 ± 0.2 | −5.5 ± −6.9 |
| methane → acetamide | −7.97 | −10.0 | −10.3 ± 0.2 | −11.8 |
| methane → thymine | −6.2 | −10.8 | −9.6 ± 1.2 | |

[a] Results obtained with our method based on macroscopic representation of the solvent with geometries, charges, and quantum mechanical contributions from refs 21 and 22. [b] Results of the combination of the quantum mechanical and free energy perturbation calculations from refs 21 and 22. [c] From refs 21 and 22. [d] Includes also nonpolar and/or quantum mechanical contributions.

the free energies of tautomerization of 2-oxopyridine, 2-oxopyrimidine, and cytosine,[21] and the second group consists of free energies of "mutation" of methane to acetamide and to thymine.[22] Experimental results are available for all corresponding calculated values except the difference in the free energy of hydration between methane and thymine which has been calculated as a prediction. Interaction parameters in these free energy calculations were not fitted to reproduce the corresponding experimental values. Point charges on nuclei of the solutes were fitted to reproduce electrostatic potentials from the wave functions,[20] and other parameters were taken from the molecular simulation program AMBER.[23] Water molecules were represented by the TIP3P model.[24] We used the quantum mechanical tautomerization energies in the gas phase obtained with 6-31G* basis set published in the original papers[21,22] and used molecular geometries and point charges from these publications as input in our calculations of hydration enthalpies based on the continuum representation of the solvent. Polarization of the solutes by the solvent was neglected both in the free energy simulations and in our calculations.

## Results

**A. Comparison with the Free Energy Perturbation Results.** To establish the reliability of our method we first represent a comparison of the results of our method with the corresponding results from the free energy perturbation studies (Tables I). The table shows that our results agree very well with the results of the free energy simulations and/or with experimental values. This provides us with some confidence in our calculations of p$K$ differences reported below.

**B. p$K$ Differences between Congeners of Imidazolium.** Results of our calculations of p$K$ differences are listed in Table II and are divided in groups according to the molecules whose p$K_a$ is taken as a reference.

In the first group, p$K_a$ of imidazolium is taken as a reference. The table shows that calculated p$K$ differences vary within only ~1 p$K$ unit with different choices of the charge distributions. However, only for 2-methylimidazolium do the calculated and experimental p$K$ differences agree within the same one p$K$ unit accuracy. For imino and chloro derivatives of imidazolium the calculated p$K$ differences deviate from their experimental values by more than four p$K$ units. These deviations are dominated by large differences in the vacuum proton affinities, suggesting that namely the calculation of this contribution is the main source of the error.

Table II. Calculated and Experimental p$K$ Differences between Congeners of Imidazolium

| | | calculated $\Delta pK_a$ | | | expt[c] |
|---|---|---|---|---|---|
| compound | charge[b] distribution | SCF[d] | hydration[e] | total | p$K_a$ |
| imidazole | | 0.00 | 0.00 | 0.00 | 0.00 |
| 2-Me-1MID | 6-31G/6-31G Mulliken | −4.33 | 3.41 | −0.92 | −0.93 |
| | 6-31G**/6-31G Mulliken | −4.33 | 3.02 | −1.31 | −0.93 |
| | 6-31G/6-31G multipole expn | −4.33 | 2.11 | −2.22 | −0.93 |
| 2-$NH_2$-1MID | 6-31G/6-31G Mulliken | −6.54 | 1.10 | −5.44 | −1.45 |
| | 6-31G**/6-31G Mulliken | −6.96 | 1.07 | −5.89 | −1.45 |
| | 6-31G/6-31G multiple expn | −6.54 | 1.08 | −5.46 | −1.45 |
| 2-Cl-1MID | 6-31G/6-31G Mulliken | 6.99 | 1.33 | 8.32 | 3.52 |
| | 6-31G**/6-31G Mulliken | 6.10 | 1.17 | 7.27 | 3.52 |
| | 6-31G/6-31G multipole expn | 6.99 | 0.58 | 7.57 | 3.52 |
| N1-Me-4-Cl-1MID | | 0.00 | 0.00 | 0.00 | 0.00 |
| N1-Me-5-Cl-1MID | 6-31G/6-31G Mulliken | −1.31 | −0.84 | −2.15 | −1.91 |
| | 6-31G**/6-31G Mulliken | −1.56 | −1.14 | −2.70 | −1.91 |
| | 6-31G/6-31G multipole expn | −1.31 | −1.67 | −2.98 | −1.91 |
| N1-Me-4-Me-1MID | | 0.00 | 0.00 | 0.00 | 0.00 |
| N1-Me-5-Me-1MID | 6-31G/6-31G Mulliken | −0.64 | −0.47 | −1.11 | −0.09 |
| | 6-31G**/6-31G Mulliken | 0.02 | −0.36 | −0.34 | −0.09 |
| | 6-31G/6-31G multipole expn | −0.64 | 0.11 | −0.53 | −0.09 |
| N1-Me-4-Cl-1MID | | 0.00 | 0.00 | 0.00 | 0.00 |
| N1-Me-4-Me-1MID | 6-31G/6-31G multipole expn | −9.60 | 0.53 | −9.07 | −4.39 |
| N1-Me-5-Cl-1MID | | 0.00 | 0.00 | 0.00 | 0.00 |
| N1-Me-5-Me-1MID | 6-31G/6-31G multiple expn | −8.94 | 2.32 | −6.62 | −3.57 |

[a] For the nomenclature see Figure 2a and ref 25; in each group the first compound with zero values of all contributions is taken as a reference. [b] Indicates whether the Mulliken charges or the multiple expansion were used; the basis set shown before the slash was used for the energy calculations; the one after the slash was used in the geometry optimization. [c] From ref 25. [d] Differences in the zero-point energies (not included) are small, 0.05–0.4 pK units, and do not exceed 0.1 pK units for all compounds with the difference between the experimental and calculated values below 1.5 pK units. [e] Calculated with our method.

In the second group we compare two chloro derivatives of methylimidazolium and find that the calculated p$K$ difference is within one p$K$ unit of its experimental value, with both vacuum and hydration contributions being rather small. The same is found in the third group where we compare two methyl derivatives of methylimidazolium.

Table II thus suggests that if p$K$ differences are calculated between compounds that have "chemically similar" substituent groups (e.g., hydrogen and methyl) or differ only in the position of the substituent in the molecule, one can expect an agreement within one p$K$ unit between the calculated and experimental values. If, however, the substituents differ significantly (as, e.g., hydrogen and chloride) large errors in the calculated p$K$ differences can be expected. This expectation is supported in groups four and five comparing methyl and chloro derivatives of methylimidazoliums.

**C. p$K$ Differences between Congeners of the Neurotransmitter Histamine.** Following the trends found in the study of the congeners of imidazolium we have chosen to calculate p$K$ differences between chemically similar congeners of the neurotransmitter histamine listed in Table III. Our calculations show that even in solution the conformation in which the side chain forms a hydrogen bond with the ring (Figure 2b) and in which the proton is on the side chain is more stable than any other conformation. Therefore this conformation of histamine and of its derivatives

(20) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(21) Cleplak, P.; Bash, P.; Singh, U. C.; Kollman, P. *J. Am. Chem. Soc.* **1987**, *109*, 6283.
(22) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564.
(23) AMBER (USCF), Version 3.0; Singh, U. C.; Weiner, P. K.; Caldwell, J. W.; Kollman, P. A. Department of Pharmaceutical Chemistry, University of California: San Francisco, CA, 1986.
(24) Jorgensen, W. L.; Chandrasechar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926.

**Table III.** Experimental and Calculated $pK_a$ Differences between Congeners of the Neurotransmitter Histamine

| compound | charge[b] distribution | calculated $\Delta pK_a$ SCF | hydration[d] | total | exptl[c] $\Delta pK_a$ |
|---|---|---|---|---|---|
| histamine | 6-31G/6-31G multipole expn | 0.00 | 0.00 | 0.00 | 0.00 |
| 4-Me-HISTMN | | -2.39 | 1.78 | -0.61 | -0.69 |
| 2-Me-HISTMN | | -4.61 | 3.05 | -1.56 | |

[a] For the nomenclature see Figure 2b and ref 25. [b] Indicates that the multipole expansion was used for the representation of the charge distribution of the solute; the basis set shown before the slash was used for the energy calculations; the one after the slash was used in the geometry optimization. [c] From ref 25. [d] Calculated with our method.

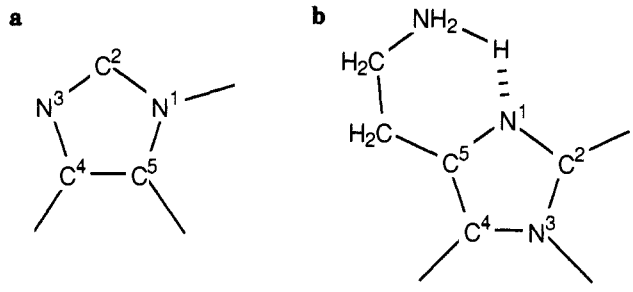


**Figure 2.** Structure and numbering of atoms in (a) imidazolium and (b) histamine.

was used in our calculations of p$K$ differences.

Tables III shows excellent agreement between the calculated and experimental p$K$ differences between histamine and its 4-methyl derivative. Our calculations suggest that the 2-methyl derivative of histamine, whose $pK_a$ has not been determined experimentally, is 1–1.5 pK unit more basic than histamine and its 4-methyl derivative. As only the monocationic form of the neurotransmitter is expected to be physiologically active, our results suggest that the 2-methyl derivative of histamine should be about one order of magnitude less active than histamine and its 4-methyl derivative. This prediction is in agreement with experimental measurements of physiological activity of these compounds.[26]

## Discussion

In our previous publications[9,10] we have shown that our method based on the continuum representation of the solvent leads to accurate predictions of hydration enthalpies for a large number of polar and charged molecules. We have also shown that in a few cases where comparison was possible the hydration enthalpies calculated with our continuum method and with molecular representations of the solvent were similar if the same charge distributions were used in both types of calculations.[9] Therefore the same degree of success in predicting tautomerization energies with our method and with the free energy perturbation technique,[21] which supports the validity of our method, comes at no surprise as entropy differences in tautomerization can be expected to be very small. Excellent agreement between differences in hydration enthalpies calculated with our continuum method, and differences in free energies calculated with the free energy perturbation technique between methane and acetamide, and between methane and thymine are more surprising. In these cases one would expect to find significant entropy differences which should be reflected in differences between results obtained with the enthalpy and free energy calculations. This may be due to the reported 10% accuracy of the free energies calculated with the perturbation method.[22] A general agreement between the results of the two methods suggests that they have similar accuracies and that either one can be used for calculations of free energy differences in cases when entropy differences are not very large. Our method can, however, be significantly less demanding computationally—computation of each value in Table I took about 1 h of VAX-11/785 CPU time. It should be noted that present calculations do not include effects of polarization of the solute by the solvent, and inclusion of such effects in our calculations as suggested in refs 10 and 11 can further improve the accuracy of our method.

A major purpose of this work is to find out whether accurate calculations of p$K$ differences in a series of derivatives of biologically active compounds are feasible. A detailed quantum mechanical study of acidities and basicities[27] shows that errors in their calculated values relative to experimental ones can easily be on the order of 10 kcal/mol even in calculations with large basis sets including diffuse functions. These errors can be rather different for atomic groups of different types (e.g., with localized and with diffuse orbitals). Therefore in calculations of differential proton affinities one might expect cancellation of errors in differences of the proton affinities between groups of the same type and significant residual errors between groups that have significant chemical differences. This expectation is supported by our calculations which show systematic cancellation of errors between derivatives of biological molecules with similar chemical types of substituents and large errors for pairs of compounds that involve substituents with chemically very different types. In the latter case it is apparent that these errors are dominated by contributions from vacuum proton affinities as can be expected on the basis of the results of earlier proton affinity calculations.[27] Similar large errors in predicted vacuum proton affinities and p$K$ differences were found in ref 28 for derivatives of ethane; however, their source, importance, and generality were not discussed. A real understanding of the nature of structural and molecular properties responsible for the errors in the calculated vacuum proton affinities may require quantum chemical studies at the level far beyond the routine level employed in this work. It should be noted that our attempts to add the correlation energy at the MP2 level made the agreement with experiment worse as has been observed in ref 21. It is possible that this is due to the large basis set superposition error that overweighs the introduced corrections at this level of calculations (Krauss, M., personal communication). At present these findings set limits on practically feasible predictions of p$K$ differences due to chemical modifications. Our results suggest that such differences can be predicted with about one p$K$ unit accuracy only for pairs of compounds that do not differ significantly in the type of substituents. This suggestion and practical usefulness of p$K$ calculations even within these limits are supported by our successful qualitative and semiquantitative prediction of differences in physiological activities of different derivatives of neurotransmitter histamine. Our results also suggest that a reliable method for calculation of hydration contribution to p$K$ differences is now available and that a further progress in the computational quantum chemistry may allow accurate computations of p$K$ differences between a wider variety of compounds.

**Acknowledgment.** We thank H. Weinstein for helpful and stimulating discussions, R. Osman for useful suggestions, and A. P. Mazurek for his help with the computations. This work has been supported by NSF Grant DMB-8519273 and NIH Grant GM-38144 (A.A.R.). The research described has been approved by the Health Effects Research Laboratory, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the agency nor does the mentioning of trade names of commercial products constitute endorsement or recommendation for use.

**Registry No.** 2-Me-1M1D, 693-98-1; 2-$NH_2$-1M1D, 7720-39-0; 2-Cl-1M1D, 16265-04-6; N1-Me-4-Cl-1M1D, 4897-21-6; N1-Me-5-Cl-1M-D, 872-49-1; N1-Me-4-Me-1M1D, 6338-45-0; N1-Me-S-Me-1M1D, 10447-93-5; 4-Me-HISTMN, 36507-31-0; 2-Me-HISTMN, 34392-54-6; imidazole, 288-32-4; histamine, 51-45-6.

(25) Ganellin, C. R. In *Molecular and Quantum Pharmacology*; Bergmann, E. D., Pullman, B., Eds; D. Reidel Publ. Co.: Boston, MA, 1975.
(26) (a) Durant, G. J.; Ganellin, C. R.; Parsons, M. E. *J. Med. Chem.* **1975**, *18*, 905. (b) Reggio, P.; Topiol, S.; Weinstein, H. *J. Med. Chem.* **1986**, *29*, 2412.
(27) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley: New York, 1986.
(28) Jorgensen, W. L.; Briggs, J. M.; Gao, J. *J. Am. Chem. Soc.* **1987**, *109*, 6857.