

Studies on tautomerism: Benchmark quantum chemical calculations on formamide and formamidine

Geza Fogarasi *

Institute of Chemistry, Eotvos University, H-1518 Budapest, Pf. 32, Hungary

ARTICLE INFO

Article history:

Received 21 December 2009

Received in revised form 23 February 2010

Accepted 23 February 2010

Available online 1 March 2010

Keywords:

Tautomerism

Proton transfer

Transition state barrier

Ab initio quantum chemistry

Coupled cluster

ABSTRACT

Tautomerism is a ubiquitous phenomenon in both chemistry and molecular biology and numerous computational studies of variable accuracy are available on it. In the present work, we report on some benchmark quantum chemical calculations on the formamide $\leftrightarrow$ formamidic acid (**1**) and formamidine $\leftrightarrow$ formamidine (**2**) systems. Some results on the acetaldehyde $\leftrightarrow$ vinyl alcohol and acetaldimine $\leftrightarrow$ vinyl amine pairs are also presented. A central question of the study is the role of water as a catalyst, investigated in form of the monohydrates of **1** and **2**. Optimized structures have been obtained for the tautomers and the transition states, including their monohydrates. The calculations represent the most comprehensive and highest level study up to now, with the goal of converging the relevant energy differences to an accuracy of 0.5 kcal/mol. To this aim, basis sets cover a range from simple 6-31G(d,p) up to 6-311++G(3df,3pd) and from cc-pVTZ up to aug-cc-pV5Z. The electron correlation treatment has been varied from RHF and DFT up to CCSD(T), plus one case with CCSDT (true triples). The convergence of the results with respect to the computational level is, unfortunately, different. The tautomerization energies of the *free* molecules converge at CCSD/aug-cc-pVTZ or /cc-pVQZ, the triples being not necessarily required. However, in the *general* case (hydrates and transition states) one needs the triples, and the basis set should be of quadruple-zeta quality, to achieve really strict convergence.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Tautomerism is a special form of isomerism in which a proton is relocated within the molecule, accompanied by appropriate changes in the bonding structure. The process plays an important role in numerous chemical and biological systems. Specifically, one of the possible mechanisms of mutation of DNA may be the tautomerization of nucleotide bases [1].

Tautomerization in various systems has been the subject of numerous quantum chemical (QC) studies. Still, it is difficult to get a clear picture about the *accuracy* of the calculations. This may be a crucial point if the relative energies of tautomers are in a narrow range of a few kcal/mol: for example, the three main tautomers of cytosine differ in energy by 2–3 kcal/mol only [2]. In fact, the present study was largely motivated by the uncertainties concerning cytosine. In cases like that a really good description of relative energies should be accurate within ~ 0.5 kcal/mol. With this in mind, we want to perform benchmark calculations to check the ultimate accuracy achievable with present day QC methods. Four simple systems, shown in Fig. 1, have been selected for the tests: formamide (FMD) $\leftrightarrow$ formamidic acid (FAC), **1**; formamidine

(FIM) $\leftrightarrow$ formamidine, **2** (in the latter, the two tautomers are identical, the question is the transition state); acetaldehyde $\leftrightarrow$ vinylalcohol, **3**; and acetaldimine $\leftrightarrow$ vinylamine, **4**. These pairs represent four fundamental cases of tautomerism: amide $\leftrightarrow$ imidic acid, amine $\leftrightarrow$ imine, keto $\leftrightarrow$ enol and imine $\leftrightarrow$ enamine. Accordingly, a large number of previous studies are available for comparison, mainly for formamide [3–23] but also for the other, less ubiquitous systems [24–30]. For the pairs in **3** and **4**, tautomerization cannot be expected, the isomers exist as independent molecules. Still, the isomerization energy in these systems can add information about the consistency of the computational methods used. An important part of the present study will be the investigation of the role of water. This will be done on the two smaller systems, in form of monohydrates: first, the change in tautomerization energies will be determined, and then the effect of water on the transition barrier will be studied.

2. Computational details

The level of QC calculations was varied systematically, from Hartree–Fock (RHF) through Moller–Plesset 2nd order perturbation (MP2) theory up to Coupled Cluster (CC) theory in the CCSD(T) approximation. (In one single case, even CCSDT – with true triples – was tested.) Beside traditional wave function methods, density

* Tel.: +36 1 3722540; fax: +36 1 3722909.

E-mail address: fg@chem.elte.hu

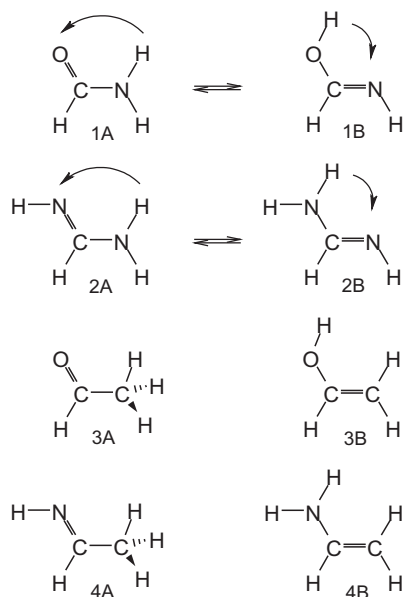


Fig. 1. The four tautomer pairs investigated: **1A** – formamide, **1B** – formimidic acid, **2A** – formamidine, **2B** – formamidine (the two tautomer forms are identical), **3A** – acetaldehyde, **3B** – vinylalcohol, **4A** – acetaldimine, **4B** – vinylamine.

functional theory (DFT) was also examined for comparison, applying the B3LYP functional. The quality of basis sets ranged from 6-31G(d,p) to aug-cc-pV5Z. (For the above standard acronyms, see, e.g. [31]). Pople-type sets were used in the lower level, RHF and DFT calculations, while Dunning's correlation consistent (cc) basis sets were chosen for the CC calculations. In between, in the MP2 calculations both types of basis sets were tested. Up to medium levels of theory, optimum geometries and the corresponding energies were calculated at the same level. For the largest calculations, the MP2/aug-cc-pVTZ optimized structures were adopted as our standard geometries and only energies were computed at higher levels. MP2 was used with frozen core. Vibrational frequencies were calculated for the transition states to check the nature of a stationary point. However, frequencies were in general not computed for the energy minima, so ZPE corrections will not be included in the energies (except for one case to check a literature result). This is because our main interest is just in the *internal* accuracy of the methods, and experimental energies are, of course, not available for comparison, anyway. Two computer program packages were used: PQS [32] for all RHF, B3LYP and MP2, and CFOUR [33] for the CC calculations.

3. Relative energies of free tautomer pairs

Results for three tautomer pairs are listed in Table 1. In very broad terms, the general picture is the same at all levels of theory: the keto form and its imine analogue is the more stable tautomer within a pair in all calculations and quantitatively, the energy differences do not change more than a few kcal/mol with the method. It is notable that even the simplest RHF method performs well in this respect.

However, as pointed out above, we are interested in an accuracy better than 0.5 kcal/mol. Therefore, in the discussion below, convergence will be considered safe if the last changes do not exceed 0.2 kcal/mol. We discuss first the role of basis sets. In the LCAO-MO-type RHF and B3LYP calculations, as seen in Table 1, convergence within the group of Pople-type basis sets has likely been achieved for system **1**, while systems **3** and **4** seem a little more sensitive to the quality of basis set. About the larger, correlation

consistent (cc) basis sets we have information in the group of MP2 results as, of course, the SCF wave function is a necessary first step in these calculations. (The geometry is fixed in these calculations – see the footnotes to Table 1 – but the changes in geometry are already negligible energetically at this point.) Within the MP2 calculations the RHF energies (not listed in Table 1) show already perfect convergence: for system **1**, the maximum change in ΔE is only 0.02(!) kcal/mol from the cc-pVTZ basis set up to aug-cc-pV5Z; for the other two systems the corresponding value is larger but still below 0.2 kcal/mol. Note that the MP2 energy itself is more sensitive to the basis set in **3** and **4** than in **1**, see next below. So, from the one-electron point of view, the aug-cc-pVTZ basis set can be considered as complete.

The size of a basis set plays, of course, another and more significant role in correlation calculations, as the volume of the configuration space increases rapidly with increasing basis sets. This can be seen explicitly in the MP2 results on **3** and **4**: in Table 1, when going from aug-cc-pVTZ to (aug)-cc-pV5Z the MP2 energies change by 0.5 kcal/mol in both systems; at the same time, the corresponding SCF energies change only 0.16 and 0.17 kcal/mol for **3** and **4**, respectively (see also above). Thus, in the total change of 0.5 kcal/mol about one third comes from the incompleteness of the one-electron basis while the role of the configuration space expansion is twice larger.

Electron correlation effects in general can be judged by comparing the perturbation theory results with the higher level coupled cluster results. For **1**, the best MP2 result is $\Delta E = 11.5$ – 11.6 kcal/mol, which should be compared with the best CCSD and CCSD(T) values of 10.6–10.7 kcal/mol (the latter seem to have converged, being stable within 0.2 kcal/mol). Thus, electron correlation beyond MP2 has an effect of about 1 kcal/mol. In the treatment of correlation one goes normally up to the CCSD(T) level, where the triple substitutions (excitations) are only approximated (by fourth-order perturbation theory). System **1** is the smallest one among those investigated here and we were able to check even the effect of *true* triples. As seen in Table 1, the CCSDT/cc-pVTZ result of 10.6 kcal/mol confirms the CCSD(T) approximation.

As already seen above, the acetaldehyde $\leftrightarrow$ vinylalcohol (**3**) and acetaldimine $\leftrightarrow$ vinylamine (**4**) systems are more sensitive to the computational method. For **3**, the best CCSD(T) result is $\Delta E = 9.3$ kcal/mol and seems very well converged with basis sets. For **4**, the aug-cc-pV5Z calculations would have been too expensive because of the lack of any symmetry (non-planar amino group). Nevertheless, the last two points in Table 1, obtained with basis sets aug-cc-pVQZ and cc-pV5Z suggest good convergence: they are listed with one decimal as 3.0–3.1 kcal/mol, but are in fact even closer, the more precise difference being 0.06 kcal/mol.

It is important to check the importance of triple substitutions. Comparing the CCSD and CCSD(T) results, it is reassuring to see that for the purpose of the present study, interested in *relative* energies within a tautomer pair, they are very close to each other: for **1** there is a difference of 0.1 kcal/mol only, for **3**, the difference is 0.2 kcal/mol, both very systematic. For **4** the two sets of data agree within 0.1 kcal/mol. Thus, one may easily jump to the conclusion that in analogous future calculations a significant part of computer time could be saved by neglecting triple excitations and using just the CCSD method. See, however, the hydrates below.

Finally, we discuss the DFT results. According to Table 1, DFT (with the popular and well established B3LYP functional) performs well, but seems less consistent than MP2. For system **1**, B3LYP is – somewhat surprisingly – practically the same as RHF, and with $\Delta E = 12.2$ – 12.3 kcal/mol overestimates the best CC-result by one and a half kcal/mol. At the same time, for system **3** DFT brings a significant improvement over RHF, with the result of $\Delta E = \sim 9.6$ kcal/mol (may be not quite converged with basis set) being close to the CC-result of 9.3 kcal/mol. Then, however, for system **4** DFT

Table 1Computed energies for tautomer pairs (energies, E in a.u. = 4.3594×10^{-18} J, differences, ΔE in kcal = 4.184 kJ).

Method ^a	Formamide 1A $E + 168$	Formamidic acid 1B $E + 168$	ΔE 1	Acetaldehyde 3A $E + 152$	Vinylalcohol 3B $E + 152$	ΔE 3	Acetaldimine 4A $E + 132$	Vinylamine 4B $E + 132$	ΔE 4
RHF									
/6-31G(d,p)//~	-0.94049	-0.92025	12.7	-0.92259	-0.90099	13.5	-1.08422	-1.07515	5.7
/6-311G(d,p)//~	-0.98228	-0.96233	12.5	-0.95792	-0.93789	12.6	-1.11134	-1.10399	4.6
/6-311++G(2d,2p)//~	-0.99477	-0.97571	12.0	-0.96887	-0.94940	12.2	-1.12218	-1.11510	4.4
/6-311++G(3df,3pd)//~	-1.00357	-0.98406	12.2	-0.97540	-0.95637	11.9	-1.12670	-1.12016	4.1
B3LYP									
/6-31G(d,p)//~	-1.89702	-1.87689	12.6	-1.83573	-1.81605	12.3	-1.96155	-1.95430	4.5
/6-311G(d,p)//~	-1.94626	-1.92570	12.9	-1.87686	-1.85968	10.8	-1.99430	-1.98980	2.8
/6-311++G(2d,2p)//~	-1.95970	-1.94026	12.2	-1.88762	-1.87187	9.9	-2.00467	-2.00106	2.3
/6-311++G(3df,3pd)//~	-1.96707	-1.94741	12.3	-1.89333	-1.87811	9.6	-2.00911	-2.00605	1.9
MP2									
/6-31G(d,p)//~	-1.42114	-1.40148	12.3	-1.37847	-1.35559	14.4	-1.53274	-1.52199	6.7
/6-311G(d,p)//~	-1.49439	-1.47634	11.3	-1.44258	-1.42260	12.5	-1.58039	-1.57192	5.3
/6-311++G(2d,2p)//~	-1.54620	-1.52793	11.5	-1.48654	-1.46941	10.7	-1.61982	-1.61415	3.6
/6-311++G(3df,3pd)//~	-1.61111	-1.59278	11.5	-1.54540	-1.52941	10.0	-1.67415	-1.66947	2.9
/PVTZ//~	-1.60545	-1.58737	11.3	-1.54215	-1.52570	10.3	-1.67071	-1.66515	3.5
/aug-PVTZ//~	-1.62077	-1.60257	11.4	-1.55433	-1.53824	10.1	-1.68169	-1.67682	3.1
/PVQZ//aug-PVTZ	-1.66265	-1.64435	11.5	-1.59281	-1.57707	9.9	-1.71534	-1.71068	2.9
/aug-PVQZ//aug-PVTZ	-1.66931	-1.65096	11.5	-1.59803	-1.58264	9.7	-1.72005	-1.71586	2.6
/PV5Z//aug-PVTZ	-1.68362	-1.66525	11.5	-1.61115	-1.59578	9.6	-1.73122	-1.72704	2.6
/aug-PV5Z//aug-PVTZ	-1.68665	-1.66823	11.6	-1.61365	-1.59835	9.6	-1.73532	-1.73126	2.6
CCSD//MP2^b									
/aug-PVTZ	-1.68002	-1.66289	10.7	-1.62548	-1.60943	10.1	-1.76155	-1.75569	3.7
/PVQZ	-1.75693	-1.73965	10.8	-1.69557	-1.68013	9.7	-1.82416	-1.81887	3.3
/aug-PVQZ	-1.76391	-1.74647	10.9	-1.70146	-1.68639	9.5	-1.82954	-1.82472	3.0
/pV5Z	-1.79244	-1.77521	10.8	-1.72857	-1.71344	9.5	-1.85439	-1.84949	3.1
/aug-pV5Z	-1.79590	-1.77865	10.8	-1.73192	-1.71672	9.5	-	-	-
CCSD(T)//MP2^b									
/aug-PVTZ	-1.70809	-1.69111	10.6	-1.65138	-1.63562	9.9	-1.78745	-1.78158	3.7
/PVQZ	-1.78671	-1.76965	10.7	-1.72315	-1.70802	9.5	-1.85169	-1.84636	3.3
/aug-PVQZ	-1.79428	-1.77703	10.8	-1.72949	-1.71476	9.3	-1.85747	-1.85265	3.0
/pV5Z	-1.82351	-1.80648	10.7	-1.75727	-1.74246	9.3	-1.88293	-1.87802	3.1
/aug-pV5Z	-1.82720	-1.81014	10.7	-1.76080	-1.74592	9.3	-	-	-
CCSDT//MP2^b									
/PVTZ	-1.68593	-1.66900	10.6	-	-	-	-	-	-

^a Notations follow standard convention (see, e.g. [31]), except that in the correlation consistent basis sets “cc” is tacitly assumed and thus omitted for brevity. After the double slash “//” the level of geometry optimization is indicated; “~” means that optimization was done at the same level as the energy calculation.

^b All coupled cluster energy calculations were done at the MP2/aug-pVTZ geometry.

underestimates the tautomerization energy, giving $\Delta E = 1.9$ as compared to the CC value of 3.1 kcal/mol.

A detailed analysis of previous reported results is not the purpose of this study but a comparison with highest level literature data may be instructive. For **1**, Markova and Enchev [8] reported on fourth-order perturbation theory (MP4) computations in a recent study. They obtained an isomerization energy of 12.13 kcal/mol at MP4/6-31G(d,p), including a Δ_{ZPE} of 0.49 kcal/mol. At MP4/6-311++G(3df,2p)//MP2/6-31+G(d) + ZPE they obtained 11.73 kcal/mol, with $\Delta_{\text{ZPE}} = 0.33$ kcal/mol. (Note that in this case we have also calculated the ZPE; at MP2/aug-cc-pVTZ we get 28.522 for **1A** and 28.981 kcal/mol for **1B**, giving a Δ_{ZPE} of 0.46 kcal/mol.) The above data from [8], when transformed to our ΔE value, i.e. without ZPE, become 11.6 and 11.4 kcal/mol, respectively. Similar results were reported earlier from Wiberg's group [21]: MP4SDQ 12.0, QCISD 11.7 kcal/mol, both with the 6-31+G(d,p) basis. These results are roughly in the range of our second-order, MP2 results of 12.3 and 11.5 kcal/mol (Table 1). Of course, the 6-31G(d,p) or 6-31+G(d,p) basis sets are relatively small for present day standards and not really adequate for higher-order correlation calculations. With the largest Gaussian type basis set, the MP4 result is already realistic but still too high as compared to CCSD or CCSD(T); in fact, this MP4 value of 11.4 agrees with our MP2 values of 11.5 kcal/mol, so higher-order perturbation did not improve the results.

For **3** and **4**, Andrés et al. [29] tested several methods, from semiempirical up to MP2, the latter with basis sets 6-31G(d,p), 6-

311++G(d,p), and 6-311++G(3df,2p). With 6-31G(d,p) our result for pair **3** agrees perfectly with theirs, as it should be. For pair **4**, however, there is a discrepancy: our value, even after checking, is 6.7, while they reported a value of 8.57 kcal/mol. For the larger basis sets their choice is slightly different from those in our Table 1, but we agree in the range of values obtained, 10–12 kcal/mol for **3** and 4–6 kcal/mol for **4**.

For **3**, most recent computational results, including CC computations, are also available from X. Yang et al. [28]. As part of a comprehensive survey of all possible structures formally derivable from acetaldehyde by the “Scaled Hypersphere” search method, they obtained for the isomerization energy **3A** $\rightarrow$ **3B** a value of 10.67 kcal/mol including a Δ_{ZPE} of 0.69 kcal/mol, corresponding to $\Delta E = 9.98$ kcal/mol, at the level CCSD(T)/cc-pVTZ//B3LYP/6-311++G(d,p). This is well confirmed by our similar but formally higher level CCSD(T)/aug-cc-pVTZ//MP2/aug-cc-pVTZ data of 9.9 kcal/mol. It should be noted that they quote C_1 symmetry for their B3LYP/6-311++G(d,p) acetaldehyde equilibrium geometry which implies nonplanarity (not discussed by the authors). If true, this is certainly a slight defect of the B3LYP method but does not change the conclusions, as the difference in energy should be insignificant. Our MP2/aug-cc-pVTZ geometry is definitely planar as proven explicitly by the vibrational frequencies (that we did calculate in this case) having no imaginary value.

For the acetalimine $\leftrightarrow$ vinylamine system **4**, there are excellent results already from 1994 by Lammertsma and Prasad [30].

Applying the “Gaussian-1” (G1) and “Gaussian-2” (G2) complex extrapolation methods [31] they obtained an isomerization energy of $\Delta E = 3.9$ kcal/mol. Note that G1 and G2 include ZPE but this effect (for the energy difference) was below 0.1 kcal/mol. Their result compares reasonably well with our best values of 3.0–3.1 kcal/mol, but the difference is not quite negligible.

4. Relative energies of the monohydrates

The effect of water was studied for the formamide $\leftrightarrow$ formamidine acid system **1** in form of the monohydrates. The optimized structures including all the relevant geometric parameters are shown in Fig. 2, with the energies listed in Table 2 (formamidine in Fig. 1 will be used only in the section on the transition states below). The basic conclusion about system **1** is that attaching one water molecule to the tautomers decreases the energy difference significantly, by $\Delta\Delta = 1$ –2 kcal/mol, depending on the level of calculation. This is not quite evident because *bulk* water has the opposite effect [20], explained by the fact that formamide's dipole moment is much larger than that of formamidine acid (for example, in our calculations 3.8–4.0 D vs 0.9–1.1 D). In the monohydrate, however, the determining factor is the strength of the H-bonds. As seen in Fig. 2, both X...H bridges are indeed definitely shorter in the acidic form (Fig. 2c) than in the amide (Fig. 2a), 1.779 and 1.908 Å vs 1.877 and 2.037 Å, respectively. It should be noted that a microwave spectroscopic study by Lovas et al. [34] undertook the difficult task of determining the formamide–water structure experimentally. They fitted the MW data with constraints taken from early RHF theoretical structures, and gave two sets of results, specifically for the hydrogen bonds. In both results, the O3...H8 distance (their r_1) is longer than H5...O7 (r_2) which is certainly the wrong order, as was already observed in previous computational results [8,9,14].

As to the energy figures, our highest level, CCSD(T) calculations give $\Delta E = 9.2$ – 9.4 kcal/mol. In the literature, Markova and Enchev [9], at MP4/6-31G(d,p), obtained 10.73 kcal/mol including a Δ_{ZPE} of 0.47 kcal/mol. This gives then $\Delta E = 10.26$ kcal/mol. Their best result, at MP4/6-311++G(3df,2p)//MP2/6-31+G(d) + ZPE was $\Delta E = 10.75 - 0.45 = 10.30$ kcal/mol. With this latter, the decrease due to water, relative to the isolated pair is $\Delta\Delta = 11.40 - 10.30 = 1.1$ kcal/mol. This result is close to our CCSD(T) results of 1.4–1.5 kcal/mol. In fact, according to Table 2, even MP2 and B3LYP work very well for $\Delta\Delta$, with converged values at 1.6 and 1.7 kcal/mol.

5. Transition states

We have determined the transition states for formamide $\leftrightarrow$ formamidine acid **1** and formamidine $\leftrightarrow$ formamidine **2** (Fig. 1), these being smaller systems where the effect of water can still be

Table 2

Energies of the monohydrates of formamide and formamidine acid.^a

Method ^a	FMD-H ₂ O ^b E + 244	FAC-H ₂ O ^b E + 244	ΔE	$\Delta\Delta^c$
RHF				
/6-31G(d,p)	−0.97948	−0.96105	11.6	1.1
/6-311G(d,p)	−1.04392	−1.02582	11.4	1.1
/6-311++G(2d,2p)	−1.06285	−1.04473	11.4	0.6
/6-311++G(3df,3pd)	−1.07405	−1.05555	11.6	0.6
B3LYP				
/6-31G(d,p)	−2.33823	−2.32212	10.1	1.5
/6-311G(d,p)	−2.41401	−2.39754	10.3	2.6
/6-311++G(2d,2p)	−2.43579	−2.41898	10.5	1.7
/6-311++G(3df,3pd)	−2.44546	−2.42849	10.6	1.7
MP2				
/6-31G(d,p)	−1.66137	−1.64471	10.5	1.8
/6-311G(d,p)	−1.77752	−1.76305	9.1	2.2
/6-311++G(2d,2p)	−1.85834	−1.84249	9.9	1.6
/6-311++G(3df,3pd)	−1.95129	−1.93559	9.9	1.6
/PVTZ//~	−1.94194	−1.92713	9.3	2.0
/aug-PVTZ//~	−1.96581	−1.95022	9.8	1.6
/PVQZ//aug-PVTZ	−2.02677	−2.01137	9.7	1.8
/aug-PVQZ//aug-PVTZ	−2.03713	−2.02134	9.9	1.6
/PV5Z//aug-PVTZ	−2.05808	−2.04230	9.9	1.6
/aug-PV5Z//aug-PVTZ	−2.06265	−2.04673	10.0	1.6
CCSD/MP2^b				
/aug-PVTZ	−2.04476	−2.02949	9.6	1.1
/PVQZ	−2.15400	−2.13869	9.6	1.2
/aug-PVQZ	−2.16454	−2.14885	9.8	1.1
CCSD(T)/MP2^b				
/aug-PVTZ	−2.08278	−2.06808	9.2	1.4
/PVQZ	−2.19426	−2.17962	9.2	1.5
/aug-PVQZ	−2.20575	−2.19067	9.4	1.4

^a For units and notation see also Table 1.

^b FMD and FAC are shorthand notations for formamide and formamidine acid, respectively.

^c $\Delta\Delta = \Delta E(\text{free molecules}) - \Delta E(\text{monohydrates})$; the lowering of ΔE relative to its value in the absence of water; for the latter see column four of Table 1.

studied at high levels of theory. The energies in Table 3 are given for the free molecules as well as for the monohydrates. In the first case, formamide and formamidine are very close to each other, with converged values for the transition barrier of 47.1–47.2 and 47.6–47.8 kcal/mol, respectively. The effect of electron correlation, as expected, is more significant here for the transition states than was the case with equilibrium structures above. Between MP2 and CCSD, Δ_{TS} increases by up to 5 kcal/mol. The change is quite significant even between CCSD and CCSD(T), Δ_{TS} goes back by 3 kcal/mol.

Comparing the monohydrates with the free molecules, already previous studies have shown [3,5–8] that water drastically lowers the barrier. According to the data in Table 3, Δ_{TS} is reduced by more than a factor of two. The effect in **2** is slightly larger than in **1**. The

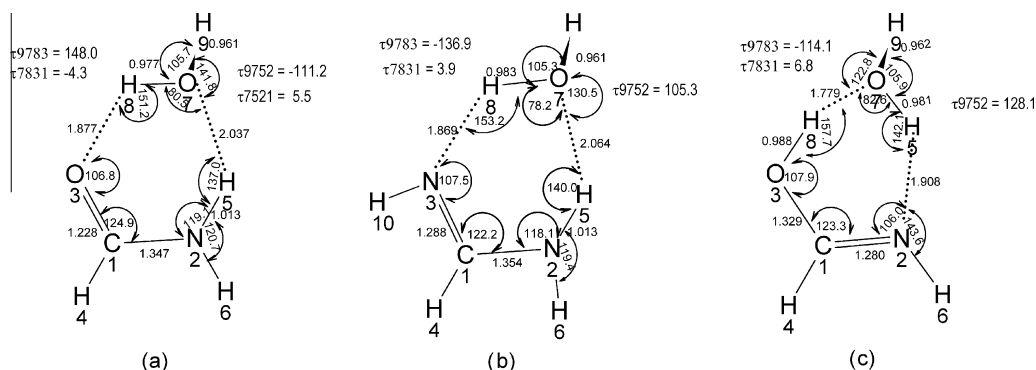


Fig. 2. Equilibrium structures optimized at MP2/aug-cc-pVTZ level: monohydrates of (a) formamide; (b) formamidine; (c) formamidine acid.

Table 3Transition state energies for tautomer pairs.^a

Method	FMD–FAC ^b <i>E</i> _{TS} + 167	Δ _{TS} ^c 1	(FMD–FAC)·H ₂ O ^b <i>E</i> _{TS} + 245	Δ _{TS} ^c 1	FIM ^b <i>E</i> _{min} + 149 ^d	FIM ^b <i>E</i> _{TS} + 149	Δ _{TS} 2	FIM·H ₂ O ^b <i>E</i> _{min} + 225 ^d	FIM·H ₂ O ^b <i>E</i> _{TS} + 225	Δ _{TS} 2
B3LYP										
/6-31G(d,p)	–2.82338	46.2	–1.30718	19.5	–1.01090	–0.93823	45.6	–1.45367	–1.42801	16.1
/6-311++G(2d,2p)	–2.88299	48.1	–1.40052	22.1	–1.06583	–0.98977	47.7	–1.54350	–1.51551	17.6
/6-311++G(3df,3pd)	–2.89097	47.7	–1.41037	22.0	–1.07186	–0.99666	47.2	–1.55192	–1.52459	17.1
MP2										
/6-31G(d,p)	–2.34658	46.8	–0.62557	22.5	–0.56288	–0.48803	47.0	–0.80413	–0.77303	19.5
/6-311++G(2d,2p)	–2.47062	47.4	–0.82158	23.1	–0.66823	–0.59244	47.6	–0.98191	–0.95111	19.3
/6-311++G(3df,3pd)	–2.53860	45.5	–0.91728	21.3	–0.72849	–0.65601	45.5	–1.07047	–1.04242	17.6
/PVTZ//~	–2.53331 (1894 cm ^{–1}) ^e	45.3	–0.90930 (1610 cm ^{–1}) ^e	20.5	–0.72209	–0.64943 (1925 cm ^{–1}) ^e	45.6	–1.06030 (1565 cm ^{–1}) ^e	–1.03150	18.1
/aug-PVTZ//~	–2.54854	45.3	–0.93230	21.0	–0.73684	–0.66421	45.6	–1.08363	–1.05589	17.4
/PVQZ//aug-PVTZ	–2.59031	45.4	–0.99318	21.1	–0.77349	–0.70087	45.6	–1.13936	–1.11101	17.8
/aug-PVQZ//aug-PVTZ	–2.59709	45.3	–1.00312	21.3	–0.77989	–0.70747	45.4	–1.14946	–1.12173	17.4
/PV5Z//aug-PVTZ	–2.61135	45.3	–1.02397	21.4	–0.79218	–0.71972	45.5	–1.16837	–1.14034	17.6
aug-PV5Z//aug-PVTZ	–2.61443	45.3	–1.02837	21.5	–0.79495	–0.72260	45.4	–1.17269	–1.14484	17.5
CCSD//MP2										
/aug-PVTZ	–2.60032	50.0	–1.00453	25.2	–0.80501	–0.72427	50.7	–1.17080	–1.13677	21.3
/PVQZ	–2.67723	50.0	–1.11353	25.4	–0.87403	–0.79316	50.7	–1.27213	–1.23750	21.7
/aug-PVQZ	–2.68423	50.0	–1.12393	25.5	–0.88063	–0.80005	50.6	–1.28237	–1.24856	21.2
PV5Z	–2.71257	50.1	–	–	–0.90701	–0.82606	50.8	–	–	–
/aug-PV5Z	–2.71611	50.1	–	–	–	–	–	–	–	–
CCSD(T)//MP2										
/aug-PVTZ	–2.63290	47.2	–1.04649	22.8	–0.83293	–0.75677	47.8	–1.20886	–1.17819	19.2
/PVQZ	–2.71169	47.1	–1.15793	22.8	–0.90362	–0.82743	47.8	–1.31238	–1.28113	19.6
/aug-PVQZ	–2.71926	47.1	–1.16920	22.9	–0.91075	–0.83486	47.6	–1.32353	–1.29317	19.1
PV5Z	–2.74835	47.2	–	–	–0.93778	–0.86156	47.8	–	–	–
/aug-PV5Z	–2.75210	47.1	–	–	–	–	–	–	–	–

^a For units and notation see also Table 1. The bold numbers in the heading refer to the systems defined in Fig. 1.^b FMD, FAC and FIM are shorthand notations for formamide, formamidic acid and formamidine, respectively.^c The difference relative to the lower-energy member of the pair, FMD (Table 1) and FMD·H₂O (Table 2).^d The minimum energy values of the optimized structures are also listed here as they are not present in the previous tables.^e Vibrational frequencies have been calculated at the MP2/PVTZ level. Each system has one single imaginary frequency as listed, indicating that the TS is indeed a first order saddle point.

optimized TS geometries of the hydrates are shown in Fig. 3. As compared to Fig. 2, it is clear that water binds much more strongly with the TS than with the minimum energy structures, thus becoming a catalyst for the tautomerization process. This is indicated by the very short X...H distances: in Fig. 3a *r*(O₇...H₈) of 1.234 Å is hardly longer than the formally bonded O₃–H₈ of 1.202 Å. (It is also seen that the TS is structurally closer to the acidic form than to the amide). Not listed in Fig. 3, the distance between the pillar atoms (O₇–N₂ and O₇–O₃) is 2.40 and 2.39 Å, respectively. As expected, the two bonds in the C–C–N skeleton indicate delocalization. Shown in Fig. 3b, the formamidine ↔ formamidine TS is, of course, symmetric (C_s symmetry group, with H₉O₇C₁H₄ defining the plane of symmetry). (It may be interesting

to note that in *free* formamidine, without water, the TS has C₂ symmetry, with the bridging hydrogen in the N–C–N plane, and one N–H hydrogen up, the other down.) The structure is otherwise quite similar to that in Fig. 3a, with slightly longer H-bond of 1.373 Å. The O...N pillar distance is 2.44 Å, close to the analogous values in the TS of formamide ↔ formamidic acid.

Acknowledgments

Financial support for this work has been provided by the Grant NKTH-OTKA-A07/K 68427 from the National Office for Research and Technology of Hungary and the Hungarian Scientific Research Fund.

References

- [1] J.D. Watson, F.H.C. Crick, *Nature* 171 (1953) 964–967.
- [2] See, for example G. Fogarasi, P.G. Szalay, *Chem. Phys. Lett.* 356 (2002) 383–390.
- [3] S. Kalia, A. Sharma, B.S. Kaith, *J. Chem. Sci.* 119 (2007) 617–624.
- [4] D.M. Du, M. Qin, A.P. Fu, Z.Y. Zhou, *Pol. J. Chem.* 81 (2007) 369–380.
- [5] W.C. Liang, H.R. Li, X.B. Hu, S.J. Han, *Acta Phys. Chim. Sin.* 22 (2006) 797–803.
- [6] M.K. Hazra, T. Chakraborty, *J. Phys. Chem. A* 109 (2005) 7621–7625.
- [7] W.C. Liang, H.R. Li, X.B. Hu, S.J. Han, *J. Phys. Chem. A* 108 (2004) 10219–10224.
- [8] N. Markova, V. Enchev, *J. Mol. Struct. (THEOCHEM)* 679 (2004) 195–205.
- [9] M.-H. Liu, C. Chen, C.-W. Liu, *Struct. Chem.* 15 (2004) 309–315.
- [10] D.M. Du, A.P. Fu, Z.Y. Zhou, *Int. J. Quantum Chem.* 99 (2004) 1–10.
- [11] Y.L. Jin, X.L. Mu, B.L. Zhang, X.Y. Wang, *Chin. J. Chem. Phys.* 16 (2003) 445–454.
- [12] A.P. Fu, H.L. Li, D.M. Du, Z.Y. Zhou, *Chem. Phys. Lett.* 382 (2003) 332–337.
- [13] Y. Kim, S. Lim, H.J. Kim, Y. Kim, *J. Phys. Chem. A* 103 (1999) 617–624.
- [14] C. Adamo, M. Cossi, V. Barone, *J. Comput. Chem.* 18 (1997) 1993–2000.
- [15] R.L. Bell, D.L. Taveras, T.N. Truong, J. Simons, *Int. J. Quantum Chem.* 63 (1997) 861–874.
- [16] C. Adamo, F. Lelj, *Int. J. Quantum Chem.* 56 (1995) 645–653.
- [17] J. Pranata, G.D. Davis, *J. Phys. Chem.* 99 (1995) 14340–14346.

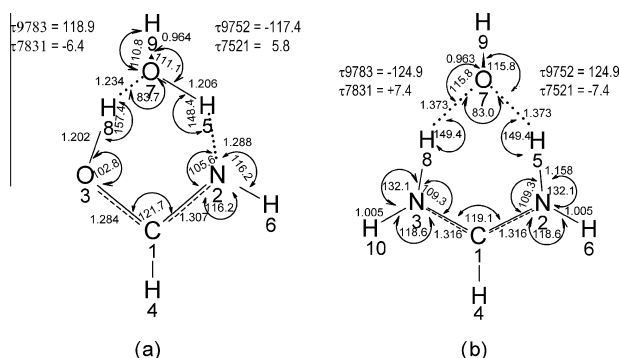


Fig. 3. Transition state structures optimized at MP2/aug-cc-pVTZ level for mono-hydrates of (a) formamide ↔ formamidic acid; (b) formamidine ↔ formamidine (two identical forms).

- [18] O.N. Ventura, J.B. Rama, L. Turi, J.J. Dannenberg, J. Phys. Chem. 99 (1995) 131–136.
- [19] J.S. Kwiatkowski, J. Leszczynski, J. Mol. Struct. 270 (1992) 67–86.
- [20] M.W. Wong, K.B. Wiberg, M.J. Frisch, J. Am. Chem. Soc. 114 (1992) 1645–1652.
- [21] X.C. Wang, J. Nichols, M. Feyereisen, M. Gutowski, J. Boatz, A.D.J. Haymet, J. Simons, J. Phys. Chem. 95 (1991) 10419–10424.
- [22] J.S. Kwiatkowski, R.J. Bartlett, W.B. Person, J. Am. Chem. Soc. 110 (1988) 2353–2358.
- [23] H.B. Schlegel, P. Gund, E.M. Fluder, J. Am. Chem. Soc. 104 (1982) 5347–5351.
- [24] M.H. Almatarnah, C.G. Flinn, R.A. Poirier, Can. J. Chem. 83 (2005) 2082–2090.
- [25] Y. Podolyan, L. Gorb, J. Leszczynski, J. Phys. Chem. A 106 (2002) 12103–12108.
- [26] Y. Kim, S. Lim, Y. Kim, J. Phys. Chem. A 103 (1999) 6632–6637.
- [27] R.L. Bell, T.N. Truong, J. Phys. Chem. A 101 (1997) 7802–7808.
- [28] X. Yang, S. Maeda, K. Ohno, J. Phys. Chem. A 111 (2007) 5099–5110.
- [29] J. Andrés, L.R. Domingo, M.T. Pitcher, V.S. Safont, Int. J. Quantum Chem. 66 (1998) 9–24.
- [30] K. Lammertsma, B.V. Prasad, J. Am. Chem. Soc. 116 (1994) 642–650.
- [31] F. Jensen, Introduction to Computational Chemistry, John Wiley & Sons, Chichester, New York, 1999.
- [32] PQS *Ab Initio* Program Package, version 2.4, Parallel Quantum Solutions, 2013 Green Acres, Suite A, Fayetteville, Arkansas 72703, USA.
- [33] CF4UR, a quantum chemical program package written by J.F. Stanton, J. Gauss, M.E. Harding, P.G. Szalay, with contributions from: A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, O. Christiansen, M. Heckert, O. Heun, C. Huber, D. Jonsson, J. Jusélius, K. Klein, W.J. Lauderdale, D. Matthews, T. Metzroth, D.P. O'Neill, D.R. Price, E. Prochnow, K. Ruud, F. Schiffmann, S. Stopkiewicz, A. Tajti, M.E. Varner, J. Vázquez, F. Wang, J.D. Watts; and the integral packages MOLECULE (J. Almlöf, P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J.Aa. Jensen, P. Jørgensen, J. Olsen), and ECP routines by A.V. Mitin, C. van Wüllen. For the current version, see <http://www.cfour.de>.
- [34] F.J. Lovas, R.D. Suenram, G.T. Fraser, J. Chem. Phys. 88 (1988) 722–729.