

On the Keto–Enol Tautomerization of Malonaldehyde: An Effective Fragment Potential Study

Mark A. Freitag,* Trenton L. Pruden, David R. Moody, James T. Parker, and Marcel Fallet

Department of Chemistry, Creighton University, 2500 California Plaza, Omaha, Nebraska 68178

Received: September 13, 2006; In Final Form: January 5, 2007

Recent studies have mapped the keto–enol tautomerization of malonaldehyde through a general transition structure that leads exclusively to the *Z* isomer of the enol. However, it will be shown that a competing general transition structure exists that leads to both the *E* and *Z* isomers of the enol at the B3LYP/6-31G(d,p) and MP2/6-31G(d,p) levels of theory. Both the RHF- and DFT-based effective fragment potential methods have been used to model solvation effects, and the results are compared with full ab initio calculations. It is found that two bridging water molecules with two discrete DFT-based effective fragment potential solvent waters at the MP2/6-31G(d,p) level of ab initio theory provides the most computationally effective model for solvent effects in this system. It is shown that the relative energies for this QM/MM model differ from the full MP2/6-31G(d,p) energies by an average absolute relative difference of 2.2 kcal mol⁻¹ across the reaction path when the zero-point vibrational energy correction is included.

I. Introduction

There has been recent interest in the acid-catalyzed mechanism for the keto–enol tautomerization of malonaldehyde. Yamabe et al.¹ have examined this reaction, which ultimately leads to an extensively studied *Z* isomer of the enol containing an intramolecular hydrogen bond.² In their study, Yamabe and co-workers focused exclusively on a tautomerization mechanism that involves what will be referred to here as a “cis” transition structure, where the carbonyl groups are adjacent to one another, as shown in Figure 1. This general transition scheme leads directly to the stabilized, intramolecular hydrogen bonded form of the *Z* isomer.

However, though this *Z* isomer is the most stable form of the enol in the gas phase, other isomers are important. Chivassa et al.³ have used high-resolution FT–IR spectroscopy to show that, of the seven other possible structural isomers of the enol, four are principally formed in the gas phase. Of these, three are *E* isomers and one is a *Z* isomer. In addition, the relative stability of the intramolecular hydrogen-bonded structure in aqueous solution is diminished because of competing hydrogen bond formation with solvent waters.¹

The formation of the *E* isomer of the enol is highly unlikely through the “cis” transition structure, because of a necessary rotation about a carbon–carbon double bond. Instead, one might consider the keto–enol tautomerization of malonaldehyde through a “trans” scheme, as shown in Figure 2. In this structure, the carbonyl groups are generally opposite one another. Although the stabilizing intramolecular hydrogen bonded form of the *Z*-enol is not formed directly through the “trans” structure, it can be recovered through rotation about the appropriate carbon–carbon single bond.

The purpose of the present work is 3-fold. First, we extend the important work of Yamabe et al., who have considered the keto–enol tautomerization of malonaldehyde at the B3LYP/6-31G(d) level of theory with the PCM continuum model for solvation, but focusing only on the “cis” transition

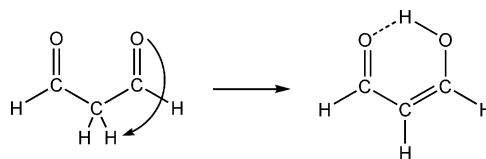


Figure 1. General “cis” transition scheme for the keto–enol tautomerization of malonaldehyde.

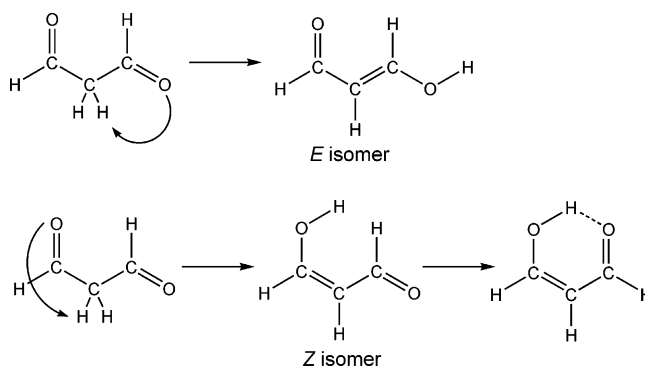


Figure 2. General “trans” transition scheme for the keto–enol tautomerization of malonaldehyde. Through this general scheme, either the *E* or the *Z* isomer of the enol can be formed.

state structure. In the present work, we focus on the “trans” transition state structures, and the *Z*- and *E*-enol isomers that result. To that end, we consider the tautomerization mechanism with one, two, and three bridging waters, which allow for the transfer of the hydrogen from the α carbon to the appropriate oxygen. Once the optimal number of bridging waters has been identified, solvent effects are considered by adding discrete waters to the bridged system. Second, we examine the performance of the new DFT-based effective fragment potential (EFP1/DFT) method^{4,5} as compared to the well-established Hartree–Fock-based^{6,7} effective fragment potential method (EFP1/HF). These QM/MM results are also compared with full ab initio B3LYP and MP2 calculations. Finally, the keto–enol tautomerization of malonaldehyde is the first step in a proposed in-vivo mechanism⁸ for the Mannich⁹ reaction. It is believed

* Corresponding author. E-mail: freitag@creighton.edu. Phone: (402) 280-2274. Fax: (402) 280-5737.

TABLE 1: Activation Energies (E_a) and Reaction Energies (ΔE) (in kcal mol⁻¹) for the Paths Leading to the *E* and *Z* Enol Isomers Discussed in the Text.

isomer	level of theory/basis	one bridging water		two bridging waters		three bridging waters	
		E_a	ΔE	E_a	ΔE	E_a	ΔE
<i>E</i>	RHF/6-31G(d,p)	47.46	3.64	39.56	4.36	38.30	3.72
	B3LYP/6-31G(d,p)	25.12	-0.48	15.27	-0.97	12.18	-2.33
	MP2/6-31G(d,p)	32.36	3.53	22.70	2.57	19.77	1.16
	RHF/6-311+G(2d,p)//RHF/6-31G(d,p)	50.99	3.26	43.86	4.48		
	B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p)	30.23	-2.48	21.13	-1.83		
	MP2/6-311+G(2d,p)//MP2/6-31G(d,p)	32.87	0.84	23.92	0.37		
<i>Z</i>	RHF/6-31G(d,p)	49.44	5.40	43.03	3.02	n/a	n/a
	B3LYP/6-31G(d,p)	26.16	1.11	17.21	-4.59	n/a	n/a
	MP2/6-31G(d,p)	32.97	6.00	24.40	-0.44	n/a	n/a
	RHF/6-311+G(2d,p)//RHF/6-31G(d,p)	51.53	4.30	46.33	1.92	n/a	n/a
	B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p)	31.62	0.45	23.18	-5.58	n/a	n/a
	MP2/6-311+G(2d,p)//MP2/6-31G(d,p)	33.36	3.30	25.09	-3.15	n/a	n/a

^a Single-point 6-311+G(2d,p)//6-31G(d,p) energies are not ZPVE corrected.

that this reaction plays a role in ethanol metabolism in the liver, a process that may ultimately be a cause of alcohol-related liver damage.¹⁰ Therefore, we identify an appropriate EFP-based QM/MM model for examining the solvated mechanism for the full Mannich reaction in a neutral medium. This choice of a non-base-catalyzed mechanism also makes direct comparison with Yamabe's previous work possible.

II. Methods

To facilitate comparison with the work of Yamabe et al., the all-electron 6-31G(d,p)¹¹ and 6-311+G(2d,p) basis sets were used for all atoms. (*p* functions have been added to Yamabe's choice for a double- ζ basis set to better model the extensive hydrogen bonding in this system, and also to compliment the basis set used to construct the fragments for the EFP1/HF method.) Analytical Hessians were obtained at the Hartree-Fock level of theory, and the double-differenced Hessians were calculated at the B3LYP and MP2 levels of theory. RHF, DFT (B3LYP),¹² and MP2 calculations were carried out using the GAMESS¹³ suite of programs. The effect of solvation was modeled using the EFP-1 method optimized for both Hartree-Fock^{6,7} and DFT (B3LYP)⁴ theory (hereafter referred to as EFP1/HF and EFP1/DFT, respectively). The EFP1/DFT model can employ DFT or MP2 methods in the active region; this is referred to as B3LYP-EFP1/DFT and MP2-EFP1/DFT, respectively. The present work uses only RHF wave functions with the EFP1/HF solvation model, which is referred to as RHF-EFP1/HF. All saddle points were confirmed with an appropriate Hessian calculation, revealing one imaginary harmonic normal-mode frequency. To confirm the connection between the suspected transition state and the desired reactants and products, the single imaginary frequencies were followed backward and forward using the intrinsic reaction coordinate (IRC) method,¹⁴ except in one case discussed below. All energies include the unscaled zero-point vibrational energy (ZPVE) correction, except where noted. Chemical intuition was used to place EFP waters in the first solvation shell for this small system with its strong hydrogen bond acceptor sites. All chemical structures were visualized using MacMolPlt.¹⁵

III. Results

A. No Solvent Waters. 1. Basis Set Effects. The activation energies and reaction energies for the transition states leading to both the *E* and *Z* isomers of the enol with one bridging water, to facilitate the hydrogen transfer, are given in Table 1. For the present work, we follow Yamabe and define the activation energy (E_a) as the ZPVE-corrected energy difference between

TABLE 2: Differences between Non-ZPVE-Corrected 6-311+G(2d,p)//6-31G(d,p) Energies and ZPVE-Corrected 6-311+G(2d,p) Energies (kcal mol⁻¹) for the Path Leading to the Formation of the *E* Isomer of the Enol^a

	one bridging water		two bridging waters	
	E_a	ΔE	E_a	ΔE
RHF	-1.09	1.31	-0.56	0.53
B3LYP	-1.32	2.10	-1.55	1.77
MP2	-1.20	1.63	-1.52	1.44

^a No structure corresponding to three bridging waters was found.

the appropriate solvated keto-malonaldehyde and the highest energy first-order saddle point. Reaction energies (ΔE) are similarly defined as the ZPVE-corrected energy difference between the solvated enol-malonaldehyde and the solvated keto-malonaldehyde.

In addition to the results shown, fully optimized 6-311+G(2d,p) structures were also found for the *E* isomer with one and two bridging waters. A comparison of the non-ZPVE-corrected single-point energies (shown in Table 1) versus the fully optimized triple- ζ geometries revealed an average absolute difference of only 0.24, 0.51, and 0.38 kcal mol⁻¹ at the stationary points for the RHF, B3LYP, and MP2 levels of theory, respectively. Table 2 shows the difference in E_a and ΔE between the ZPVE-corrected triple- ζ geometries and non-ZPVE-corrected triple- ζ single-point energies at double- ζ geometries. The difference is remarkably consistent, and near 1.5 kcal mol⁻¹ for all levels of theory. It is therefore concluded that single-point triple- ζ energies at double- ζ geometries are a good indicator of full triple- ζ energetics, and further expensive full triple- ζ calculations were discontinued.

As seen in Table 1, the *E* isomer activation energy differences between the double- ζ and triple- ζ single-point structures at the RHF level of theory are 3.5 kcal mol⁻¹, 5.1 kcal mol⁻¹ for B3LYP, and only 0.5 kcal mol⁻¹ for MP2. For all levels of theory, E_a was lower for the double- ζ structure. The differences in ΔE were 2.7 kcal mol⁻¹ or less for all methods, but in this case the double- ζ values were high compared to the triple- ζ single points. (The bulk of this difference goes away when one compares the double- ζ geometries with the fully optimized triple- ζ geometries; there, the difference in ΔE is 1.0 kcal mol⁻¹ or less for all methods.) For the *Z* isomer, the full double- ζ energies were compared to the triple- ζ single-point energies and revealed a difference of 2.1, 5.5, and 0.4 kcal mol⁻¹ for the RHF, B3LYP, and MP2 levels of theory, respectively. Again, the differences in ΔE were 2.7 kcal mol⁻¹ or less for all methods.

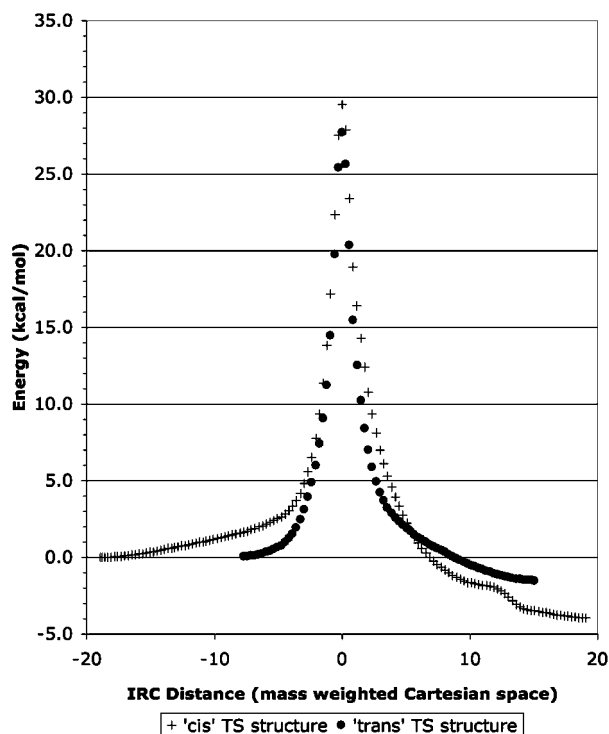


Figure 3. IRC pathways at the B3LYP/6-31G(d,p) level of theory for the “cis” and “trans” configurations. The zero of energy has been defined as the reactant “cis” geometry. The structures corresponding to these paths are shown in Figures 4 and 5.

These basis set investigations with one and two bridging waters suggest that the E_a accuracy in the DFT and MP2 methods using the 6-31G(d,p) basis set, compared to 6-311+G-(2d,p), are on the order of 6 kcal mol⁻¹ for the B3LYP method, and 1 kcal mol⁻¹ for MP2. The accuracy in ΔE is on the order of 1 kcal mol⁻¹ for B3LYP, and 3 kcal mol⁻¹ for MP2. Although the difference in the B3LYP method is troubling, it may be an artifact of the single-point approximation to the full triple- ζ energies: for the two reaction paths where fully optimized triple- ζ energetics are available (the *E* isomer with one and two bridging waters), the B3LYP difference relative to the double- ζ structures drops to 4 kcal mol⁻¹. This is consistent with Yamabe’s finding that the double- ζ /triple- ζ difference is on the order of 4 kcal mol⁻¹ for the “cis” transition structure.

For the purposes of comparing with previous work that employed the B3LYP/6-31G(d) level of theory, and because of the relatively small basis set effects discussed above, all of the following results are for fully optimized 6-31G(d,p) structures with a ZPVE correction included. Because the basis set difference for the MP2 method appears to be on the order of 1–3 kcal mol⁻¹, these second-order calculations are used as a benchmark throughout this work to gauge the performance of the B3LYP method, where the basis set differences for the activation energy (relative to triple- ζ) are higher.

2. One Bridging Water Molecule. The clearest evidence demonstrating that the “trans” transition state (TS) structure should be considered as well as the “cis” structure is seen in Figure 3: using the B3LYP/6-31G(d,p) level of theory, the “cis” and “trans” IRC paths between the keto and enol forms of malonaldehyde with one bridging water molecule have been plotted. The “trans” TS structure is more stable than the “cis” by 1.8 kcal mol⁻¹, yet both transition states lead back to the same reactant geometry, as shown in Figures 4 and 5. To form the “cis” transition state geometry, the reactant must first rotate

about the C2–C3 bond. This rotation requires about 2–3 kcal mol⁻¹, as shown by the IRC plot in Figure 3. By contrast, the geometry of the malonaldehyde is virtually unchanged in going from reactant to the “trans” TS. Assuming Arrhenius-type behavior and using the definition of activation energy given above, E_a is 1.9 kcal mol⁻¹ less for the “trans” TS, which leads to a ratio of $k_{\text{trans}}/k_{\text{cis}} = 24.7$ at 298 K. For one bridging water, the “trans” transition state must be considered as a competing, if not slightly favored, reaction pathway. It should be pointed out that small errors in the predicted activation energies (on the order of 1–2 kcal/mol) can lead to significant errors in rate constants, and care should be taken when considering kinetic questions.

Yamabe et al. used the B3LYP/6-31G(d) level of theory (no ZPVE-correction was indicated) coupled with Onsager’s self-consistent reaction field method with a dielectric constant of 78.39 for water. They found an activation energy of 31.1 kcal mol⁻¹ for the “cis” transition state with one bridging water. The “trans” TS structure that leads to the *E* isomer has a very comparable activation energy, as shown in Table 1: the B3LYP/6-31G(d,p) is 6.0 kcal mol⁻¹ lower than the “cis” structure, and the MP2/6-31G(d,p) is 1.3 kcal mol⁻¹ higher. For the formation of the *Z* isomer, the differences of the “trans” relative to the “cis” activation energies are 5.0 kcal mol⁻¹ lower and 1.9 kcal mol⁻¹ higher, respectively, for B3LYP and MP2.

3. Two Bridging Water Molecules. For the formation of the *E* isomer through a transition state with two bridging waters, the activation energy drops to 15.3 and 22.7 kcal mol⁻¹ for the B3LYP and MP2 methods, respectively. For the *Z* isomer, the barriers are 17.2 and 24.4 kcal mol⁻¹, respectively. These compare favorably to the activation energy of 20.4 kcal mol⁻¹ found by Yamabe et al. for the “cis” transition state with two bridging waters. The reaction energies are tabulated in Table 1.

4. Three Bridging Water Molecules. Despite numerous efforts, no three bridging water transition state was found leading to the *Z* isomer; instead, the two-water bridge was always recovered. Indeed, the energetic benefit of adding the third water is small even for the *E* isomer: the activation energy only dropped by 3.1 kcal mol⁻¹ for B3LYP (to 12.2 kcal mol⁻¹) and 2.9 kcal mol⁻¹ for MP2 (to 19.8 kcal mol⁻¹). For the “cis” TS structure, Yamabe et al. found a barrier of 13.1 kcal mol⁻¹.

As discussed in section IV, these results support Yamabe’s general conclusion that a two-water hydrogen transfer bridge is a good model for the tautomerization process. We therefore investigate the effect of solvation on the two bridging-water structure by adding discrete solvent waters in the form of EFP waters based on the RHF and B3LYP levels of theory.

B. Two Bridging Waters and One Solvent Water. With the addition of a single solvent water molecule, as shown in Figures 6 and 7, the potential energy surface becomes more interesting (see Figures 8 and 9). At the RHF-EFP1/HF, MP2-EFP1/DFT, full B3LYP, and full MP2 levels of theory, there is a very shallow minimum in the potential energy surface corresponding to a stable hydronium ion. As with all stationary points, these points were confirmed by either an analytical Hessian or a double-differenced numerical Hessian. However, for all levels of theory, the minimum is shallow enough that the addition of the ZPVE energy lifts the “minimum” above the level of at least one of the two adjacent transition states. These two general transition state structures are denoted TS1, which corresponds to the motion of H7 between C2 and H10; and TS2, which corresponds to the coupled motion of H12 between O10 and O13, and H14 between O13 and O4 for the

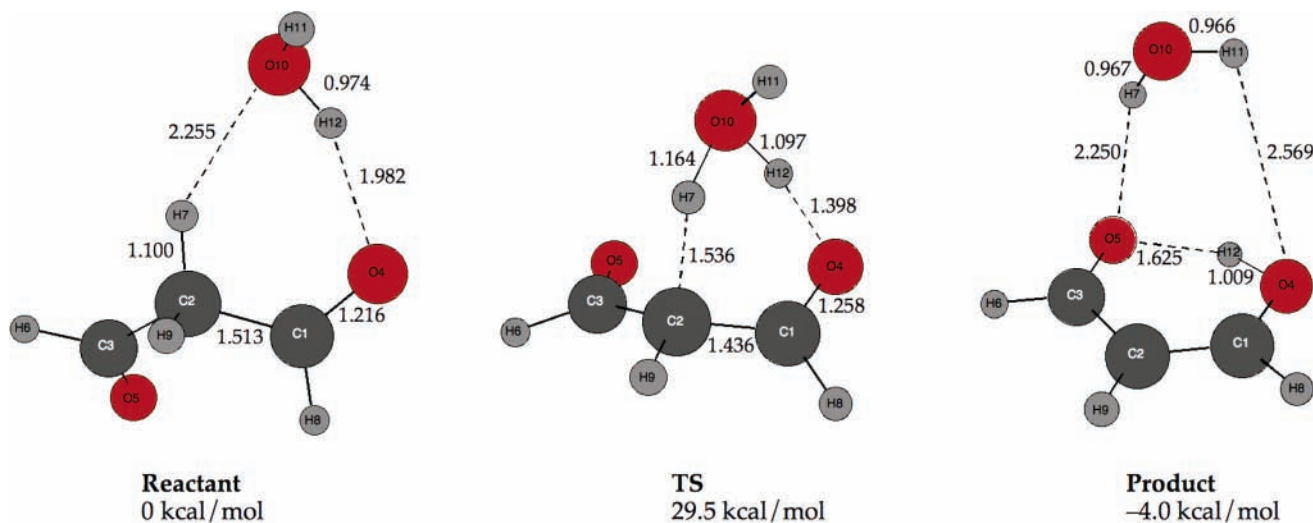


Figure 4. Reactant, transition state, and product geometries for the “cis” transition state structure from the IRC pathway shown in Figure 3.

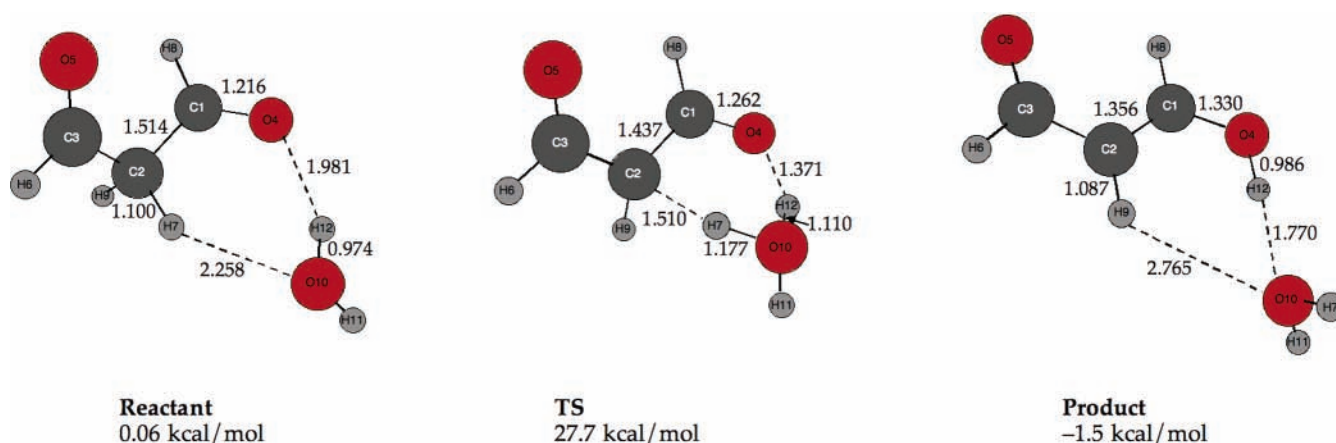


Figure 5. Reactant, transition state, and product geometries for the “trans” transition state structure from the IRC pathway shown in Figure 3.

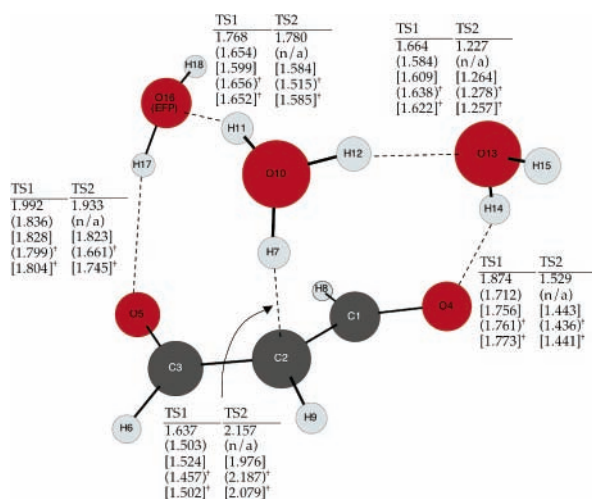


Figure 6. Saddle point for two bridging waters and one EFP solvent water at the RHF-EFP1/HF/6-31G(d,p), B3LYP-EFP1/DFT/6-31G(d,p) (parentheses), MP2-EFP1/DFT/6-31G(d,p) (brackets), B3LYP/6-31G(d,p) ([†]), and MP2/6-31G(d,p) ([‡]) levels of theory leading to the *E* isomer. Bond distances are given in angstroms (Å).

formation of the *E* isomer (see Figure 6). For the formation of the *Z* isomer, the coupled TS2 motion is H11 between O10 and O13 and H14 between O13 and O5 (see Figure 7).

The only exception to this qualitative picture is at the B3LYP-EFP1/DFT level of theory, where no intermediate and no TS2 structures were found. This may be due to the fact that this

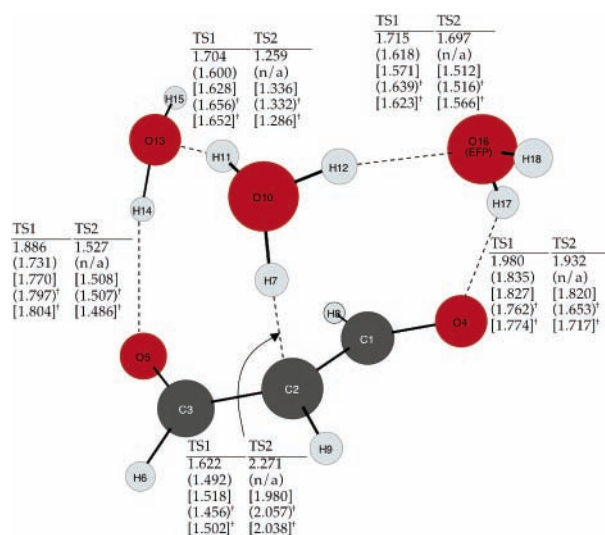


Figure 7. Saddle point for two bridging waters and one EFP solvent water at the RHF-EFP1/HF/6-31G(d,p), B3LYP-EFP1/DFT/6-31G(d,p) (parentheses), MP2-EFP1/DFT/6-31G(d,p) (brackets), B3LYP/6-31G(d,p) ([†]), and MP2/6-31G(d,p) ([‡]) levels of theory leading to the *Z* isomer. Bond distances are given in angstroms (Å).

method underestimates the hydrogen bond distances between H12 and O13, and again between H14 and O4 in the first transition state, as seen in Figure 6. The first transition state is therefore biased more toward the final products than the other

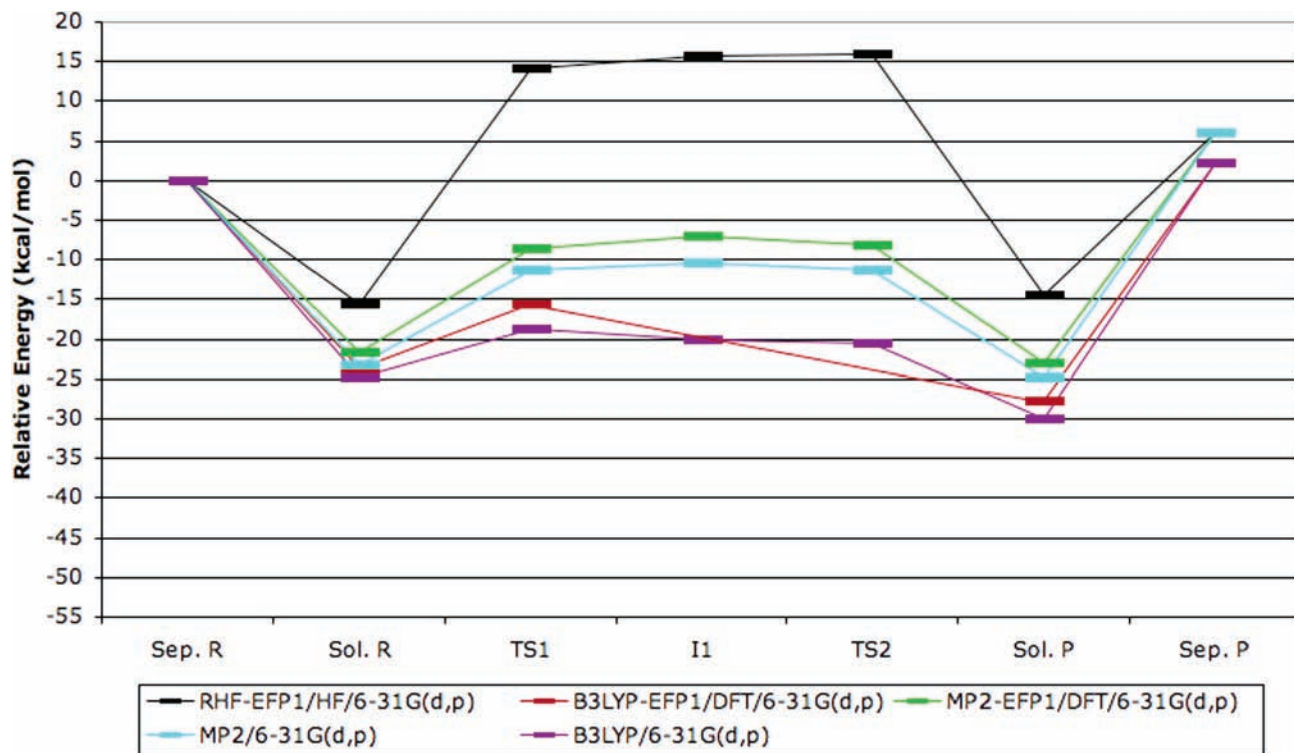


Figure 8. Potential energy surfaces relative to the separated reactants for two bridging waters and one solvent water leading to the *E* isomer. The transition state structures are shown in Figure 6.

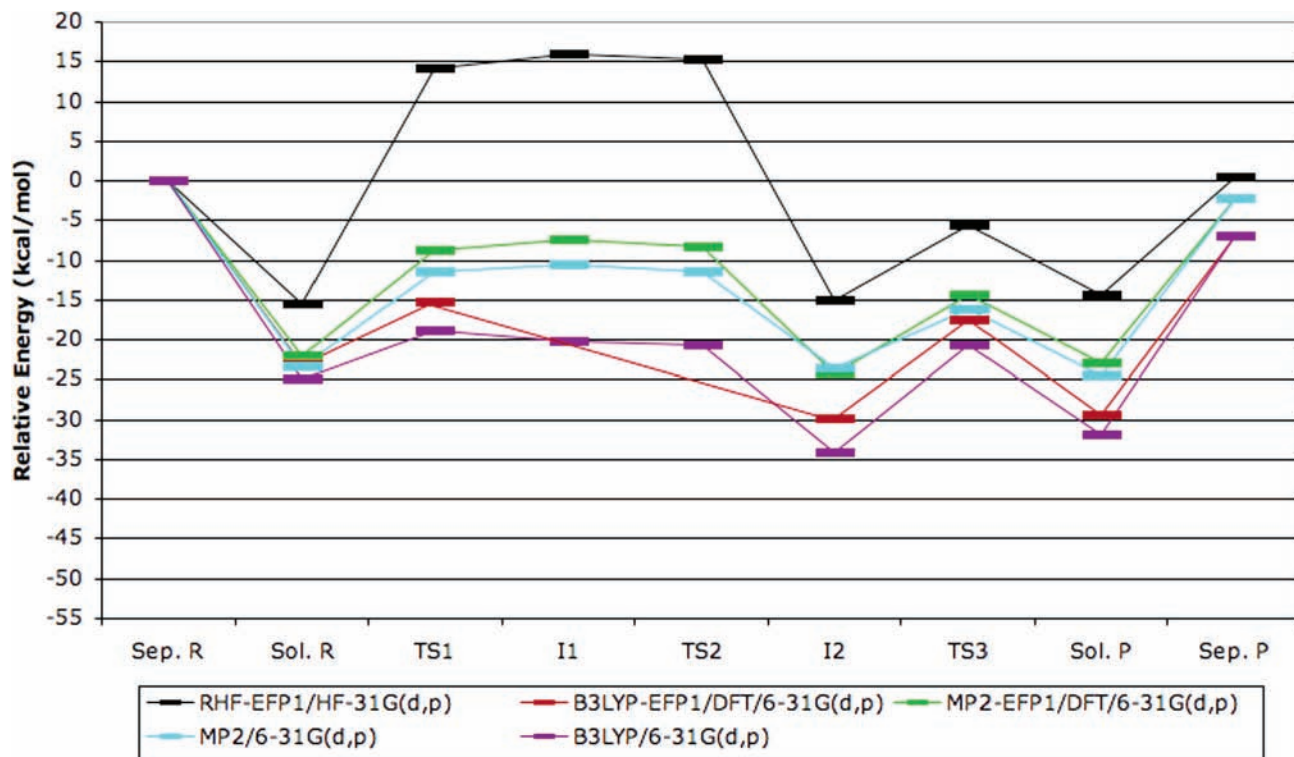


Figure 9. Potential energy surfaces relative to the separated reactants for two bridging waters and one solvent water leading to the *Z* isomer. The transition state structures are shown in Figure 7.

methods, and although the IRC was very shallow energetically, it never located an intermediate minimum.

This general PES landscape was also found for the formation of the *Z* isomer, as seen in Figure 9. This path allows for the formation of an intramolecular hydrogen bond through a carbon–carbon single bond rotation. The additional transition state corresponding to this rotation (TS3) is shown, making the full path completely comparable to that considered by Yamabe.

Again, at all levels of theory save B3LYP-EFP1/DFT, a shallow intermediate corresponding to a hydronium ion is found on the potential energy surface, but the addition of the ZPVE correction shows the intermediate to be unphysical. In the numbers that follow, the activation energy is defined as the difference in ZPVE-corrected energy between the higher of the two transition states (if two exist) and the energy of the solvated keto form of malonaldehyde.

Focusing on the EFP results, the addition of a single solvent water reduces the activation barrier to the formation of the *E* isomer by 7.3 kcal mol⁻¹ relative to the unsolvated mechanism, to an E_a of 8.0 kcal mol⁻¹ using B3LYP. For the MP2 method, the activation energy drops by 9.1 kcal mol⁻¹ over the unsolvated, to 13.6 kcal mol⁻¹. For the formation of the *Z* isomer, the B3LYP activation energy drops by 9.5 kcal mol⁻¹, to 7.7 kcal mol⁻¹, whereas the MP2 drops by 13.8 kcal mol⁻¹ to an E_a of 13.6 kcal mol⁻¹. These results are similar to the activation barrier of 13.6 kcal mol⁻¹ found by Yamabe through the “cis” transition state.

Although the performance of the EFP methods is the primary focus of this work, it is interesting to note the landscape of the IRC paths for the fully ab initio B3LYP/6-31G(d,p) and MP2/6-31G(d,p) methods leading to the *E* and *Z* isomers. As can be seen from the geometries shown in Figures 6 and 7, the TS1 transition states are almost certainly identical for both the *E* and *Z* paths; indeed, the non-ZPVE-corrected energies only differ by 0.015 kcal mol⁻¹, strongly suggesting that the two points are identical. In addition, the I1 intermediates are also the same, differing by only 0.033 kcal mol⁻¹. The second transition state for the *Z* path is only 0.006 kcal mol⁻¹ above the intermediate I1 and is characterized by an imaginary frequency of only 145 cm⁻¹. These two points are so close energetically that it was impossible to find an IRC path between them (IRC searches displaced both directions along the imaginary frequency always led to the enol product). The MP2 path displays a similar behavior, although an IRC path between the second transition state and the intermediate could be found. In this case, the difference between the two intermediates is only 0.002 kcal mol⁻¹, and there is no question that they are the same point on the PES.

C. Two Bridging Waters and Two Solvent Waters. The PES for two bridging waters and two EFP solvent waters is qualitatively similar to that for one EFP solvent water, but the activation energy leading to the *E* isomer is reduced by 3.2 kcal mol⁻¹ for both B3LYP and MP2 methods relative to the corresponding single solvent water structures. For the B3LYP surface, E_a is reduced to 4.8 kcal mol⁻¹, and for the MP2 surface the E_a is 10.4 kcal mol⁻¹. For the formation of the *Z* isomer, the barrier drops by 1.9 kcal mol⁻¹ for B3LYP, to 5.8 kcal mol⁻¹, and the MP2 drops by 3.0 kcal mol⁻¹, to 10.6 kcal mol⁻¹. By way of comparison, Yamabe et al. report a barrier of 7.4 kcal mol⁻¹ for the “cis” transition state.

D. Two Bridging Waters and Three Solvent Waters. The potential energy surfaces for two bridging waters and three EFP solvent waters are qualitatively similar to those for the two EFP solvent waters. One notable exception to this generalization is the presence of a second transition state (TS2) for the B3LYP-EFP1/DFT path leading to the *E* isomer. Of the six paths investigated, this is the only one using the B3LYP-EFP1/DFT method that found the second transition state. The effect of the third discrete solvent water on the activation energy is minimal: for the formation of the *E* isomer, E_a drops by 0.5 and 1.5 kcal mol⁻¹ for the B3LYP and MP2 methods, respectively, and it *increases* by 0.2 and 0.4 kcal mol⁻¹, respectively, for the formation of the *Z* isomer. These results are consistent with those found by Yamabe for the “cis” structure; they found an E_a drop of only 0.3 kcal mol⁻¹ in going from two to three solvent waters.

IV. Discussion

A. QM/MM Model. It has been shown previously that B3LYP-EFP1/DFT method tends to overbind the transition

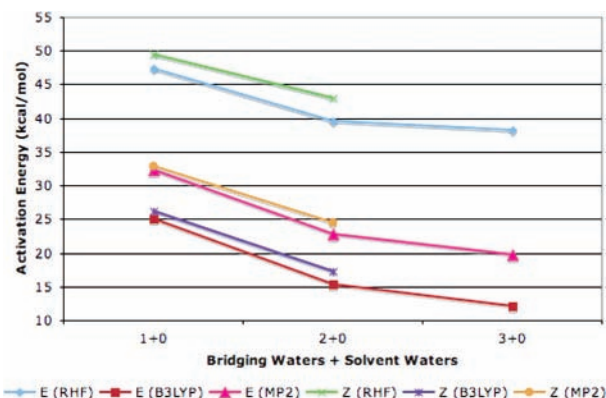


Figure 10. Activation energies as a function of the number of bridging water molecules for the formation of the *E* and *Z* isomers.

state,⁵ leading to activation barriers that are consistently lower than MP2-EFP1/DFT. Although the B3LYP-EFP1/DFT method differs qualitatively from the full B3LYP method (the former not locating the shallow intermediate on the potential energy surface), the MP2-EFP1/DFT method matches the full MP2 path very well for the formation of the *E* isomer, overestimating the relative energies by an average of 2.5 kcal mol⁻¹ across the reaction path, as seen in Figure 8. For the formation of the *Z* isomer, this method again performs well, with an absolute average deviation of 2.0 kcal mol⁻¹ across the reaction path, as seen in Figure 9. Again, all energies at all stationary points are overestimated except for the second intermediate (I2), which is underestimated by 0.8 kcal mol⁻¹. These results are consistent with the recommendation by Adamovic and Gordon⁵ that the MP2-EFP1/DFT model be preferred over B3LYP-EFP1/DFT.

Figure 10 shows the lowering of the activation energies for the formation of the *E* and *Z* isomers as the number of bridging waters increases. (Again, no stable three-water bridge was found for the *Z* isomer.) Relative to the correlated methods, Hartree–Fock clearly does not perform well, consistently overestimating the activation barriers for well-known reasons.¹⁶ Focusing then on the B3LYP and MP2 results, the addition of a second bridging water drops the barrier by an average of 9.8 kcal mol⁻¹ for the *E* isomer, and an average of 8.8 kcal mol⁻¹ for the *Z*. The addition of a third bridging water for the *E* isomer only drops the activation energy by an average of 3.0 kcal mol⁻¹, which is within the accuracy we can expect to achieve by using a double- ζ basis set, as compared to a triple- ζ . Therefore, two bridging waters are taken to be the optimal model for hydrogen transfer in malonaldehyde for all transition structures considered to date (both “cis” and “trans”).

Figure 11 shows the effect of additional EFP solvent water molecules on the two bridging water model. Again, the Hartree–Fock barriers are too high, as expected.¹⁶ For the B3LYP and MP2 methods, the addition of a solvent water molecule lowers the barrier by about 8 kcal mol⁻¹ for the *E* isomer and about 10 kcal mol⁻¹ for the *Z*. There is less benefit to be gained by adding a second solvent water, but still important given the expected accuracy of the double- ζ basis set: the E_a lowers by about 3 kcal mol⁻¹ for both isomers. The addition of a third solvent water does not have any significant affect on the activation energy, lowering it by only 1 kcal mol⁻¹ for the *E* isomer, and *raising* it by about 1 kcal mol⁻¹ for the *Z*. The addition of a third solvent water is typically in the second solvation shell, thus not affecting the stabilization of the solvated malonaldehyde directly.

B. Mechanism. Yamabe and co-workers make the argument that the presence of a water tetramer in the solvated keto form

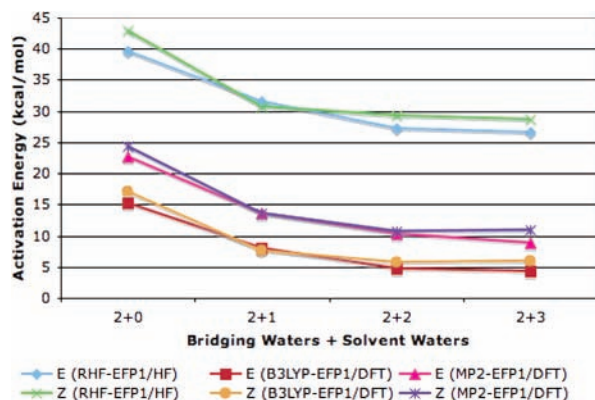


Figure 11. Activation energies as a function of the number of EFP solvent water molecules for the formation of the *E* and *Z* isomers with two bridging water molecules.

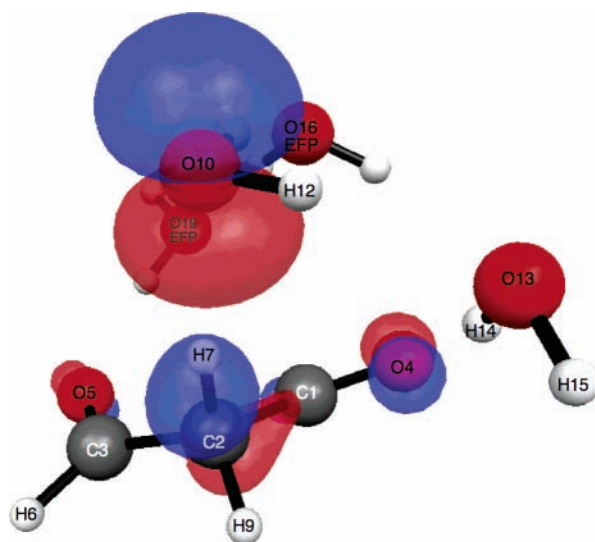


Figure 12. HOMO of the keto form at the MP2-EFP1/DFT/6-31G(d,p) level of theory. The position of the fragments leads to the formation of the *E* isomer.

makes the O10 oxygen (see, for example, Figures 6 and 7) highly nucleophilic. They point out that the HOMO of gas-phase water is a lone pair on oxygen with an energy of -0.50 au, and the HOMO of the keto form of malonaldehyde (with the water tetramer) is localized at this same lone pair on oxygen, but with a much higher energy of -0.42 au.¹

Focusing on the MP2-EFP1/DFT/6-31G(d,p) results, it is clear that the presence of the water tetramer is not necessary to increase the HOMO energy of the lone pair localized at O10. Figure 12 shows the HOMO form of the keto with two bridging waters and two EFP solvent waters leading to the *E* isomer of the enol. The energy of this orbital is also significantly higher in energy: -0.439 au as compared to -0.497 au for the free water at the same level of theory. Because of the increased nucleophilicity, the C2–H7 bond length is elongated to 1.109 Å over the unsolvated bond length of 1.092 Å. The LUMO+1 (0.003 au higher in energy than the LUMO) is delocalized but does have a noteworthy AO component associated with H7. The keto form leading to the *Z* isomer displays similar frontier orbital properties, having an energy of -0.439 au and a C2–H7 bond length of 1.110 Å. However, in this case the MO analogous to the HOMO shown in Figure 12 is the (HOMO–1). (The HOMO itself is only 0.007 au higher in energy.)

Even in the case of two bridging waters and only one fragment, the (HOMO–1) has similar atomic orbital contributions but has a slightly lower energy of -0.446 au. (By way of comparison, the full MP2/6-31G(d,p) method has the same (HOMO–1) at -0.448 au.) In this case, the C2–H7 bond length is 1.104 Å. Therefore, with increasing energy in the HOMO [or (HOMO–1)], corresponding to increased nucleophilicity, the C2–H7 bond length becomes more elongated, as expected. For the “cis” transition structure, Yamabe reports a C2–H7 bond length of 1.124 Å.

It is interesting to note that the molecular orbital shown in Figure 12 also has a significant AO component connecting the carbons where the double bond eventually forms. For the *E* isomer, this is between C1 and C2, and for the *Z* isomer, it is between C2 and C3. This, along with the unphysical nature of the hydronium ion “intermediate”, indicates that the tautomerization mechanism may be more concerted than previously suggested.

V. Conclusions

A mechanism for the keto–enol tautomerization through a competing “trans” transition state structure has been investigated using ab initio and the RHF- and DFT-based effective fragment potential methods. This “trans” structure leads to activation energies that are on par with, or slightly more favorable than, the previously studied “cis” structure, and the “trans” structure has the added flexibility of leading to either the *E* or *Z* isomer of the enol.

As with the general “cis” transition structure, the “trans” reaction paths are driven by the increased nucleophilicity of the HOMO (or HOMO–1) molecular orbitals of the keto form of malonaldehyde. However, it has been shown that the presence of a water tetramer is a sufficient, but not a necessary, condition for the observed increase in the frontier orbital energy.

Although there are qualitative differences in the potential energy surfaces between the B3LYP-EFP1/DFT/6-31G(d,p) method and the full B3LYP/6-31G(d,p) method, the average absolute relative deviation between the MP2-EFP1/DFT/6-31G(d,p) method and the full MP2/6-31G(d,p) calculations is 2.2 kcal mol⁻¹ and will be used to further investigate the energetics of the solvated Mannich reaction. The addition of two or three discrete EFP solvent waters is expected to recover the majority of the relaxation effects due to solvation.

Acknowledgment. This research was supported by an award from Research Corporation. Thanks to James E. Hougas III for early preliminary work related to this project and to Prof. Mark L. Kearly for introducing us to the Mannich reaction.

Supporting Information Available: Figures of the transition states analogous to Figures 6 and 7, and potential energy surfaces analogous to Figures 8 and 9 for additional solvent waters. A figure of the (HOMO–1) of the keto form at the MP2-EFP1/DFT/6-31G(d,p) level of theory, and a table of the all the imaginary frequencies for the transition states discussed in the text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Yamabe, S.; Tsuchida, N.; Miyajima, K. *J. Phys. Chem. A* **2004**, *108*, 2750.
- (2) See, for example: Tomaoyuki, H.; Mukamel, S. *J. Phys. Chem. A* **2003**, *107*, 9113. Sobolewski, A. L.; Domcke, H. B. *J. Phys. Chem. A* **1999**, *103*, 4494. Perrin, C. L.; Kim, Y. *J. Am. Chem. Soc. A* **1998**, *120*, 12641. Chiavassa, T.; Roubin, P.; Pizzala, L.; Verlaque, P.; Allouche, A.; Marinelli,

- F. J. Phys. Chem.* **1992**, *96*, 10659. Latajka, Z.; Scheiner, S. *J. Phys. Chem.* **1992**, *96*, 9764. Hutchinson, J. S. *J. Phys. Chem.* **1987**, *91*, 4495. Frisch, M. J.; Scheiner, A. C.; Schaefer, H. F., III; Binkley, J. S. *J. Chem. Phys.* **1985**, *82*, 4194.
- (3) Chiavassa, T.; Verlaque, P.; Pizzala, L.; Allouche, A.; Roubin, P. *J. Phys. Chem.* **1993**, *97*, 5917.
- (4) Adamovic, I.; Freitag, M. A.; Gordon, M. S. *J. Chem. Phys.* **2003**, *118*, 6725.
- (5) Adamovic, I.; Gordon, M. S. *J. Phys. Chem. A* **2005**, *109*, 1629.
- (6) Day, P. N.; Pachter, R.; Gordon, M. S.; Merrill, G. N. *J. Chem. Phys.* **2000**, *112*, 2063.
- (7) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293.
- (8) See, for example, the following reviews: Heaney, H. The Bimolecular Aromatic Mannich Reactions. *Comprehensive Organic Synthesis*; Trost, B. M., Hrsg.; Pergamon Press: Oxford, U.K., 1993; Vol. 2, Part 2, Chapter 4.2. Tramontini, M. *Synthesis* **1973**, 703. Thompson, B.B. *J. Pharm. Sci.* **1969**, *57* (5), 715.
- (9) Mannich, C.; Krosche, W. *Arch. Pharm.* **1912**, *250*, 647.
- (10) J. D. Tuma; M. L. Kearley; G. M. Thiele; S. Worrall; L. W. Klaassen; M. F. Sorrell; Elucidation of Reaction Scheme Describing Malondialdehyde – Acetaldehyde – Protein Adduct Formation. *Chem. Res. Toxicol.* **2001**, *14* (7), 822–832.
- (11) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257. Pople, J. A. *J. Chem. Phys.* **1975**, *62*, 2921.
- (12) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098; *J. Chem. Phys.* **1993**, *98*, 1372, 5648.
- (13) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
- (14) Fukui, K. In *The World of Quantum Chemistry*; Daudel, R., Pullman, B., Eds.; Reidel: Dordrecht, The Netherlands, 1974; p 113. Fukui, K.; Kato, S.; Fujimoto, H. *J. Am. Chem. Soc.* **1975**, *97*, 1. Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 36. Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523. Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1990**, *95*, 5853.
- (15) Bode, B. M.; Gordon, M. S. *J. Mol. Graph. Modeling* **1998**, *16*, 133.
- (16) See, for example: Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons: Chichester, U.K., 1999; p 111.