Theoretical Studies of the Tautomeric Equilibria for Five-Member N-Heterocycles in the Gas Phase and in Solution

Peter I. Nagy,*,† Frederick R. Tejada,‡ and William S. Messer, Jr.*,‡

Center for Drug Design and Development, The University of Toledo, Toledo, Ohio 43606-3390, and Department of Pharmacology, The University of Toledo, Toledo, Ohio 43606-3390

Received: June 8, 2005; In Final Form: September 30, 2005

Tautomeric equilibria have been studied for five-member N-heterocycles and their methyl derivatives in the gas phase and in different solvents with dielectric constants of $\epsilon = 4.7-78.4$. The free energy changes differently for tautomers upon solvation as compared to the gas phase, resulting in a shift of the equilibrium constant in solution. Solvents with increasing dielectric constant produce more negative solute−solvent interaction energies and increasing internal energies. The methyl-substituted imidazole and pyrrazole form delicate equilibria between two tautomeric forms. Depending on the solvent, the methyl-substituted triazoles and tetrazole have one or two major tautomers in solution. When estimating the relative solvation free energies by means of an explicit solvent model and using the FEP/MC method, one observes that the preferred tautomers differ in several cases from those predicted by the continuum solvent model. The 1,2-prototropic shift, as an intramolecular tautomerization path, requires about 50 kcal/mol activation energy for imidazole in the gas phase, and this route is also disfavored in a solution. The calculated activation free energy along the intramolecular path is 48−50 kcal/mol in chloroform and water as compared to a literature value of 13.6 kcal/mol for pyrrazole in DMSO. A molecular dynamics computer experiment favors the formation of an imidazole chain in chloroform, making the 1,3-tautomerization feasible along an intermolecular path in nonprotic solvents. In aqueous solution, one strong $N-H\cdots O_w$ hydrogen bond is formed for each species, whereas all other nitrogens in the ring form weaker, $N\cdots H_wO_w$ type hydrogen bonds. The tetrahydrofuran solvent acts as a hydrogen bond acceptor and forms $N-H\cdots O$(ether) bonds. Molecules of the dichloromethane solvent are in favorable dipole−dipole interactions with the solute. The results obtained are useful in the design of N-heterocyclic ligands forming specified hydrogen bonds with protein side chains.

Introduction

Most biologically active molecules, natural and synthetic, are basically polar structures. They form hydrogen bonds with biopolymers, establishing one of the major stabilizing factors for the ligand−receptor complex. The ligand molecule must take an appropriate conformation for such interactions, but the structural change may not be favorable from the point-of-view of the ligand itself. The activation free energy must be covered by the free energy gain for the formation of the ligand−receptor complex.

Protonation and deprotonation result in a switch in the hydrogen bond donor and acceptor character of a ligand site, which is determined by the local pH. It is difficult to anticipate the protonation state of a ligand with a medium-strength basic or acidic site in the binding cavity. The cavities, in the absence of the ligand, may be filled in by water molecules. Although the shape and the size of the binding cavities are generally not known precisely (and even in that case one has to consider the structural changes upon the thermal motion), still it is generally accepted that the binding cavity has structural peculiarities and imposes constraints on the structure of a well-binding ligand. To fit this requirement, some ligands may change their hydrogen-bond donating and accepting sites. Water molecules residing in the cavity could mediate this process.

The above "metamorphosis" is possible through the protonation/deprotonation process or through a tautomeric transformation. Ligands with special substructures can undergo such isomerism. A family of polar molecules of this type is comprised of N-heterocycles with oxo group(s) connected to the ring. In fact, pyrimidine and purine derivatives are the best-known subjects for such oxo−hydroxy (lactam−lactim) tautomerism.[1,3b] 3-OH and 5-OH isoxazoles also show this sort of proton relocation, which was studied in aqueous solution by applying the free-energy perturbation method[2a] and different continuum solvent approaches.[2a−c]

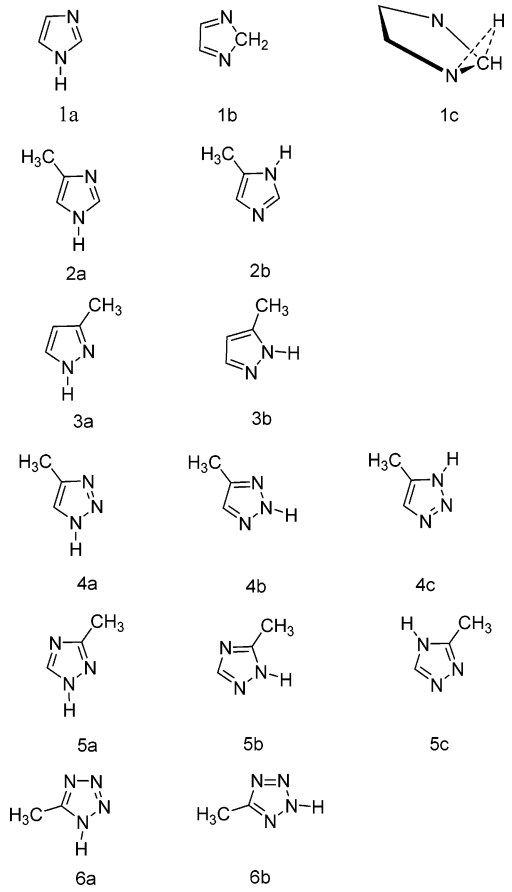
Another important class of molecules with the above potential consists of five-member heterocycles with two or more N atoms in the ring. In the neutral form, one nitrogen is bound to a mobile hydrogen, forming the so-called pyrrolic N−H. The other nitrogen(s) is(are) basic with the lone pair(s) in the ring plane. Scheme 1 shows five representatives of this family. Because of the nonsymmetrical methyl substitution, there are 12 distinguishable tautomers for these molecules.

Imidazole is the most famous representative of this family. The imidazole segment, as part of histamine and the histidine residue in proteins, plays major roles in important biological interactions. The imidazole ring with a nonsymmetrical substitution exhibits an $HN_1\cdots HN_3$ tautomerism studied both experimentally and theroretically.[3] Histamine activates the H_1 and H_2

* Corresponding authors. E-mail: (P.I.N.) pnagy@utnet.utoledo.edu; (W.S.M.) wmesser@utnet.utoledo.edu.
† Center for Drug Design and Development, The University of Toledo.
‡ Department of Pharmacology, The University of Toledo.

10.1021/jp053076w CCC: $30.25
Published on Web 11/04/2005

SCHEME 1: Structures of Compounds[a]



[a] Codes in parentheses are those used in Tables 5 and 6: 1a, 1*H*-imidazole; 1b, 2*H*-imidazole; 1c, imidazole TS; 2a, 4Me-imidazole-1*H* (4Mim1H); 2b, 4Me-imidazole-3*H* (4Mim3H); 3a, 3Me-pyrrazole-1H (3Mpy1H); 3b, 3Me-pyrrazole-2*H* (3Mpy2H); 4a, 4Me-triazole-1,2,3-1*H* (4Mtr123,1H); 4b, 4Me-triazole-1,2,3-2*H* (4Mtr123,2H); 4c, 4Me-triazole-1,2,3-3*H* (4Mtr123,3H); 5a, 3Me-triazole-1,2,4-1*H* (3Mtr124,1H); 5b, 3Me-triazole-1,2,4-2*H* (3Mtr124,2H); 5c, 3Me-triazole-1,2,4-4*H* (3Mtr124,4H); 6a, 5Me-tetrazole-1*H* (5Mte1H); 6b, 5Me-tetrazole-2*H* (5Mte2H).

receptors[4a] whereas its α(*R*)-methyl[4b] and α(*R*),β(*S*) dimethyl[4c] derivatives are highly potent agonists of the H_3 receptor. In the histidine residue, for example, the imidazole ring is part of the catalytic triad for serine proteinases.[5]

3-OH pyrrazole exhibits both N−H···N and oxo−hydroxy tautomerism in aqueous solution. The equilibria were studied by continuum solvent approaches and by using the free energy perturbation method both by Parchment et al.[6a] and by Cao et al.[6b] The latter group used atomic charges fitted to the HF/6-31G** molecular electrostatic potential calculated by the corresponding SCRF wave function. This is a method similar to that used in our present study.

In-solution tautomerism and relative free energies for the 1,2,3- and 1,2,4-triazoles were studied by Cox et al.[7a] and by Murdock et al.[7b] by utilizing the free energy perturbation method and the thermodynamic integration method implemented in variants of the Monte Carlo method. All these studies were performed for the unsubstituted triazole molecules, where only two distinguishable tautomers exist. The in-solution behavior of tetrazole was studied by Wong et al.[8] by utilizing a continuum solvent approach.

However, the mechanism of the tautomerism raises problems. A simple reaction route is possible in protic solvents like water and methanol. The solvent can catalyze the proton relocation by forming an $R-OH_2^+$ cation at the N−H site and returning the proton in a concerted process to the aromatic ring at the basic N-site by forming a deprotonated $R-O^-$ solvent anion. The mechanism is presumably different in nonprotic solvents as dichloromethane (CH_2Cl_2) or tetrahydofuran (THF), where no mobile hydrogen is expected in the solvent. Furthermore, the proton relocation is more complicated if 1−3 tautomerism emerges (imidazole, 1,2,4-triazole) in comparison with the 1−2 tautomerism in pyrrazole, 1,2,3-triazole, and tetrazole.

NMR studies for triazoles by Lunazzi et al.[9a] indicate tautomeric equilibria in CH_2Cl_2 and THF. An intramolecular path based on a 1,2-prototropic shift was proposed in nonpolar solvents. Circumstantial arguments by Albert and Tayler[9b] concluded that the 2*H* form for 1,2,3-triazole is favored over the 1*H* tautomer by a factor of about two in water. Alkorta and Elguero[10] studied theroretically, at the B3LYP/6-31G* level, the transition state (TS) for the 1,2-tautomerization path for pyrrazole. It was pointed out that the relocating proton stays on the outskirt of the ring throughout the intramolecular rearrangement, but is out of the heavy-atom plane in the TS structure. This finding has been utilized in our present study when a two-step intramolecular tautomerization path is proposed for 1−3 tautomerisms..

Recently, different muscarinic agonists have been synthesized in our laboratory.[11] Common in these structures is a five-member heterocyclic ring with, however, only hydrogen-bond acceptor sites. An extension of the heterocyclic rings to ones with both proton donor and acceptor sites raises the problem of determining the prevalent tautomer or estimating the population of different species in a tautomeric equilibrium. Since it is not known whether the N-heterocycles interact with protein side chains by water mediation or enter direct contact with polar or nonpolar side chains in a nonaqueous medium, molecular environments of different polarities were studied in the present investigation. Calculations were performed by utilizing the integral equation formalism of the polarizable continuum method, where the possible different immediate molecular environments were mimicked by a continuum solvent of various dielectric constants. The N-heterocycles are intended to connect to saturated chains in our future molecules, so all five-member rings were substituted by a methyl group in a way that any mirror-image symmetry was destroyed in the molecules in scrutiny.

Previous investigations for studying the tautomeric equilibria used different theoretical approaches and generally only the water solvent.[2,6−8] To the best of our knowledge, this study is the first consistent theoretical procedure considering electron correlation and solvent effects for 12 tautomers in the gas phase and in four different solvents: chloroform, acetone, methanol, and water. For testing our approach, studies for unsubstituted triazoles also were performed and compared with the available experimental results. Finally, Monte Carlo simulations were carried out for estimating the relative solvation free energies for several tautomeric transformations in slightly polar solvents and water. These simulations allow for the characterization of the first solvation shell around the solute, which is not obtainable from the continuum approach. Overall, our aim was to explore what changes in the tautomer composition may be expected when a potential drug with a $-CH_2-$N-heterocycle moiety leaves the aqueous phase and enters the less polar (receptor) environment. Also this study seeks a combination of theoretical approaches to meet this goal at a generally affordable computational effort and cost.

TABLE 1: Optimized Geometric Data for the Gas-Phase 1*H*-Imidazole[a]

	MP2			B3LYP			
	6-31G*	6-311++G**	aug-cc-pvtz	6-31G*	6-311++G**	aug-cc-pvtz	expt
N_1C_2	1.366	1.367	1.362	1.367	1.367	1.363	1.364
C_2N_3	1.325	1.325	1.322	1.315	1.312	1.310	1.314
N_3C_4	1.377	1.377	1.373	1.379	1.377	1.375	1.382
C_4C_5	1.378	1.382	1.377	1.372	1.371	1.367	1.364
C_5N_1	1.376	1.376	1.372	1.381	1.380	1.376	1.377
RMS	0.008	0.009	0.008	0.004	0.004	0.004	
$N_1C_2C_3$	111.6	111.6	111.4	111.8	111.5	111.5	112.0
$C_2N_3C_4$	104.9	105.1	105.2	105.2	105.5	105.6	104.9
$N_3C_4C_5$	111.0	110.9	110.8	110.8	110.6	110.5	110.7
$C_4C_5N_1$	104.9	104.9	105.0	105.0	105.1	105.2	105.5
$C_5N_1C_2$	107.6	107.5	107.7	107.2	107.2	107.3	106.9
RMS	0.5	0.4	0.5	0.3	0.4	0.5	

[a] Distances in Å, angles in deg.

Methods and Calculations

B3LYP/6-31G* calculations provide molecular geometries in good agreement with the experimental gas-phase structure for methylamine, aniline, and phenol,[12] formaldoxime,[13] methyl acetate, and methyl thioformate.[14] Good agreement was found also for methanol, whereas the derived dipole moments compared favorably with the experimental values (in the absence of experimental structures) for different conformers of ethanol and isopropyl, isobutyl, and *tert*-butyl alcohols.[15b] B3LYP/6-31G* predicts probably good molecular geometries for the 2-(4OH-phenyl)-ethylamine (tyramine) conformers, but underestimates the gauche−trans energy separation[16] relative to the MP2/6-31G* value and the energy separation concluded from free-jet microwave spectroscopy data.[17]

In the present study, the theoretical structure for imidazole was compared with geometric parameters derived on the bases of its microwave spectra.[18a] The geometry optimizations were carried out at the DFT/B3LYP[19] and ab initio/MP2[20] levels of theory by utilizing the 6-31G*, 6-311++G**, and the aug-cc-pvtz[21] basis sets in both series (Table 1). Optimization for aromatic heterocycles at correlated level is necessary, since HF optimizations either at the 3-21G or the 6-31G* levels predict significantly shorter distances for some bonds in the ring, compared to the experimental values.[2b,18b] Investigating the 1*H*−2*H* tautomerization, the optimized 2*H* as well as the TS structure (separating 1*H*- and 2*H*-imidazoles) were determined. Gas-phase species were considered at the MP2 and B3LYP levels using both the 6-31G* and the 6-311++G** basis sets. The geometries were reoptimized at the IEF-PCM/B3LYP/6-31G* level with chloroform ($CHCl_3$) and water solvents (see below). Energy minima and TS structures were certified by vibrational frequencies; all positive values for minima and one imaginary value for each TS. Geometric parameters and energy results are provided in Tables 2 and 3.

In studies for methyl-substituted N-heterocycles, molecular geometries were optimized at the B3LYP/6-31G* and B3LYP/aug-cc-pvtz levels in the gas-phase (Table 4). In-solution calculations started with geometry optimization for 12 tautomers of methyl-substituted N-heterocycles (Scheme 1) in the continuum dielectric approach.[22] The integral equation formalism[23a−c] was utilized in the polarizable continuum method[22a,b] (IEF-PCM), where the internal energy, E^s, and the electrostatic solute−solvent interaction energy, V_{elst} can be calculated as follows:

$$E^s_{int} = [\langle \Psi_s | H^\circ | \Psi_s \rangle] \quad (1a)$$

$$V_{elst} = [\langle \Psi_s | V_R^{sol} | \Psi_s \rangle] \quad (1b)$$

Here H° is the solute's Hamiltonian, V_R^{sol} is the solvent reaction field generated by the fully polarized solute in solution, and Ψ_s is the converged wave function of the solute obtained from the in-solution calculation. The cavity accommodating the solute was generated by overlapping spheres with Bondi values for atomic radii.[23d] Four solvents, namely $CHCl_3$, acetone (CH_3COCH_3), methanol (CH_3OH), and water were considered with dielectric constants of $\epsilon = 4.71$, 20.49, 32.61, and 78.39, respectively, at $T = 25$ °C.[24] Molecular geometries and energies were obtained through IEF-PCM/B3LYP/6-31G* geometry optimizations and IEF-PCM/B3LYP/6-311++G**//IEF-PCM/B3LYP/6-31G* single-point calculations, respectively (Tables 4−6). 1*H*−2*H* and 1*H*−4*H* tautomerizations for 1,2,3-triazole and 1,2,4-triazole, respectively, were also studied in different solvents including THF ($\epsilon = 7.43$) and CH_2Cl_2 ($\epsilon = 8.82$). Geometry optimization and frequency analysis were performed at the IEF-PCM/B3LYP/6-31G* level, single point energies were calculated at the IEF-PCM/MP2/6-31G*//IEF-PCM/B3LYP/6-31G* level (Table 7). Calculations were performed with the Gaussian 03 software[25] running at the Ohio Supercomputer Center.

Relative free energies, ΔG, were calculated for the methyl-substituted tautomers of N-heterocycles at the IEF-PCM level as

$$\Delta G = \Delta E^s_{int} + {}^1/_2 \Delta V_{elst} + \Delta G_{drc} \quad (2)$$

where ΔG_{drc} is the relative dispersion-repulsion-cavitation free energy provided by the Gaussian 03 version of the method.

Using the identity for the internal energy upon solvation, $E^s_{int} = (E^s_{int} - E^g) + E^g = E^{sg} + E^g$, where $E^{sg} = E^s_{int} - E^g$, and E^g is the optimized energy of the given molecule in the gas phase, the relationship $\Delta E^s_{int} = \Delta E^{sg} + \Delta E^g$ holds for the relative values. Table 4 summarizes the tautomerization energy in the gas phase, ΔE^g, whereas Table 5 provides the E^{sg} values themselves. While only the relative energies are relevant in the gas phase, the E^{sg} values reveal the different reactions of the tautomers to the solvent environment changing the solutes' geometry and electron distribution. E^{sg} is always a positive number, because the internal energy of a molecule necessarily

TABLE 2: Optimized Geometric Data for the 1*H*- and 2*H*-Imidazoles and the TS (1*H* to 2*H*) Structure in the Gas Phase, in Chloroform, and in Aqueous Solution[a]

Gas Phase

	MP2/6-31G*		B3LYP/6-31G*		MP2/6-311++G**		B3LYP/6-311++G**	
	2H	TS	2H	TS	2H	TS	2H	TS
N_1C_2	1.458	1.452	1.458	1.444	1.457	1.455	1.456	1.442
C_2N_3	1.458	1.350	1.458	1.346	1.457	1.352	1.456	1.346
N_3C_4	1.300	1.354	1.287	1.347	1.299	1.352	1.284	1.343
C_4C_5	1.468	1.396	1.478	1.397	1.474	1.402	1.478	1.398
C_5N_1	1.300	1.361	1.287	1.358	1.299	1.358	1.284	1.353
HN_1		1.277		1.289		1.287		1.298
HC_2	1.097	1.288	1.098	1.298	1.096	1.275	1.096	1.287
$N_1C_2N_3$	110.7	113.6	110.2	113.5	110.6	113.5	109.6	113.0
$C_2N_3C_4$	103.7	103.8	104.1	104.0	103.9	104.0	104.6	104.4
$N_3C_4C_5$	111.0	111.2	110.8	111.2	110.8	111.1	110.6	111.0
$C_4C_5N_1$	111.0	109.7	110.8	109.3	110.8	109.8	110.6	109.3
$C_5N_1C_2$	103.7	101.5	104.1	101.9	103.9	101.6	104.6	102.2
N_1HC_2		69.0		67.9		69.2		67.8
$HN_1C_2N_3$		109.1		108.7		108.4		108.6

IEF−PCM/B3LYP/6-31G*

	$CHCl_3$			water		
	1*H*	2*H*	TS	1*H*	2*H*	TS
N_1C_2	1.361	1.458	1.438	1.358	1.458	1.436
C_2N_3	1.322	1.458	1.348	1.325	1.458	1.350
N_3C_4	1.383	1.289	1.350	1.385	1.290	1.351
C_4C_5	1.373	1.476	1.395	1.374	1.475	1.395
C_5N_1	1.379	1.289	1.360	1.378	1.289	1.361
HN_1	1.011		1.295	1.012		1.299
HC_2		1.098	1.297		1.098	1.294
$N_1C_2N_3$	111.7	109.6	113.1	111.6	109.4	112.9
$C_2N_3C_4$	105.1	104.6	104.2	105.1	104.8	104.4
$N_3C_4C_5$	110.6	110.6	111.0	110.4	110.5	110.9
$C_4C_5N_1$	105.1	110.6	109.1	105.2	110.5	109.1
$C_5N_1C_2$	107.5	104.6	102.4	107.7	104.8	102.6
N_1HC_2			67.4			67.3
$HN_1C_2N_3$			108.5			108.4

[a] For the gas-phase 1H-imidazole geometry, see Table 1. Distances in Å, angles in deg.

increases if its geometry and/or charge distribution changes as compared to its optimized structure in the isolated form.

Monte Carlo simulations[26] for calculating relative solvation free energies by utilizing the free energy perturbation method (FEP)[27] were performed with the BOSS 3.6 program[28] in NpT (isothermal−isobaric) ensembles at $T = 25$ °C and $p = 1$ atm. For the aqueous solution model, a water box including 506 TIP4P water molecules[29] and a single solute were considered. For solution models in $CHCl_3$, CH_2Cl_2 and THF, a single solute and 264 solvent molecules were considered. Using the CH_x ($x = 1, 2$) united atom model for these solvent molecules, 4, 3, and 5 point-models for $CHCl_3$, CH_2Cl_2, and THF were applied, respectively. The corresponding solvent boxes were available from the program library.[28] The solute geometries were previously optimized quantum-mechanically at the IEF-PCM level in the given solvent, and remained rigid throughout the MC moves. The interaction energy of the solution elements was calculated by using the 12−6−1 OPLS-AA pair potential.[30] In this approach, each atom in the solution (with CH_x united atoms for the solvents) is represented by a set of charge and steric parameters. Steric OPLS parameters were taken from the program library. The solvent−solvent cutoff (RCUT) and the solute−solvent cutoff (SCUT) were set to 9.75 and 12 Å, respectively. By using the ICUT = 2 option, every solvent molecule is seen by the solute if the solvent's central atom is within a sphere of $R = \text{SCUT} = 12$ Å around any solute atom. Using SCUT as large as 12 Å, the relative interaction energy between the solute tautomers and the solvent molecules out of the cutoff region is negligible.[12,15a] Other simulation parameters were close to those used in recent calculations.[15]

As was discussed earlier,[15a,31] in cases when an 12−6−1 OPLS-AA effective pair-potential is used for atomic interactions of rigid molecules in MC, the solute's internal energy is considered to be constant calculated by the IEF-PCM method at the optimized geometry. Then the atomic point-charges for the solute are to be chosen for mimicking the in-solution charge distribution at the in-solution relevant geometries. To this aim, the atomic charge parameters for the solutes were fitted, using the CHELPG procedure,[32] to the in-solution molecular electrostatic potential (ELPO) generated by the IEF-PCM/B3LYP/6-31G* wave function, or in some cases to the in-solution ELPO derived by means of the IEF-PCM/B3LYP/6-311++G** single-point wave function. The solvent parameters were fitted to produce good solvent densities and heats of vaporization.[26d,29]

The FEP method was applied along nonphysical perturbation paths, by developing and annihilating hydrogen atoms at the required sites for the given tautomers. Since the free energy is a state function, the final result is independent of the perturbation path, which can be even a nonphysical route. The interaction potential parameters and geometry parameters were calculated as a linear function of the reference parameters defined at the two end-points of the perturbation step. Using double-wide sampling, the perturbation step parameter was chosen so that the free energy increments did not exceed 1 kcal/mol. 3500 and

TABLE 3: Energies and Free Energies of 2*H*-imidazole and the TS (1*H* to 2*H*) Structures Relative to Those for 1*H*-imidazole in the Gas Phase, in Chloroform, and in Aqueous Solution[a]

2*H*-Imidazole						
	ΔE^{g}	ΔZPE	$\Delta(H(T) - \mathrm{ZPE})$	$-T\Delta S(T)$	$\Delta G(\mathrm{solv})$	ΔG_{tot}
Gas Phase						
B3LYP/6-31G*	13.84	−0.52	−0.02	−0.03		13.27
MP2/6-31G*//	15.02					14.45
B3LYP/6-31G*						
MP2/6-31G*	14.87	−0.50	−0.02	−0.03		14.32
B3LYP/6-311++G**//	17.54					16.97
B3LYP/6-31G*						
B3LYP/6-311++G**	17.52	−0.67	0.00	−0.05		16.80
MP2/6-311++G**	17.11	−0.26	−0.10	0.09		16.84
	ΔE^{s}_{int}	ΔZPE	$\Delta(H(T) - \mathrm{ZPE})$	$-T\Delta S(T)$	$\Delta G(\mathrm{solv})$	ΔG_{tot}
$CHCl_3$						
B3LYP/6-31G*	13.00	−0.57	0.01	−0.07	2.75	15.12
MP2/6-31G*//	14.57				2.61	16.55
B3LYP/6-31G*						
B3LYP/6-311++G**//	16.63				2.34	18.34
B3LYP/6-31G*						
Water						
B3LYP/6-31G*	11.81	−0.41	0.00	−0.05	5.11	16.46
MP2/6-31G*//	13.85				4.64	18.03
B3LYP/6-31G*						
B3LYP/6-311++G**//	15.33				4.60	19.47
B3LYP/6-31G*						
TS (1*H* to 2*H*)						
	ΔE^{g}	ΔZPE	$\Delta(H(T) - \mathrm{ZPE})$	$-T\Delta S(T)$	$\Delta G(\mathrm{solv})$	ΔG_{tot}
Gas Phase						
B3LYP/6-31G*	49.38	−3.09	−0.16	0.22		46.35
MP2/6-31G*//	49.50					46.47
B3LYP/6-31G*						
MP2/6-31G*	49.47	−3.02	−0.16	0.22		46.51
B3LYP/6-311++G**//	50.66					
B3LYP/6-31G*						
B3LYP/6-311++G**	50.68	−3.16	−0.15	0.21		47.58
MP2/6-311++G**	48.26	−2.77	−0.25	0.35		45.59
	ΔE^{s}_{int}	ΔZPE	$\Delta(H(T) - \mathrm{ZPE})$	$-T\Delta S(T)$	$\Delta G(\mathrm{solv})$	ΔG_{tot}
$CHCl_3$						
B3LYP/6-31G*	48.53	−3.19	−0.14	0.18	2.64	48.02
MP2/6-31G*//	48.62				2.83	48.30
B3LYP/6-31G*						
B3LYP/6-311++G**//	49.70				2.48	49.03
B3LYP/6-31G*						
Water						
B3LYP/6-31G*	47.34	−3.27	−0.12	0.16	4.93	49.04
MP2/6-31G*//	47.29				5.28	49.34
B3LYP/6-31G*						
B3LYP/6-311++G**//	48.35				4.84	49.96
B3LYP/6-31G*						

[a] Energies in kcal/mol. Corrections for free energies in single-point calculations utilize the values obtained for the corresponding geometry optimization.

5000 K configurations were considered in the equilibration and averaging phases, respectively. Results are summarized in Tables 7 and 8.

For studying the possibility of the intermolecular 1*H*−3*H* proton relocation, molecular dynamics simulations for a tetramer of 4-methylimidazole were carried out in $CHCl_3$. Four solvent molecules were placed in a box of 355 $CHCl_3$ molecules and *NpT* molecular dynamics simulations were performed at $T =$ 298 K and 1 atm by utilizing the Tripos force-field in the Sybyl 6.9.1 software.[33] Two solute molecules were arranged in an antiparallel stacking position at the onset of simulations, whereas the other two were placed in perpendicular planes regarding the first two at distances of 7−8 Å between the closest nonbonded heavy atoms. Atomic charges were set to values used in the MC simulations both for the solute and solvent molecules. Using the SHAKE algorithm for all single bonds, a 2 fs time step, periodic boundary conditions and a nonbonded cutoff of 10 Å were applied in the 500 ps long equilibration phase followed by a production phase of 1 ns. Previous experiences with the CHELPG charges indicate[15a,31] that the electrostatic interaction energy is negligible for neutral, moderately polar molecules beyond this cutoff limit. The list of the nonbonded pairs was updated every 10 steps; temperature and pressure coupling constants were set to 0.1 ps each.

Results and Discussion

Gas-Phase Structures and Tautomerism. The structure of 1*H*-imidazole was determined by utilizing data from the

TABLE 4: Relative Internal Energies in the Gas-phase at the B3LYP Level with the 6-31G* and the aug-cc-pvtz Basis Sets[a]

	ΔE^g (gas phase)	
	6-31G*	aug-cc-pvtz
4-Me-imidazole-1*H*	0.00	0.00
4-Me-imidazole-3*H*	0.09	0.52
3-Me-pyrrazole-1*H*	0.00	0.00
3-Me-pyrrazole-2*H*	−0.05	0.29
4-Me-triazole-1,2,3−1*H*	0.00	0.00
4-Me-triazole-1,2,3−2*H*	−4.72	−4.08
4-Me-triazole-1,2,3−3*H*	−0.20	0.12
3-Me-triazole-1,2,4−1*H*	0.00	0.00
3-Me-triazole-1,2,4−2*H*	−0.38	−0.10
3-Me-triazole-1,2,4−4*H*	6.41	6.14
5-Me-tetrazole-1*H*	0.00	0.00
5-Me-tetrazole-2*H*	−2.20	−1.90

[a] Energies in kcal/mol.

microwave spectra by Christen et al.[18a] Optimized geometries for this molecule at the B3LYP and MP2 levels with the 6-31G*, 6-311++G**, and aug-cc-pvtz basis sets can be compared with the experimental values in Table 1. The calculations, in any combination of the two theoretical levels and basis sets, provide geometries close to the experimental structure. The root-mean-square (rms) values for the ring distances and the bond angles still indicate the best overall agreement between the B3LYP/6-31G* structure and the experimental one. The good performance of the B3LYP/6-31G* optimization may be attributed to an error cancellation due to the smaller basis set and the applied approach for considering the electron correlation. Use of the B3LYP level and the 6-31G* basis set in geometry optimizations is advantageous, however, from a practical point-of-view: the calculations speed up by a factor of about 2 compared to the MP2 level.

All five-member N-heterocycles in the present study exhibit tautomeric forms. The N to N proton relocation does not lead to an observable new structure with unsubstituted rings in several cases. The possibility for a tautomeric equilibrium becomes obvious, however, if ring substituents destroy the equivalence of the N-sites. It is then a central problem; what is the reaction path for the tautomerism?

In the gas-phase, the possible reaction routes are either the intermolecular proton jump throughout collision or an intramolecular rearrangement. In dilute solution, the solvent-mediated tautomerism may be a favorable reaction path if the solvent can take up the proton at the N−H site, and, after one or more proton jumps involving the solvent molecule(s), returns the proton to the basic N-site. Such a mechanism is readily available in protic solvents (water, CH_3OH, etc.) but not in solvents like $CHCl_3$ (see section Equilibration Mechanism in Solution). The intramolecular path is still a reasonable variant in nonpolar solvents, and this reaction route was studied first in the gas phase.

If there are two neighboring nitrogens in the ring (pyrrazole, triazoles, tetrazole) then the 1,2-prototropic shift is a plausible intramolecular mechanism. Alkorta and Elguero[10] found a transition state (TS) with relative energy of 51 kcal/mol at the B3LYP/6-31G* level for pyrrazole and with the proton above the ring plane and out of the ring area.

There are no neighboring nitrogens, however, in imidazole and for the 1−4 tautomerization in 1,2,4-triazole (or for the 2−4 tautomerization with a substituent on the ring carbon atom). All our attempts have failed for finding a transition state for a one-step, 1,3 proton relocation for imidazole in the gas phase. No TS structure was found with a proton somewhere between the two nitrogens either for the isolated imidazole or the imidazole···$CHCl_3$ complex. This latter dimer was studied in order to explore whether the $CHCl_3$ facilitates the proton relocation by involvement of the solvent in a nonpolar medium.

Next, consecutive 1,2-prototropic shifts were considered for imidazole. In this postulated mechanism, the $H(N_1)$ proton moves to the C_2 atom in the first step and the second step is the forward motion of the proton from C_2 to N_3. The same mechanism was expected for the tautomerization in 1,2,4-triazole.

Table 2 summarizes the geometric results for 2*H* imidazole and the TS structure separating the 1*H* and 2*H* tautomers. Lynch and Truhlar found[34] that the TS structures in simple reactions could be predicted erroneously at the B3LYP level. The present study results in similar geometric parameters at the MP2 and the B3LYP levels, with both the 6-31G* and 6-311++G** basis sets. The 1*H* (see Table 1) and the 2*H* structures have C_s and C_{2v} symmetries, respectively, whereas the TS structure is without any symmetry. The relocating H atom stays above the general ring plane, which is still almost entirely flat even in the TS, but leans outside the ring area in a fashion similar to that in the pyrrazole TS.[10] (The TS geometry may be characterized by the $HN_1C_2N_3$ torsion angle: values larger or smaller than 90° refer to a hydrogen leaning outside or inside the ring area, respectively. The $HN_1C_2N_3$ torsion angle was calculated at 108−109° depending upon the method.)

Relative energies (Table 3) reveal basis set effects rather than dependence on the method. ΔE_{int} is 14−15 and 17−18 kcal/mol for 2*H* imidazole in the gas phase, as calculated with the 6-31G* and 6-311++G** basis sets, respectively. Interestingly, the relative TS energies are in a fairly narrow range of 48.3−50.7 kcal/mol in all present calculations. Relative free energies, ΔG_{tot} (calculated in the rigid rotor-harmonic oscillator approximation[35a]), are 0.3−0.7 and about 3 kcal/mol smaller than the energy values for 2*H* imidazole and TS, respectively, mainly due to the reduction of the zero point energy (ZPE) term. Thermal corrections for enthalpy and entropy mostly cancel each other both for the 2*H* and the TS structures.

Geometry optimizations for tautomers of methyl-substituted N-heterocycles led to planar ring structures both at the B3LYP/6-31G* and the B3LYP/aug-cc-pvtz levels. The molecules have symmetries very close to C_s with mirror plane through the heavy atoms. Relative energies are summarized in Table 4. The basis set effect, 6-31G* vs aug-cc-pvtz, is generally small. Changes in the ΔE^g values are no more than 0.4 kcal/mol (with one exception), but the variations were large enough to change the sign for 3-Me-pyrrazole-2*H* and 4-Me-1,2,3-triazole-3*H*. Thus, in cases when ΔE^g is within a range of only a few tenths of a kilocalorie per mole, the sign of the relative energies is uncertain.

A rough estimate of the equilibrium composition on the basis of the ΔE^g values predicts detectable fractions for both tautomers of 4-Me-imidazole and 3-Me-pyrrazole. Only the 2*H* tautomer is expected to appear for 4-Me-1,2,3-triazole and the 2*H* tautomer must be the overwhelming form for 5-Me-tetrazole. Two tautomers, the 1*H* and the 2*H* forms for 3-Me-1,2,4-triazole are expected in comparable concentrations in the gas phase, whereas the 4*H* tautomer is too high in relative energy for participating in the equilibrium. The above predictions have been made without considering corrections for the zero-point vibrational energies and the thermal contributions. These terms amount to a few tenths of a kilocalorie per mole (see Table 7 for triazole tautomers in different solvents) and are in the range of the uncertainty for ΔE^g due to the basis set effect.

TABLE 5: IEF-PCM/B3LYP/6-31G* Relative Energy and Free Energy Terms in Different Solvents[a]

	$CHCl_3$				CH_3COCH_3			
	E^{sg}	$^1/_2V_{elst}$	$^1/_2\Delta V_{elst}$	ΔG	E^{sg}	$^1/_2V_{elst}$	$^1/_2\Delta V_{elst}$	ΔG
4Mim1H	1.66	−9.62	0.00	0.00	3.28	−14.28	0.00	0.00
4Mim3H	1.80	−10.03	−0.39	−0.13	3.49	−14.81	−0.53	−0.16
3Mpy1H	1.06	−7.71	0.00	0.00	2.02	−11.09	0.00	0.00
3Mpy2H	1.11	−7.89	−0.18	−0.13	2.08	−11.31	−0.22	−0.16
4Mtr123,1H	2.18	−11.66	0.00	0.00	4.25	−17.23	0.00	0.00
4Mtr123,2H	0.85	−7.96	3.70	−2.37	1.51	−10.97	6.26	−1.21
4Mtr123,3H	2.22	−11.98	−0.32	−0.43	4.18	−17.44	−0.21	−0.40
3Mtr124,1H	1.57	−10.78	0.00	0.00	2.93	−15.39	0.00	0.00
3Mtr124,2H	1.54	−10.67	0.11	−0.25	2.87	−15.21	0.18	−0.21
3Mtr124,4H	3.28	−15.54	−4.76	3.46	6.44	−23.43	−8.04	1.98
5Mte1H	2.69	−14.38	0.00	0.00	5.09	−21.03	0.00	0.00
5Mte2H	1.39	−10.44	3.94	0.35	2.56	−14.72	6.31	1.48
	CH_3OH				water			
	E^{sg}	$^1/_2V_{elst}$	$^1/_2\Delta V_{elst}$	ΔG	E^{sg}	$^1/_2V_{elst}$	$^1/_2\Delta V_{elst}$	ΔG
4Mim1H	3.57	−15.03	0.00	0.00	3.92	−15.87	0.00	0.00
4Mim3H	3.80	−15.57	−0.54	−0.16	4.15	−16.42	−0.55	−0.15
3Mpy1H	2.17	−11.59	0.00	0.00	2.36	−12.16	0.00	0.00
3Mpy2H	2.23	−11.76	−0.17	−0.09	2.44	−12.37	−0.21	−0.10
4Mtr123,1H	4.61	−18.12	0.00	0.00	4.73	−18.69	0.00	0.00
4Mtr123,2H	1.62	−11.41	6.71	−1.02	1.73	−11.86	6.83	−0.92
4Mtr123,3H	4.51	−18.24	−0.12	−0.35	4.91	−19.26	−0.57	−0.49
3Mtr124,1H	3.19	−16.15	0.00	0.00	3.51	−17.06	0.00	0.00
3Mtr124,2H	3.11	−15.93	0.22	−0.19	3.37	−16.72	0.34	−0.11
3Mtr124,4H	6.98	−24.66	−8.51	1.79	7.72	−26.30	−9.24	1.51
5Mte1H	5.52	−22.16	0.00	0.00	6.03	−23.40	0.00	0.00
5Mte2H	2.75	−15.37	6.79	1.71	2.95	−16.03	7.37	1.96

[a] Energies in kcal/mol. $\Delta G = \Delta E^s_{int} + {}^1/_2\Delta V_{elst} + \Delta G_{drc}$, $\Delta E^s_{int} = \Delta E^{sg} + \Delta E^g$. ΔE^g from Table 4, ΔG_{drc} from Table 6.

The gas-phase results provide starting points for evaluating the equilibrium conditions in solution both for structures and energies. Consideration of the ΔE^g, E^{sg} and ΔE^{sg} terms helps point out the inherent energy difference, existing already for the isolated tautomers (ΔE^g), and the absolute (E^{sg}) and relative (ΔE^{sg}) increases of the internal energy upon solvation.

Tautomeric Equilibrium in Solution. *Continuum Solvent Model.* Molecular geometries of methyl-substituted five-member N-heterocycles underwent minor changes relative to the gas-phase values upon B3LYP/6-31G* optimization in solution. Results obtained for imidazole (Tables 1 and 2) are typical for other heterocycles in the present study. Generally systematic deviations for the bond-length and bond-angle parameters were calculated for the series: gas-phase molecule, solute in $CHCl_3$, and solute in water. Bond lengths changed by some thousandths of an ångstrom; bond angles varied by some tenths of a degree. Values both increased and decreased, but the changes were consistent within the series. This is also true for the transition state geometry, at least for imidazole. Relative ZPE's and thermal corrections terms are moderate for local energy minimum structures either for imidazole (Table 3) or for different isomers and tautomers of triazole (Table 7).

Energy terms calculated in $CHCl_3$, CH_3COCH_3, CH_3OH, and water are summarized in Table 5. The general conclusion from this table is that the energy components change monotonically as a function of the dielectric constant (see the Methods section) for the present molecules. The more polar the solvent, the larger the polarization, and E^{sg} increases with solvent polarity. The readiness for polarization is, however, a feature of the solute's chemical structure, and is different not only for different solutes but also for their tautomers. Accordingly, the E^{sg} values form a range for any tautomeric set. For example, the values vary between 0.8 and 2.2 kcal/mol for 4-Me-1,2,3-triazole in $CHCl_3$, thus ΔE^{sg} can be as large as 1.4 kcal/mol in this solvent. The E^{sg} values for the same molecule are in the range of 1.7−4.9 kcal/mol in water, with a maximum ΔE^{sg} of 3.2 kcal/mol.

The $^1/_2V_{elst}$ values (half of the solute−solvent electrostatic interaction energy) are all negative, thus the interaction of the mutually polarized solute and solvent stabilizes the system formed by immersing a mole of an N-heterocycle in a dielectric medium. The $^1/_2V_{elst}$ values vary between about −8 kcal/mol in $CHCl_3$ to −26 kcal/mol in water. Furthermore, the solute−solvent electrostatic interaction energies could be very different for tautomeric pairs even in a given solvent. The $^1/_2\Delta V_{elst}$ values are considerable for the tetrazole tautomers, but they are small for either the methylated imidazole or pyrrazole. The $^1/_2\Delta V_{elst}$ values are small for the 1*H* and 3*H* tautomers of 4-Me-1,2,3-triazole as well as for the 1*H* and 2*H* tautomers of 3-Me-1,2,4-triazole in any solvent. In contrast, the $^1/_2\Delta V_{elst}$ values are strongly positive and negative for the 4-Me-1,2,3-triazole-2*H* and 3-Me-1,2,4-triazole-4*H* tautomers, respectively.

The in-solution equilibrium depends on, however, the total relative free energy, ΔG, (see eq 2). Values for ΔG were small in absolute value for the imidazole and pyrrazole tautomers in any studied solvent. Because of the uncertainties discussed regarding the gas-phase equilibria, one may only say on the basis of the present results that both tautomers might be detectable for these heterocycles in a wide range of solvents. For 4-Me-1,2,3-triazole, the 2*H* tautomer must be prevalent in low polarity solvents, but the 3*H* tautomer will form a remarkable fraction in CH_3COCH_3 and CH_3OH and even more in water. Two prevailing tautomers, 1*H* and 2*H*, are expected for 3-Me-1,2,4-triazole. The 4*H* tautomer must be of a small fraction even in aqueous solution. For the 5-Me tetrazole, the

TABLE 6: IEF-PCM/B3LYP/6-311++G//IEF-PCM/B3LYP/6-31G* Relative Energy and Free Energy Terms in Different Solvents[a]**

	$CHCl_3$				CH_3COCH_3			
	ΔE^s_{int}	$^1/_2V_{elst}$	ΔG_{drc}	ΔG	ΔE^s_{int}	$^1/_2V_{elst}$	ΔG_{drc}	ΔG
4Mim1H	0.00	−10.03	0.00	0.00	0.00	−15.09	0.00	0.00
4Mim3H	0.70	−10.57	0.05	0.21	0.78	−15.78	0.07	0.16
3Mpy1H	0.00	−8.00	0.00	0.00	0.00	−11.61	0.00	0.00
3Mpy2H	0.34	−8.34	0.04	0.04	0.36	−12.04	0.05	−0.02
4Mtr123,1H	0.00	−12.39	0.00	0.00	0.00	−18.55	0.00	0.00
4Mtr123,2H	−6.18	−8.33	−0.02	−2.14	−7.86	−11.55	−0.02	−0.88
4Mtr123,3H	0.20	−12.82	0.05	−0.18	0.06	−18.87	0.07	−0.19
3Mtr124,1H	0.00	−11.32	0.00	0.00	0.00	−16.28	0.00	0.00
3Mtr124,2H	−0.12	−11.26	0.06	0.00	−0.18	−16.18	0.06	−0.02
3Mtr124,4H	8.63	−16.57	0.10	3.48	10.83	−25.37	0.10	1.84
5Mte1H	0.00	−15.37	0.00	0.00	0.00	−22.70	0.00	0.00
5Mte2H	−4.01	−11.03	−0.09	0.24	−5.45	−15.68	−0.10	1.47
	CH_3OH				water			
	ΔE^s_{int}	$^1/_2V_{elst}$	ΔG_{drc}	ΔG	ΔE^s_{int}	$^1/_2V_{elst}$	ΔG_{drc}	ΔG
4Mim1H	0.00	−15.88	0.00	0.00	0.00	−16.78	0.00	0.00
4Mim3H	0.81	−16.63	0.06	0.12	0.82	−17.57	0.08	0.11
3Mpy1H	0.00	−12.16	0.00	0.00	0.00	−12.76	0.00	0.00
3Mpy2H	0.35	−12.53	0.07	0.05	0.37	−13.20	0.09	0.02
4Mtr123,1H	0.00	−19.56	0.00	0.00	0.00	−20.21	0.00	0.00
4Mtr123,2H	−8.17	−12.04	−0.02	−0.67	−8.21	−12.53	−0.03	−0.56
4Mtr123,3H	0.02	−19.78	0.07	−0.13	0.34	−20.94	0.09	−0.30
3Mtr124,1H	0.00	−17.09	0.00	0.00	0.00	−18.07	0.00	0.00
3Mtr124,2H	−0.20	−16.98	0.06	−0.03	−0.21	−17.85	0.07	0.08
3Mtr124,4H	11.19	−26.76	0.10	1.62	11.71	−28.61	0.13	1.30
5Mte1H	0.00	−23.97	0.00	0.00	0.00	−25.37	0.00	0.00
5Mte2H	−5.73	−16.39	−0.10	1.75	−6.09	−17.15	−0.13	2.00

[a] Energies in kcal/mol. $\Delta G = \Delta E^s_{int} + {}^1/_2\Delta V_{elst} + \Delta G_{drc}$.

1*H* form may have the larger population in $CHCl_3$, and becomes dominant in more polar solvents. It should be noted that the solvent effect causes the largest shift in the equilibrium composition for this molecule within the present series. As calculated, the 2-H tautomer is the overwhelming form in the gas phase, whereas the solvent effect reverses the tautomeric preference and makes the 2-H structure a minor component in solvents with a dielectric constant of at least 20 (CH_3COCH_3).

The basis set effect on the relative energies was studied by performing IEF-PCM/B3LYP/6-311++G**//IEF-PCM/B3LYP/6-31G* single-point calculations (Table 6). The trends for the energy changes have been reserved at this level, although nonnegligible numerical deviations were obtained. The $^1/_2V_{elst}$ values are consistently more negative with the 6-311++G** basis set compared to the 6-31G* set in any solvent (Tables 5 and 6). Thus, the larger basis set makes the solute−solvent interaction more negative for any solute. It is to be emphasized that only the relative values have an influence on the total relative free energy, thus a constant shift of the $^1/_2V_{elst}$ values calculated with different basis sets would be irrelevant in this respect. The increased basis set changes, however, even the relative values, $^1/_2\Delta V_{elst}$. The sign of this term has been always maintained as calculated with the two basis sets, but the absolute values become larger with the extended basis set in cases of strong solute−solvent interactions. For example, the $^1/_2\Delta V_{elst}$ value of 7.37 kcal/mol for the tetrazole derivative with the 6-31G* set in water increases to 8.22 kcal/mol with the larger basis set. For 3-Me-1,2,4-triazole-4*H*, the corresponding values are −9.24 and −10.54 kcal/mol. Interestingly, however, the total ΔG values are in good agreement from the two sets. The sign of ΔG is still uncertain in the ±0.2 kcal/mol range, but total relative free energies are close to each other for other cases from the two sets of calculations, and make similar predictions for the largest and lowest populations for the tautomers in the equilibrium mixture. This important finding becomes possible due to a compensation effect: the more negative $^1/_2\Delta V_{elst}$ value entails a more positive ΔE^s_{int} value. For making the solute−solvent interaction energy more negative, the solute must be more polarized. The polarization of the solute increases, however, its internal energy. Thus, E^s_{int} increases, and ΔE^s_{int} also increases with respect to a reference structure.

Explicit Solvent, Monte Carlo Simulations. Although relative free energies can be determined by using a continuum solvent approximation, this approach is unable to explore the solution structure around the solute. Knowledge of the structure of the first and second solvent shells is important for understanding the mechanisms of chemical transformations including tautomerism and more complicated reactions. To obtain these information, consideration of an explicit solvent model is required. In the ideal case, both solvent models provide the same energy results. Comparisons from previous studies suggest, however, that neither quantitative, nor even qualitative agreement can be reached. Cox et al. found opposite preference for the 1,2,3- and 1,2,4-triazole tautomers in aqueous solution by using explicit and continuum solvent models.[7a] Also disagreement was found in results with the two solvent models by Parchment et al. studying the tautomeric equilibrium for 3-hydroxypyrrazole in aqueous solution.[6a] Essential differences were obtained by Nagy et al. for the zwitterionic−neutral form equilibria of substituted phenylethylamines in aqueous solution.[12,16] In contrast, good agreement was reached by applying the two methods for hydroxyisoxazoles by Woodcock et al.[2a] Thus, test calculations seemed to be necessary and were

TABLE 7: Comparison of the IEF-PCM/B3LYP/6-31G*, IEF-PCM/MP2/6-31G*//IEF-PCM/B3LYP/6-31G* and Monte Carlo Relative Solvation and Total Free Energies for the Tautomerizations of 1,2,3-triazole and 1,2,4-triazole in Different Solvents[a]

1,2,3-Triazole, 1*H* to 2*H*

	IEF−PCM[b]							MC[c] CHELPG	
	B3LYP/6-31G*				MP2/6-31G*//B3LYP/6−31G*				
	ΔE^s_{int}	ΔG_{th}	ΔG(solv)	ΔG	ΔE^s_{int}	ΔG(solv)	ΔG	ΔG(solv)	ΔG
$CHCl_3$	−6.06	0.28	3.90	−1.88	−6.34	4.18	−1.88	1.72 ± 0.05	−4.06
CH_2Cl_2	−6.81	0.23	5.33	−1.25	−7.10	5.70	−1.18	4.49 ± 0.07	−2.09
water	−7.98	0.18	7.43	−0.37	−8.30	7.92	−0.20	9.10 ± 0.10	1.30

1,2,4-Triazole, 1*H* to 4*H*

	IEF−PCM[b]							MC[c] CHELPG	
	B3LYP/6-31G*				MP2/6-31G*//B3LYP/6-31G*				
	ΔE^s_{int}	ΔG_{th}	ΔG(solv)	ΔG	ΔE^s_{int}	ΔG(solv)	ΔG	ΔG(solv)	ΔG
$CHCl_3$	8.61	−0.38	−4.83	3.40	9.09	−5.53	3.18		
THF	9.29	−0.34	−6.13	2.82	9.98	−7.08	2.56	−3.83 ± 0.10	5.12
water	11.11	−0.28	−9.39	1.44	12.42	−11.00	1.14	−10.41 ± 0.12	0.42

[a] Energies in kcal/mol. $\Delta G = \Delta E^s_{int} + \Delta G_{th} + \Delta G$(solv). [b] Continuum dielectric solvent results from IEF−PCM/B3LYP/6-31G* calculations. [c] Atomic charges for MC simulations were derived by ELPO fitting of the PCM/B3LYP/6-31G* wave function, using the CHELPG charge-fitting procedures. ΔE^s_{int} from IEF−PCM/B3LYP/6-31G* results.

performed in the present study for selected N-heterocycles with available experimental results.

Table 7 summarizes the results of the IEF-PCM and the MC calculations for the 1*H* to 2*H* and 1*H* to 4*H* tautomeric equilibria for 1,2,3- and 1,2,4-triazoles, respectively. Total relative free energies in the FEP/MC approach utilize the IEF-PCM/B3LYP/6-31G* relative internal energies calculated after geometry optimization in different solvents and the corresponding thermal corrections. Thus, differences in ΔG's are merely due to differences in the ΔG(solv) terms.

ΔG's are all negative for the 1*H* to 2*H* tautomeric transformation for 1,2,3-triazole in the three different solvents from continuum calculations. A ΔG value of −1.88 kcal/mol in $CHCl_3$ predicts 96% of the 2*H* tautomer in the equilibrium mixture at $T = 298$ K. NMR results by Lunazzi et al.[9a] show 80% 2*H* tautomer in CD_2Cl_2 at room temperature, in good agreement with our predicted 89% 2*H* fraction, as derived from $\Delta G = -1.25$ kcal/mol. In aqueous solution, the predicted ratio of 2*H*:1*H* was 65:35.

Whereas ΔG varies between −1.88 and −0.37 kcal/mol for three solvents, the corresponding MC values vary in the range of −4.06 to +1.30 kcal/mol. The derived compositions are in qualitative agreement with those from the IEF-PCM calculations in slightly polar solvents. In both $CHCl_3$ and CH_2Cl_2, the 2*H* form exists almost exclusively. The derived 2*H* fraction in CH_2Cl_2 is 97% at $T = 298$ K, still not too far from the experimental estimate of 80%. The essential difference was calculated in aqueous solution. MC predicts 10% 2*H* tautomer as compared to 65% from IEF-PCM/B3LYP/6-31G* calculations.

There is no experimental result available for the equilibrium in aqueous solution. Using two circumstantial arguments, Albert and Taylor[9b] estimate an equilibrium mixture with 65−67% 2*H* tautomer. Using the explicit solvent model, Cox et al.[7a] concluded a similar composition in aqueous solution, but their reaction field continuum model predicted a reversed composition. This example confirms that the continuum and explicit solvent models may provide qualitatively different results.

In FEP/MC calculations, the charge parameters play the decisive role for polar systems. Former studies suggest that the CHELPG fit to the molecular electrostatic potential generated by the in-solution 6-31G* wave function may result in atomic charges that predict overly polar solute charge distributions.[15a,16,31] As a result, the ΔG(solv) term may be overestimated for tautomers with considerable differences in the dipole moment. Indeed, the dipole moment was calculated at 6.2 and 0.1 D for the 1*H* and 2*H* tautomers, respectively. RESP charges lead to a less polar charge distribution for the solute and predict the free energy of hydration closer to the experimental value than the CHELPG based calculations.[15a]

Since some concern emerged regarding the reliability of the PCM/B3LYP/6-31G* conformer/tautomer relative energies in a previous study,[16] Table 7 includes the IEF-PCM/MP2/6-31G*//IEF-PCM/B3LYP/6-31G* single point results, as well. Compared to the B3LYP values, both the ΔE^s_{int} and ΔG(solv) terms increase in absolute values, but because of their opposite signs the net changes in ΔG's are only a few tenths of a kcal/mol, and the derived compositions are very similar to those from the IEF-PCM/B3LYP/6-31G* calculations.

IEF-PCM and FEP/MC results are in qualitative agreement for the 1*H* to 4*H* tautomerism of 1,2,4-triazole. Large positive ΔE^s_{int} values are partially compensated by negative ΔG(solv) terms giving positive total ΔG's, using any method for any solvent. ΔG_{th} values are relatively small. In their experiments, Lunazzi et al.[9a] did not observe the 4-H tautomer in THF. Our ΔG values of 2.56−2.82 kcal/mol calculated in the IEF-PCM approximation and an even more positive value of 5.12 kcal/mol from a FEP/MC estimation predict less than 1% population for the 4*H* conformer at $T = 298$ K in the mixture. Thus, experimental and theoretical results are in good agreement in this case.

Considering the aqueous solution, the IEF-PCM methods still predict at least 87% of the 1*H* tautomer. The value from the MC calculations is reduced to 67%. The reason again is the ΔG(solv) term, which is more negative by 1 kcal/mol at the MC than at the IEF-PCM/B3LYP level. In this case, in contrast to the case of 1,2,3-triazole, however, the ΔE^s_{int} term remains the prevalent one, and its positive sign determines ΔG for all studied solvents. Overall, the present preliminary studies indicate that in cases where ΔG is less than 1 kcal/mol in absolute value, differences in the ΔG(solv) term calculated by the two solvent models may reverse ΔG itself.

The IEF-PCM and FEP/MC results for some methyl-substituted tautomers of the N-heterocycles are compared in Table 8. Since the ΔG_{th} values are generally small for the present ring systems and cause only a shift of the total relative free energies calculated with the continuum and explicit solvent

TABLE 8: Comparison of the IEF-PCM/B3LYP/6-31G* and Monte Carlo Relative Solvation and Total Free Energies for Methyl-Substituted Five-Member N-Heterocycles in Chloroform and Aqueous Solution[a]

	$CHCl_3$				water			
	IEF-PCM[b]		MC[c]		IEF-PCM[b]		MC[c]	
	ΔG(solv)	ΔG	ΔG(solv)	ΔG	ΔG(solv)	ΔG	ΔG(solv)	ΔG
4-Me-imidazole								
1*H* to 3*H*								
6-31G*	−0.36	−0.13	−0.22 ± 0.12	0.01	−0.47	−0.15	−0.18 ± 0.15	0.14
6-311++G**// 6-31G*	−0.49	0.21	−0.54 ± 0.12	0.16	−0.71	0.11	1.04 ± 0.14	1.92
4-Me-triazole-1,2,3								
1*H* to 2*H*								
6-31G*	3.68	−2.37	2.08 ± 0.06	−3.97	6.80	−0.92	8.75 ± 0.11	1.03
6-311++G**// 6-31G*	4.04	−2.14			7.65	−0.56		
1*H* to 3*H*								
6-31G*	−0.27	−0.43	−0.26 ± 0.07	−0.42	−0.48	−0.50	1.18 ± 0.13	1.16
6-311++G**// 6-31G*	−0.38	−0.18			−0.64	−0.30		
5-Me-tetrazole								
1*H* to 2*H*								
6-31G*	3.85	0.35	2.06 ± 0.04	−1.44	8.24	1.96	5.59 ± 0.12	0.31
6-311++G**// 6-31G*	4.25	0.24			8.09	2.00		

[a] Energies in kcal/mol. $\Delta G = \Delta E^s + \Delta G$(solv). [b] The IEF−PCM ΔG(solv) values were calculated as $^1/_2\Delta V + \Delta G_{drc}$. [c] The basis for MC indicates the source of the wave function used in the CHELPG fit of the atomic charges to the molecular electrostatic potential.

models, these terms have not been considered in this table. Instead, the two focal points here are comparisons of the corresponding ΔG(solv) values calculated with different models as well as the total relative free energies, ΔG, obtained using different basis sets in derivation of the internal energy terms.

The IEF-PCM/B3LYP/6-31G* ΔG value is −0.13 kcal/mol for the 1*H* to 3*H* tautomerism of the 4-Me-imidazole in $CHCl_3$, in good agreement with the ΔG value of 0.01 ± 0.12 kcal/mol from the FEP/MC approach. (The ΔG value bears the same standard deviation as the ΔG(solv) term calculated in the FEP process.) The agreement is also good in aqueous solution: the corresponding values are −0.15 and 0.14 ± 0.15 kcal/mol. Neither result differs significantly at the 2 SD level.

Using the IEF-PCM/B3LYP/6-311++G** basis set in single point calculations for obtaining ΔE^s_{int}, the ΔG values are still very close with the chloroform solvent−0.21 and 0.16 ± 0.12 kcal/mol from the two solvent models. The deviation becomes, however, unacceptably large in aqueous solution: $\Delta G = 0.11$ vs 1.92 ± 0.14 kcal/mol. Furthermore, the sign of the ΔG value changes when the 6-311++G** set is used instead of the 6-31G* basis set in IEF-PCM/B3LYP calculations for the 1*H* and 3*H* tautomers of 4-Me-imidazole.

In contrast, the basis set effect is relatively small in the IEF-PCM calculations for the 1*H* to 2*H* tautomerism of 4-Me-1,2,3-triazole. Both basis sets predict the overwhelming amount of the 2*H* tautomer both in $CHCl_3$ and water solvent. FEP/MC results predict even larger preference for the 2*H* tautomer in chloroform, but according to this model the equilibrium is shifted in favor of the 1*H* tautomer in aqueous solution. Whereas the FEP/MC ΔG(solv) value was more negative by 1.6 kcal/mol than the IEF-PCM value in chloroform, the FEP/MC ΔG(solv) becomes more positive by 2 kcal/mol in aqueous solution.

The same qualitative conclusions may be drawn for the 1*H* to 3*H* tautomerization of 4-Me-1,2,3-triazole. A relatively small basis set effect was found in IEF-PCM, and both models predicted favorably the 3*H* tautomer in chloroform as a result of the agreeing ΔG(solv) value. In contrast, the continuum method prefers the 3*H* tautomer, and the FEP/MC model predicts the prevalence of the 1*H* form in aqueous solution due to a deviation in ΔG(solv) as much as of about 1.7 kcal/mol.

The 1*H* to 2*H* tautomerism of 5-Me-tetrazole leads to qualitative disagreement in $CHCl_3$, where the 1*H* and 2*H* forms are favored by the continuum and the explicit model, respectively. The difference in the ΔG(solv) values is 1.8 kcal/mol. In contrast, both models favor the 1*H* tautomer in aqueous solution, although the population of this form is more than 96% according to the IEF-PCM method, but a detectable equilibrium of 1*H*:2*H* = 63:37 has been predicted by the FEP/MC approach.

Overall, the results in Table 8 suggest no clear trends in predicting relative total free energies of tautomers by using different theoretical approaches. Signs of the ΔG values calculated by the two models agree in some cases and disagree in some other. No consistent agreement was calculated either in $CHCl_3$ or in aqueous solution. In the applied approaches, ΔG(solv) is responsible for these differences; thus, a thorough analysis of this term in both methods is necessary. This analysis must explore the effect of the cavity radii in the IEF-PCM method in conjunction with an analysis of the theoretical level (e.g., B3LYP, MP2 or something higher), as well as the effect of the applied basis set at least in single point calculations. On the FEP/MC side, the most important problem is finding the useful fitting method (e.g., CHELPG, RESP, or something else) of the net atomic charges to the molecular electrostatic potential (ELPO) of the solute in the given solvent. An important question is, which wave function to be used in generating the in-solution ELPO. At this point, the two investigations regarding the IEF-PCM method described above and the FEP/MC parameter studies meet. A series of IEF-PCM calculations are needed for equilibria wherein the experimental values are available. The best IEF-PCM method should be selected and different fitting processes of the atomic charges to the derived ELPO should be studied. A series of FEP/MC calculations with different atomic charges should be performed and compared with the experimental results. The best agreement for the relative free energy would indicate the favorable charge fitting method to the ELPO generated by the wave function, which was obtained

zat the previously selected "best" IEF-PCM level and applying the "best" basis set. The outlined systematic analysis is in progress.

Equilibration Mechanism in Solution. As was mentioned in relation to the gas-phase results, the 1,2-prototropic shift was studied theoretically for pyrrazole by Alkorta and Elguero.[10] Lunazzi et al.[9a] argued in favor of this mechanism not only for pyrrazole but also for 1,2,3-triazole and for the 1,2-rearrangement in 1,2,4-triazole. But what is the equilibration mechanism for the 1,3-relocation, as it appears in imidazole or in 1,2,4-triazole (called 1,4-tautomerization)?

In protic solvents, the solvent-mediated route could be simple. The real problem is the mechanism in low polarity solvents, such as $CHCl_3$. An intermolecular proton jump is always a theoretical possibility. If only one proton jumps over, however, then an ion-pair is formed which may or may not be stabilized by the solvent. Since $CHCl_3$ is unlikely to stabilize this reaction path, a possible alternative is two simultaneous proton jumps. To make it feasible, a dimer must be formed and the TS structure must be stabilized by the solvent. Our molecular dynamics study aimed to investigate this possibility.

In the starting solution configuration, two 4-Me-imidazole molecules were arranged in stacked antiparallel position, favorable for the simultaneous double proton-jump, in a box of $CHCl_3$. Two other solutes were set at a distance of 7−8 Å from the dimer in different perpendicular planes. After a short time a N−H···N bonded chain of the four molecules was formed and maintained throughout the simulation up to a total time of 1.5 ns. Our MD did not allow a proton jump, thus the present study only suggests that 4-Me-imidazole favorably forms hydrogen-bonded chain(s) in $CHCl_3$. The chain was not entirely linear, raising the possibility that a large cycle may be formed at a satisfactory concentration. In this case, simultaneous proton jumps can proceed as a circle current without the need for a remarkable support from the solvent. If this association is strong enough, the dilution does not affect the solution structure after the hydrogen-bonded rings have formed. Our results showed that about 10−12 4-Me-imidazole molecules might form a closed ring.

The above speculation may appear needless until our results for the intramolecular proton relocation have been analyzed. As mentioned in the section on Gas-Phase Tautomerism, no transition state for the one-step 1*H* to 3*H* intramolecular proton jump has been identified either for imidazole or for the imidazole···$CHCl_3$ complex. Then the 1,2-prototropic shift was studied when the proton from the N_1-site relocates to the C_2 atom, and in a next step it reaches the N_3 atom. Although this process is symmetrical in imidazole, these calculations may serve as a simplified model for the 1,3 tautomerism for other systems (1,2,4-triazole) and their substituted derivatives.

Table 2 compares the 1*H*, 2*H*, and TS imidazole geometries in $CHCl_3$ and water. Whereas the bond angles formed by the ring atoms are not very different in the three structures, the bond lengths undergo essential changes. Comparing, however, these geometric parameters with the corresponding gas-phase structures (Tables 1,2) it reveals that the geometric change in the $1H \rightarrow TS \rightarrow 2H$ relocation is an inherent feature of the imidazole system, and these geometric changes are hardly affected by the solvent.

The solvent effect is relatively small, not only on the molecular geometry but on the free energy change of the 1,2-prototropic shift in imidazole. Table 3 shows that the reaction barrier, and thus the free energy difference between the TS and the 1*H* form, is 45−48 kcal/mol in the gas phase whereas the values are 48−49 and 49−50 kcal/mol in $CHCl_3$ and water, respectively. The proton relocation in water obviously need not proceed along the intramolecular path with such a high barrier. For the equilibration, however, in $CHCl_3$, this route is an alternative as compared to the intermolecular proton transform.

$\Delta(E+ZPE)$ for the pyrrazole TS was calculated as 47.3 kcal/mol at the B3LYP/6-31G* level in the gas phase by Alkorta and Elguero.[10] The corresponding value from the present study for imidazole was very similar: 46.3 kcal/mol (Table 3). These close values suggest that the energy barrier is insensitive to whether the process is a N−H···N or a N−H···C proton relocation, and implies similar activation parameters for pyrrazole and imidazole in solution, as well.

Lunazzi et al.[9a] found an activation free energy of 13.6 kcal/mol for pyrrazole in DMSO. The theoretical gas-phase $\Delta(E+ZPE)$ value is 47.3 kcal/mol. This latter term does not contain the relative vibrational thermal corrections and the entropy changes, which are, however, negligible according to data presented in Table 3. Thus, the 34 kcal/mol difference may indicate three things: (a) there is a large solvent effect, (b) much higher level calculations than the present one are needed for theoretical characterization of the process, or (c) the experimental results refer to a tautomerization reaction path different from that studied by Alkorta and Elguero.

In our study we investigated both the gas-phase and the in-solution process. Unless the quantum-mechanical terms are in error of about 30 kcal/mol, we found that the TS activation free energy is at least about 45 kcal/mol for imidazole and increases rather than decreases upon solvent effects. Thus, the 1,2-prototropic shift requires a large activation free energy even in solution, and either only more sophisticated calculations could predict a value of about 13 kcal/mol activation free energy or the Lunazzi experiment reflects a net activation free energy of a complicated mechanism, different from a simple intramolecular 1,2-shift. In this respect, the present MD simulations may be a starting point for discovering an alternative and energetically favorable tautomerization path for imidazole derivatives.

Structure of the First Solvation Shell. Although the FEP/MC results may suggest that the atomic charges are exaggerated, MC simulations still can provide useful information regarding the solution structure. Atomic charges change generally by some hundredths of a unit for modeling the solute in low polarity solvent and in water. Relative free energies are more sensitive to these small changes than general parameters used for characterization of the solution structure.

The structure of the first solvation shell, crucial in chemical reactions or even in conformational changes, could be usefully characterized by radial distribution functions[35b] and coordination numbers. For the characterization of the interaction between the solute and solvent molecules, the pair-energy distribution function (pedf), $dN(E)/dE$, is a powerful theoretical tool. $dN(E)/dE$ gives the number of the solvent molecules in energy interaction of $E \pm {}^1/_2\Delta E$ with the solute. With polar solute and solvents, the dominating contribution to the interaction energy (excluding arrangements with atom−atom distances much shorter than the sum of the van der Waals radii) is the Coulomb term when the OPLS-AA 12−6−1 interaction potential is used. In favorable orientations and at a separation of the O··H and N··H atoms of about 1.7−2.0 Å, the interaction energy is generally the most negative. The solvent molecules with such an arrangement around the solute comprise the first solvation shell. Thus, the number of the solvent molecules with the most negative interaction energy nearly corresponds to the number of molecules in the first solvation shell. This number can be

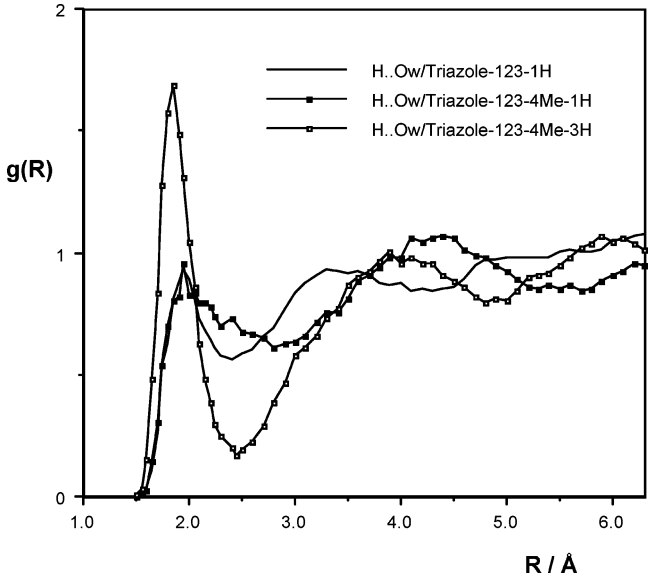


Figure 1. H···O(water) radial distribution functions for 1,2,3-triazole-1*H* and the tautomers of 4Me-triazole-1,2,3.

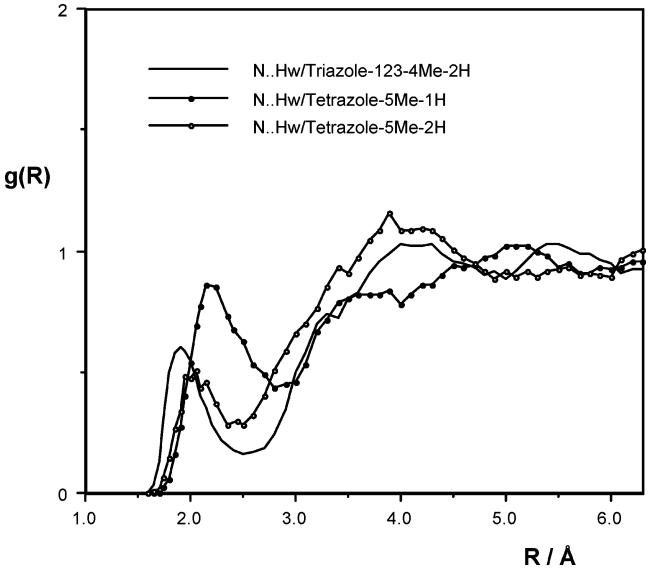


Figure 2. N···H(water) radial distribution functions for 4Me-triazole-1,2,3-1*H* and the tautomers of 5Me-tetrazole. The selected basic N atom is the fourth ring atom along the path starting at the methyl-substituted ring carbon and including the N−H site.

determined by the integration of the pedf until some distinguishing point.

Before analyzing the pedfs, study of the radial distribution functions (rdf) provides useful information about the composition of the first solvation shell. Figure 1 shows the (N)H···O_w (water oxygen) rdf's for 1,2,3-triazole tautomers. The 1*H* tautomers for both the unsubstituted ring and the 4-Me derivative provide rdf's with nearly equal peak values of $g(R) \approx 0.95$ at $R = 1.95$ Å, but the first minimum becomes less defined for the methyl derivative. This character of the $g(R)$ suggests a minor effect due to the methyl substitution of the imidazole ring in the case of the 1*H* structure, where the substituent represents a nonneighbor group with respect to the N_1−H site. In the case of the N_3−H tautomer, however, the methyl substituent corresponds to a neighboring group and the $g(R)$ curve changes dramatically. The peak appears at a shorter distance than before, $R = 1.85$ Å, and the peak value is almost doubled: $g(R) = 1.7$. The first minimum is well-defined at $R = 2.45$, and $g(R)$ has only a small value here. This character corresponds to strongly localized water molecules at the N_3−H site. Although integration of the $g(R)$ functions to their first minima provides about 1 water oxygen in the first hydration shell of the N−H group in all three cases, this water must be much more strongly localized at about N_3−H···$O_w = 1.85$ Å for the 3*H* tautomer of the 4-Me derivative than the hydrogen bonded water molecules for either the 4-Me-1,2,3-triazole-1*H* or the unsubstituted 1*H* triazole tautomers. The finding indicates a "structure making" effect of the near alkyl susbstituent.

Hydrogen bonds of the N···H_wO_w type (Figure 2) must be less strong than the N−H···O_w types above. The nitrogen atoms in Figure 2 were selected as the fourth ring atom along the path starting at the substitution site and including the N−H site. The peak values of $g(R)$'s are considerably less than 1, thus the probability of finding a water hydrogen close to an aromatic nitrogen is smaller than finding it in the bulk water. This means that the hydrogen bonds to nitrogens of five-membered heterocycles must be weaker than those in the bulk water. The closeness of the N−H site influences, however, remarkably the localization strength of the water hydrogen. For 5-Me tetrazoles, the N_3···H_w $g(R)$ peak is higher and appears at a larger R for the 1*H* than for the 2*H* tautomer. The neighboring N−H seems to reduce the chance for the formation of the N···H_w bond. The peak is similarly low for the triazole derivative in this figure, where the 2*H* tautomer also represents a neighboring N−H site with respect to the N_1···H_wO_w hydrogen bond.

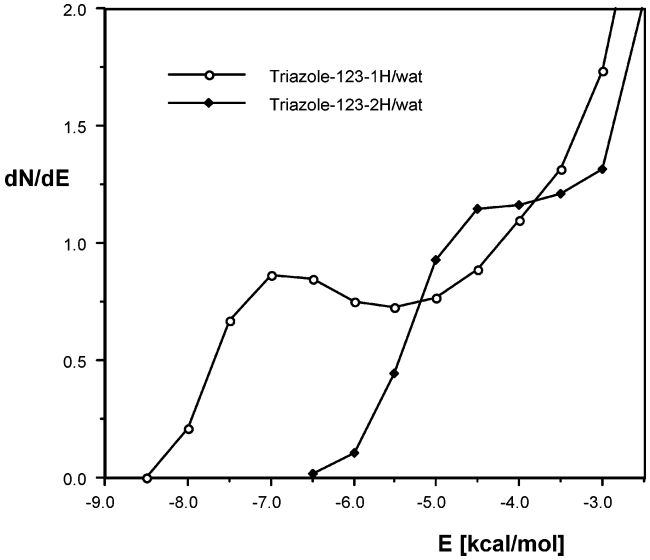


Figure 3. Pair-energy distribution functions for 1,2,3-triazole tautomers in water.

The pedfs for 1,2,3-triazole in water (Figure 3) indicate definite differences in the pair energies for the tautomers. The most negative interaction energy is −8.5 to − 8 kcal/mol for the 1*H* form. The threshold value for the 2*H* tautomer is −6.5 kcal/mol. The pedf has a maximum−minimum shape for 1*H* but exhibits only a shoulder for 2*H*. Integration of pedf/1*H* until $E = -5.5$ kcal/mol provides two water molecules in strong interaction with the solutes. From the analysis of the rdf's above, one water molecule is assigned for each hydrogen bond of the N_1−H···O_w and the N···H_wO_w type. For the latter, N may be both N_2 and N_3; the results do not provide the ratio for them. On the basis of the finding for the role of the neighboring N−H, the N_3 atom likely binds the water molecules more strongly. Also two strongly bound water molecules are calculated for the 2*H* tautomer, by integration of the corresponding pedf until the end of plateau at $E = -3.5$ kcal/mol. These water molecules

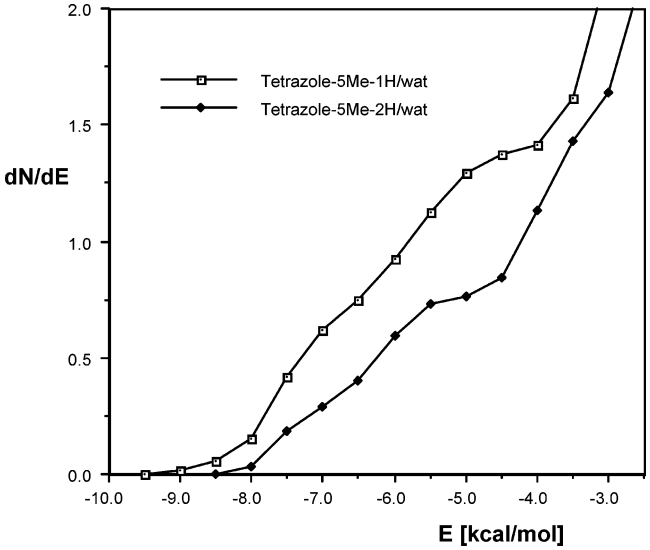


Figure 4. Pair-energy distribution functions for the 5Me-tetrazole tautomers in water.

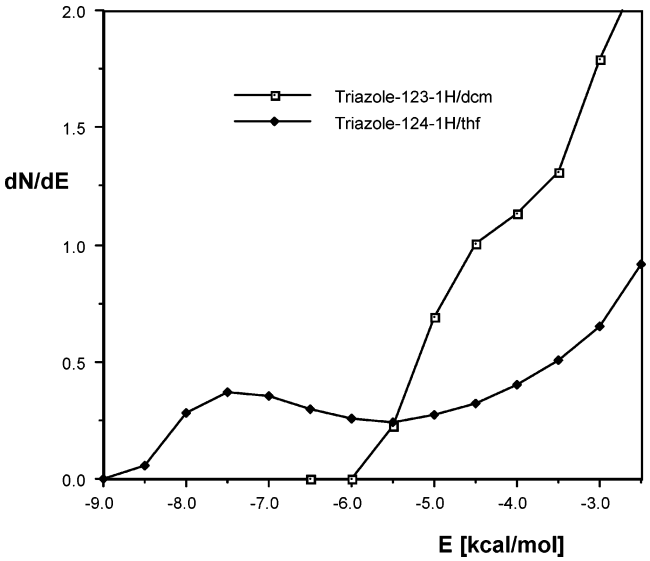


Figure 5. Pair-energy distribution functions for 1,2,3-triazole-1*H* in tetrahydrofuran and for 1,2,4-triazole-1*H* in dichloromethane.

are clearly in weaker interaction with the solute than in the case of the 1*H* tautomer. They are expected to form both N−H···O_w and N···H_wO_w bonds, where the basic nitrogens are equivalent for the 2*H* tautomer.

The pedfs in Figure 4 for the tautomers of 5-Me-tetrazole in water do not show a maximum-minimum character. The shoulders indicate the coexistence of different N··H··O hydrogen bonds with different interaction energies. The rightward shift of the pedf for the 2*H* tautomer implies weaker interactions between this tautomer and the water molecules. The slopes of the curves become monotonically increasing with $E > -4$ and -3 kcal/mol for the 1*H* and 2*H* tautomers, respectively. This shape of the pedfs indicates no further special range of the interaction energy between the solute and a solvent molecule. Integration of the distribution function up to these limits provides four water molecules strongly bound to either solute, allowing a classical interpretation of one N−H···O_w and three N···H_wO_w for hydrogen bonds in each case.

Figure 5 compares the pedfs for triazoles in slightly polar solvents for which experimental results were available.[9a] The two curves are qualitatively different as a consequence of the solvent's chemical structure. There is a resolved, although stretched maximum-minimum shape with the THF solvent. This solvent can be considered as a fairly strong hydrogen-bond acceptor forming N−H···O(ether) intermolecular bonds.

The pedf for 1,2,3-triazole-1*H* shows only a shoulder instead of a resolved structure. This pattern is attributed to the character of the CH_2Cl_2 solvent rather than to that of the solute. The CH_2Cl_2 solvent must be very weak hydrogen bond acceptor if at all. Former studies with $CHCl_3$ and CH_2Cl_2 solvents did not show any peak for the H···Cl rdfs with polar solutes.[36] The CH_2Cl_2 pedf in Figure 5 shows a so-called bipolar distribution, characteristic for polar, nonprotic solvents.[26d,37]

Integration of the pedf in THF until $E = -5.5$ kcal/mol provides 0.9 water molecules. This means that most of the time there is about one THF molecule bound to triazole, which interaction may be considered to be of the N−H···O(ether) type. Integration of the CH_2Cl_2 pedf until the end of the plateau at $E = -3.5$ kcal/mol provides three "strongly bound" CH_2Cl_2 molecules. These molecules may be considered, however, to be only in a favorable dipole−dipole interaction with the triazole solute.

Conclusions

The obtained results may be usefully applied for designing N-heterocyclic ligands that could form hydrogen bonds with protein side chains. Specific conclusions from this study have been derived as follows:

1. Gas-phase molecular geometries for small neutral species can be determined in good agreement with the experimental values by optimization at the B3LYP/6-31G* level.[12−15] Consideration of the electron correlation is important throughout the geometry optimization for aromatic heterocyles, because HF optimizations lead to too short bond lengths for the ring. For imidazole, the rms deviation with reference to the experimental bond lengths and bond angles in the ring is smaller at the B3LYP/6-31G* level than either at the MP2/aug-cc-pvtz or the B3LYP/aug-cc-pvtz level. The better performance of the B3LYP/6-31G* optimization compared to higher level calculations may be attributed to an error cancellation due to the smaller basis set and the applied approach for considering the electron correlation.

2. For a consistent consideration of the change of the internal energy, solute geometries are to be optimized in solution. Ring geometries of five-member N-heterocycles were found to change only slightly in solutions of different solvents. The rings maintain a planar heavy atom skeleton and the carbon atom of the methyl substituent lies in this plane, as well.

3. Gas-phase relative free energies and related equilibrium constants for the tautomers of N-heterocycles can dramatically change in solution, in line with previous results obtained mainly for aqueous solutions.[2,6−8] From the present study, the internal energy increases and the solute−solvent interaction energies becomes more negative when the dielectric constant increases from 4.7 to 78.4. The resulting free energy in solution is a balance of the two effects. The free energy changes differently for tautomers in solution as compared to the gas phase, resulting in a shift of the equilibrium constant upon solvation.

4. The methyl-substituted imidazole and pyrrazole form delicate equilibria between two tautomeric forms for each. IEF-PCM/B3LYP/6-31G* results predict detectable populations for both tautomers of each ring system. For 4-Me-1,2,3-triazole, the 2*H* form is the most stable one even in aqueous solution, despite the small dipole moment of this tautomer. The 4*H* tautomer of 3-Me-1,2,4-triazole is unstable in any solvent within

the $\epsilon = 4.7-78.4$ range. For 5-Me-tetrazole, the population of the 1*H* form increases with increasing dielectric constant, according to the IEF-PCM analysis.

5. Considering explicit solvent models and using the FEP/MC method for estimating the relative solvation free energies, results are in contrast to those from the continuum solvent model in several cases. This finding indicates the need for an analysis of the effects of the cavity radii, the basis set, and the selected theoretical level in the IEF-PCM calculations, and application of an appropriate method for fitting the atomic charges to the in-solution molecular electrostatic potential at the in-solution optimized solute geometry. The final combination of the theoretical approaches should be based on the best agreement of the theoretically predicted equilibrium constants with available experimental ones.

6. Analyses of radial distribution and pair-energy distribution functions suggest that there is always one water molecule that forms a strong $N-H\cdots O_w$ hydrogen bond, whereas all other nitrogens in the ring form $N\cdots H_wO_w$ type hydrogen bonds. This latter type seems to be weaker than the $N-H\cdots O_w$ bond. Tetrahydrofuran acts as a hydrogen bond acceptor and forms $N-H\cdots O$(ether) bonds, slightly weaker than the water-containing counterpart. Nonprotic solvents, not even with an acceptor site as in tetrahydrofuran, are in favorable dipole−dipole interactions with the solute.

7. The 1,2-prototropic shift, as a reaction path for the intramolecular tautomerization, requires about 50 kcal/mol activation energy for imidazole and for pyrrazole[10] in the gas phase. A solvent environment even disfavors this tautomerization path for imidazole, where no transition state has been found for a direct 1,3 relocation. The calculated activation free energy is 48−50 kcal/mol in chloroform and water, as compared to 13.6 kcal/mol, measured by Lunazzi et al.[9a] for pyrrazole in DMSO. A molecular dynamics computer experiment favors the formation of an imidazole chain in chloroform, making feasible the 1,3-tautomerization along an intermolecular path in nonprotic solvents.

Acknowledgment. The authors are indebted to the Ohio Supercomputer Center for computer time granted for the quantum-mechanical calculations and to Professor Jorgensen for permission to use the BOSS 3.6 software. This work was supported by an NIH grant (NS 31173).

References and Notes

(1) Eicher, T.; Hauptmann, S. *The Chemistry of Heterocycles*; Wiley-VCH GmbH & Co. KGaA: Weinheim, Germany, 2003.

(2) (a) Woodcock, S.; Green, D. V. S.; Vincent, M. A.; Hillier, I. H.;. Guest, M. F.; Sherwood, P. *J. Chem. Soc., Perkin Trans. 2* **1992**, 2151. (b) Gould, I. R.; Hillier, I. H. *J. Chem. Soc., Perkin Trans. 2* **1993**, 1771. (c) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 8810.

(3) For reviews, see, e.g.: (a) Grimmett, M. R. In *Comprehensive Heterocyclic Chemistry II*; Katritzky, A. R., Rees, C. W., Scriven, E. F. V., Eds.; Pergamon Press: Oxford, U.K., 1996; Vol. 3, p 77. (b) Rauhut, G. In *Advances in Heterocyclic Chemistry*; Katritzky, A. R., Ed.; Academic Press: San Diego, CA, 2001; Vol. 81, p 1.

(4) (a) Cooper, D. G.; Young, R. C.; Durant, G. J.; Ganellin, C. R. In *Comprehensive Medicinal Chemistry*; Pergamon Press: Oxford, U.K., 1990; Vol. 3, p 323. (b) Arrang, J.-M.; Garbarg, M.; Lancelot, J.-C.; Lecomte, J.-M.; Pollard, H.; Robbs, M.; Schunack, W.; Schwartz, J.-C. *Nature (London)* **1987**, *327*, 117. (c) Lipp, R.; Arrang, J.-M.; Garbarg, M.; Luger, P.; Schwartz, J.-C.; Schunack, W. *J. Med. Chem.* **1992**, *35*, 4434.

(5) Goldblum, A. In *Computational Approaches to Biochemical Reactivity*; Náray-Szabó, G., Warshel, A., Eds.; Kluwer Academic Publishers: New York, 2002; p 295.

(6) (a) Parchment, O. G.; Green, D. V. S.; Taylor, P. J.; Hillier, I. H. *J. Am. Chem. Soc.* **1993**, *115*, 2352 (b) Cao, M.; Teppen, B. J.; Miller, D. M.; Pranata, J.; Schäfer, L. *J. Phys. Chem.* **1994**, *98*, 11353.

(7) (a) Cox, J. R.; Woodcock, S.; Hillier, I. H.; Vincent, M. A. *J. Phys. Chem.* **1990**, *94*, 5499. (b) Murdock, S. E.; Lynden-Bell, R. M.; Kohanoff, J.; Margulis, C. J.; Sexton, G. J. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5281.

(8) Wong, M. W.; Leung-Toung, R.; Wentrup, C. *J. Am. Chem. Soc.* **1993**, *115*, 2465.

(9) (a) Lunazzi, L.; Parisi, F.; Macciantelli, D. *J. Chem. Soc., Perkin Trans. 2* **1984**, 1025. (b) Albert, A.; Taylor, P. *J. Chem. Soc., Perkin Trans. 2* **1989**, 1903.

(10) Alkorta, I.; Elguero, J. *J. Chem. Soc., Perkin Trans. 2* **1998**, 2497.

(11) (a) Ojo, B.; Dunbar, P. G.; Durant, G. J.; Nagy, P. I.; Huzl, J. J., III; Periyasamy, S.; Ngur, D. O.; El-Assadi, A. A.; Hoss, W. P.; Messer, W. S., Jr. *Bioorg. Med. Chem.* **1996**, *14*, 1605. (b) Messer, W. S., Jr.; Abuh, Y. F.; Liu, Y.; Periyasamy, S.; Ngur, D. O.; Edgar, M. A. N.; El-Assadi, A. A.; Sbeih, S.; Dunbar, P. G.; Roknich, S.; Rho, T.; Fang, Z.; Ojo, B.; Zhang, H.; Huzl, J. J., III; Nagy, P. I. *J. Med. Chem.* **1997**, *40*, 1230. (c) Huang, X. P.; Nagy, P. I.; Williams, F. E.; Peseckis, S. M.; Messer, W. S., Jr. *Br. J. Pharmacol.* **1999**, *126*, 735. (d) Rajeswaran, W. G.; Cao, Y.; Huang, X. P.; Wroblewski, M. E.; Colclough, T.; Lee, S.; Liu, F.; Nagy, P. I.; Ellis, J.; Levine, B. A.; Nocka, K. H.; Messer, W. S., Jr. *J. Med. Chem.* **2001**, *44*, 4563. (e) Cao, Y.; Zhang, M.; Wu, C.; Lee, S.; Wroblewski, M. E.; Whipple, T.; Nagy, P. I.; Takács-Novák, K.; Balázs, A.; Tõrös, S.; Messer, W. S., Jr. *J. Med. Chem. Soc.* **2003**, *46*, 4273 2003.

(12) Nagy, P. I.; Takács-Novák, K. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2838.

(13) Nagy, P. I.; Kökösi, J.; Gergely, A.; Rácz, Á. *J. Phys. Chem. A.* **2003**, *107*, 7861.

(14) Nagy, P. I.; Tejada, F. R.; Sarver, J. G.; Messer, W. S., Jr. *J. Phys. Chem. A.* **2004**, *108*, 10173.

(15) (a) Nagy, P. I. *J. Phys. Chem. B* **2004**, *108*, 11105. (b) Nagy, P. I.; Erhardt, P. W. *J. Phys. Chem. B* **2005**, *109*, 5855.

(16) Nagy, P. I.; Völgyi, G.; Takács-Novák, K. *Mol. Phys.* **2005**, *103*, 1589.

(17) Melandri, S.; Maris, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2863.

(18) (a) Christen, D.; Griffiths, J. H.; Sheridan, J. Z. *Z. Naturforsch., Teil A* **1981**, *36*, 1378. (b) Nagy, P. I.; Durant, G. J.; Smith, D. A. *J. Am. Chem. Soc.* **1993**, *115*, 2912.

(19) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(20) (a) Hehre, W. J.; Radom, L.; Schleyer, P.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986. (b) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618. (c) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem.* **1976**, *10s*, 1.

(21) (a) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796. (b) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1994**, *100*, 2975. (c) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1995**, *103*, 4572.

(22) (a) Miertus, S.; Scrocco, E.; Tomasi, *J. Chem. Phys.* **1981**, *55*, 117. (b) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027. (c) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210. (d) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161. (e) Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, *100*, 4187. (f) Curutchet, C.; Cramer, C. J.; Truhlar, D. G.; Ruiz-Lopez, M. F.; Rinaldi, D.; Orozco, M.; Luque, F. J. *J. Comput. Chem.* **2003**, *24*, 284.

(23) (a) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032. (b) Cancès, E.; Mennucci, B. *J. Chem. Phys.* **1998**, *109*, 249. (c) Cancès, E.; Mennucci, B. *J. Chem. Phys.* **1998**, *109*, 260. (d) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

(24) *CRC Handbook of Chemistry and Physics*, 85th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, FL, 2004.

(25) Gaussian 03, Revision B.05: Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Pittsburgh, PA, 2003.

(26) (a) Jorgensen, W. L.; Madura, J. D. *J. Am. Chem. Soc.* **1983**, *105*, 1407; (b) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 1489. (c) Jorgensen, W. L.; Gao, J. *J. Phys. Chem.* **1986**, *90*, 2174. (d) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. *J. Phys. Chem.* **1990**, *94*, 1683.

(27) (a) Zwanzig, R. W. *J. Chem. Phys.* **1952**, *22*, 1420. (b) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.

(28) Jorgensen, W. L. BOSS, Version 3.6. *Biochemical and Organic Simulation System User's Manual*; Department of Chemistry, Yale University: New Haven, CT, 1995.
(29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(30) (a) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225. (b) Rizzo, R. C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1999**, *121*, 4827.
(31) Alagona, G.; Ghio, C.; Nagy, P. I. *Int. J. Quantum Chem.* **2004**, *99*, 161.
(32) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.
(33) *Sybyl 6.91*; Tripos Inc.: St. Louis, MO.
(34) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936.
(35) (a) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA, 2000. (b) Cramer, C. J. *Essentials of Computational Chemistry*; Wiley: New York, 2002; pp 78−80.
(36) (a) Nagy, P. I. *Acta Chim. Hung.* **1992**, *129*, 429. (b) Nagy, P. I.; Takács-Novák, K. K *J. Am. Chem. Soc.* **2000**, *122*, 6583. (c) Nagy, P. I.; Takács-Novák, K.; Ramek, M. K. *J. Chem. Phys.* B **2001**, *105*, 5772.
(37) Jorgensen, W. L.; Briggs, J. M. *Mol. Phys.* **1988**, *63*, 547.