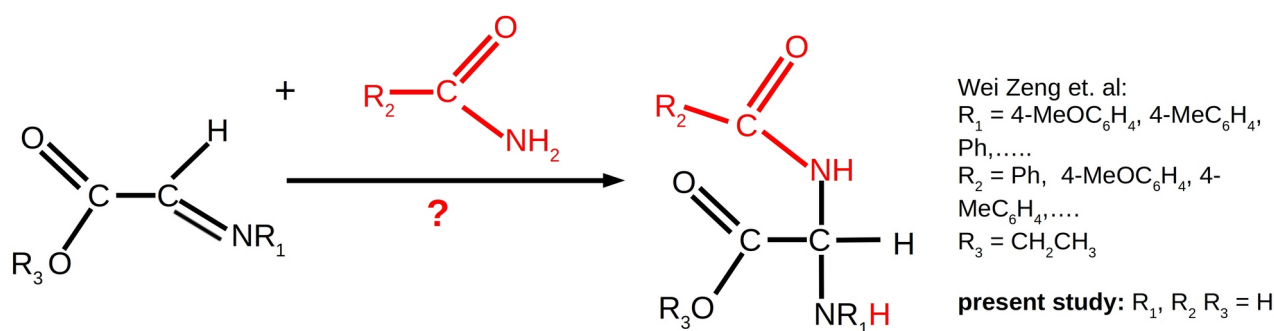# Reaction Mechanism of 2-Amido-2-Aminoacetic Acid-Formation from Amides and 2-Iminoacetic Acid: A DFT-study and Wavefunction Analysis

Georg Dazinger

Abstract

Based on a study of Wei Zeng et. al.[7], where the synthesis of *gem*-diamino acid esters from 2-iminoacetic acid esters and amides, with various N- and C-substituents, respectively, is reported, a modeled reaction, where the latter substituents were replaced by H, was simulated by means of DFT. A reasonable reaction mechanism was found for the formation of 2-amido-2-aminoacetic acid from formamide and 2-iminoacetic acid. Moreover, possible side reactions were simulated and discussed.

Wei Zeng et. al:
R_1 = 4-MeOC$_6$H$_4$, 4-MeC$_6$H$_4$, Ph,.....
R_2 = Ph, 4-MeOC$_6$H$_4$, 4-MeC$_6$H$_4$,....
R_3 = CH$_2$CH$_3$

present study: R_1, R_2 R_3 = H

*picture 1: the investigated reaction*

1. Introduction

The reaction of amides with imines is a well known method in order to synthesize *gem*-diamines and its derivatives [1, 2]. *gem*-diamino carboxylic acid derivatives play an important role as biologically active molecules [3, 4], for example, as inhibitors of HIV-1 protease [5] and E. coli [6]. Wei Zeng et. al. [7] synthesized gem-diamino acid esters from 2-iminoacetic acid esters and amides, with various N- and C-substituents, respectively. As amide substituents R_2, phenyl, 4-methylphenyl and others were used, the iminoacetic acid esters were N-substituted with phenyl, 4-methoxyphenyl and others (substituents R_1) as R_3, ethyl was used. Wei Zeng et. al. used Cu(II) and various phosphanes as catalyst, resulting in "good to excellent" [7] yields. However, the reaction mechanism has hitherto remained unclear.

The present study provides insights into this reaction mechanism. Interestingly, Wei Zeng et. al. report that *gem*-diamino acid esters were also yielded "in traces" [7], if no Cu(II)-catalyst was used. For the latter reaction (without catalyst), several mechanistic pathways were simulated by theoretical means. It turned out that the investigation of the non-catalyzed synthesis of *gem*-diamino acetic esters also provides information for the catalyzed reaction. Obviously, one or more side reactions draw responsible for the low yield of the desired product. In the present study, a possible mechanism for one of these competing reactions is proposed in order to elucidate inhibiting and competing factors. All substituents R_x (with x=1,

2 or 3), were modeled as H. The replacement of larger substituents by H or CH_3 is commonly done in order to reduce computational requirements, as for the calculation of enzymes [8]

For R_1, R_2 and R_3 set as H, picture 2 gives a schematic overview over the calculated reaction pathways. A reaction mechanism, where formamide directly adds to 2-iminoacetic acid, was calculated. This reaction mechanism proved to be highly unlikely due to a high kinetic barriers. This is why it was searched for another reaction pathway. Indeed, an alternative pathway was found, which is considerably more probable.

Amide-imidic acid- tautomerism has been investigated in a wide variety of contexts [9, 10, 11, 12], including the author's own work [13], as well as experimentally [14, 15]. In the alternative pathway, amide undergoes a tautomerism reaction, which is catalyzed by 2-iminoacetic acid. The resulting imidic acid then adds to the former, resulting in 2-amido-2-aminoacetic acid, where the amide- and amine moieties are in *gem*-position. A possible side-reaction was also elucidated. It was found that imidic acid may attack on the a-C, as shown in picture 2, resulting in N-(1,1- dihodroxy-2-iminoethyl)amide.
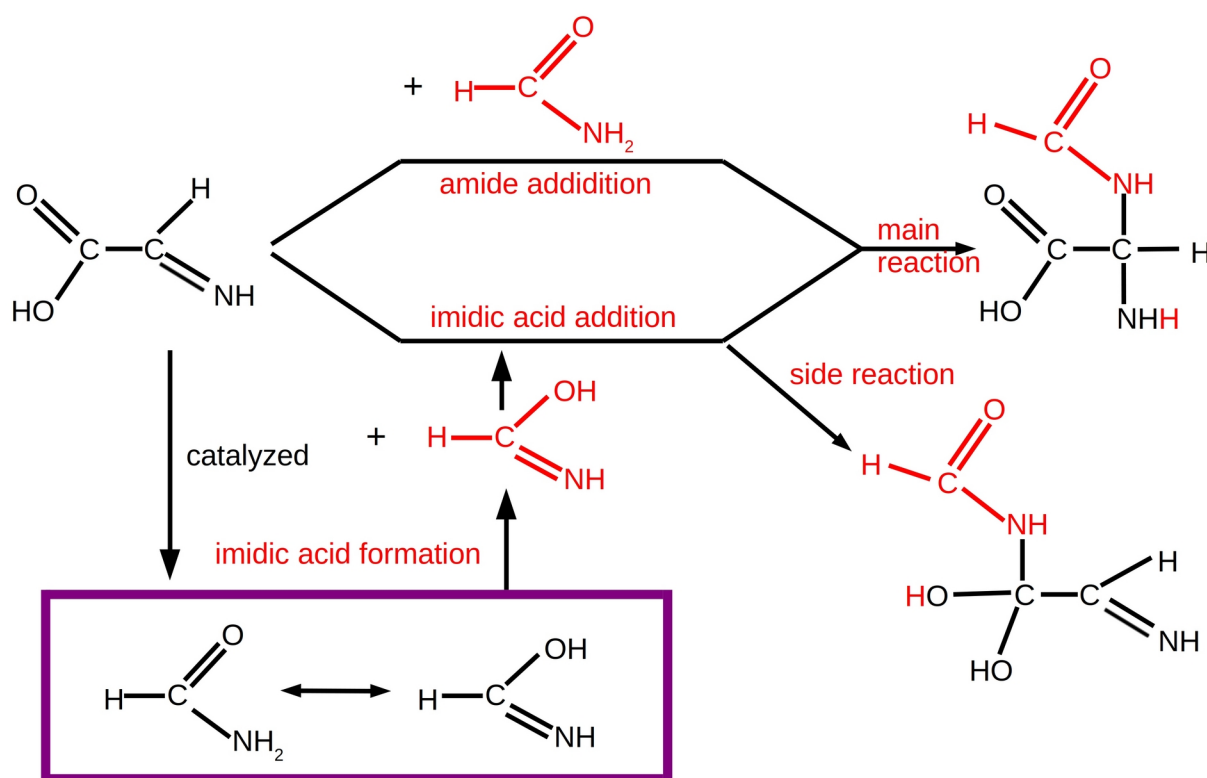
picture 2: schematic representation of the main reaction pathways and the side reaction

The results provide explanations why the yield of the main product 2-amido-2-aminoacetic acid and its substituted derivatives is low, and why a transition metal catalyst is needed.

1.1. Methods and Theory

All molecules were modeled by means of DFT, applying the range-separated hybrid-meta-GGA-functional M11 designed by Truhlar et. al, which has been proven to perform well for DFT calculations [16, 17], as well as for TDDFT [18]. For all molecular states in chapter 2.1., where the reaction mechanisms are discussed, 6-31G(d,p) was used as basis set [19], polarization functions (d,p) were added [20].

The triple-zeta basis 6-311++G(d,p) [21], including diffuse functions [22], was applied for certain decisive molecular states. These states were fully re-optimized with the latter basis and used for wavefunction analysis in chapter 2.2. Moreover, the triple-zeta optimized states were compared to their respective double-zeta analogues in order to check the reliability of 6-31G(d,p). The basis sets were partly retrieved from EMSL basis set exchange homepage [23, 24, 25].

All molecular states were geometry optimized and frequency analyzed, obtaining no negative eigenvalue of the resulting Hessian matrix for ground states and one negative eigenvalue for transition states. Gibbs free energies were calculated by adding entropy correction terms obtained from frequency analysis to the respective molecular electronic energy.

All ground states were abbreviated as **GS** and distinguished by capital letters, for example: **GS_A**, **GS_B**, and so on. Transition states were abbreviated as **TS**, followed by two capital letters, which are assigned to their respective ground states. For example, the transition state between **GS_A** and **GS_B** was named **TS_AB**. Single atoms were named by their element symbol and their numbering within the respective molecule. For example, nitrogen with the number 5 was named as N5, oxygen with number 2 was named as O2 (not to be confused with O_2).

Transition states were verified as connecting links between their respective ground states by IRC calculation [26].

Dimer coordination energies were counterpoise corrected in order to diminish basis set superposition error [27, 28].

Optimization and frequency calculations were carried out with GAMESS-US [29, 30].

In order to obtain chemically intuitive information, wavefunction analysis, including Mayer bond order, ADC-Hirshfeld charges, electrostatic potential surface and Ruedenberg orbitals, was performed for certain decisive molecular states. All wavefunction-analyzed states were calculated and optimized with 6-311+ +G(d,p). Atomic dipole-corrected Hirshfeld charges [31] were calculated from GAMESS-output. Atomic dipole corrected (ADC) Hirshfeld charges are taking the atomic dipole moments into account, using the assumption that the sum of all atomic dipole moments equal the molecular dipole moment. The resulting correction term is added to the conventional Hirshfeld charges [32].

Mayer bond orders [33, 34] were calculated from the latter geometry-optimized states in order to gain further information about their molecular structure. Moreover, electrostatic potential surfaces [35] were calculated in order to elucidate nucleophility and electrophility of the molecules analyzed. Ruedenberg orbitals [36] were directly calculated with GAMESS-US. These localized orbitals provide better insights into intramolecular electronic interaction. Ruedenberg orbitals were visualized with Jmol [37].

ADC Hirshfeld charges, Mayer bond orders, charge-transfer-analysis and electrostatic potential surfaces (EPS) were calculated with MULTIWFN [38], which is a highly recommendable tool for a deeper understanding of the electronic structure of molecules. Electrostatic potential surfaces, calculated using

the latter, were visualized with VMD[39]. Molecular geometries were visualized with chemcraft ® [40]. Chapter 1.1. of this study and the acknowledgments were largely adopted from Lit. [13].

2. Results and Discussion

2.1.1. Overview

Four reaction mechanisms were simulated: H-rotation, direct amide addition, imidic acid formation and addition, and finally a side reaction. They are connected as illustrated picture 3. Outgoing from 2-iminoacetic acid and formamide, loosely connected to each other (**GS_A**), a H-rotation takes place, leading to **GS_B**. Continuing from **GS_B**, there are two different pathways for the formation of 2-amido-2-aminoacetic acid **GS_F**: the first pathway describes addition of amide, the second one is the formation and addition of imidic acid. These two pathways describe the main reaction, since **GS_F** is referred to as the desired product in lit. [7].
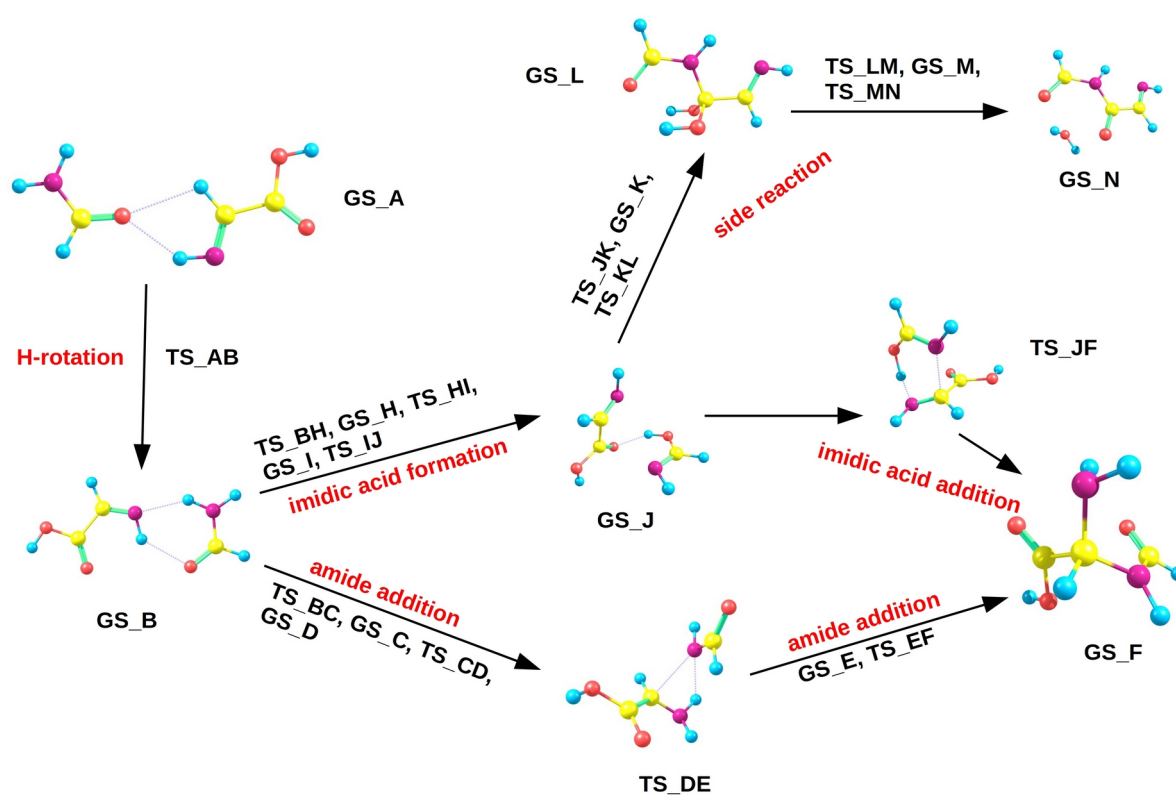
picture 3: outline of all reaction pathways

From **GS_J**, which is an intermediate of the imidic acid addition-pathway, the side reaction bifurcates, leading to N-(1,1- dihodroxy-2-iminoethyl)amide **GS_L**, and, via elimination of H_2O, to N-formyl-2-iminoethanamide and H_2O, **GS_N**.

4
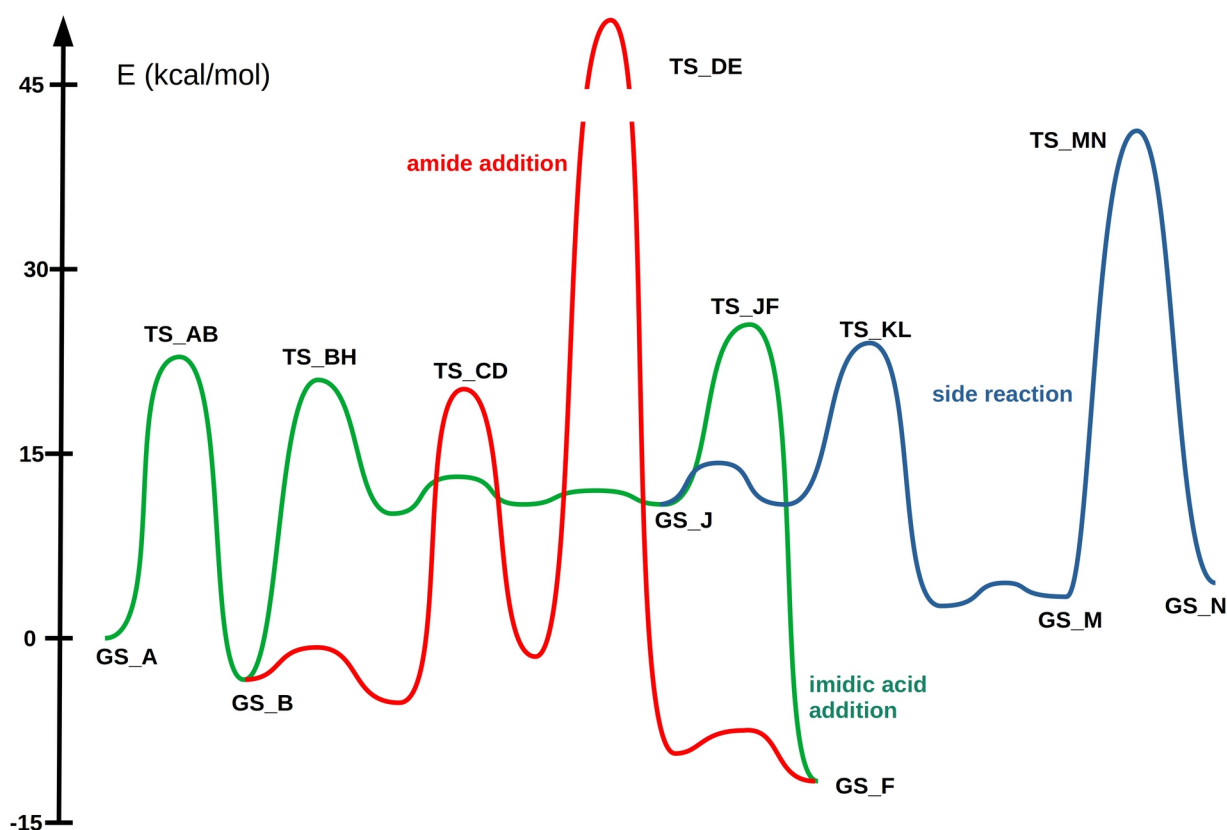
*picture 4: comparative energy profile of all reaction pathways. H-rotation is assigned to imidic acid addition. The energy barrier of **TS_DE** is out of scale.*

An energy comparison (Gibbs free energies, basis set: 6-31G(d,p)) of all reaction pathways is shown in picture 4. Noteworthy, **TS_KL**, which is the rate-determining step for the side reaction, is slightly lower than **TS_JF**. All reaction mechanisms are discussed in detail in the following chapters.

2.1.2. H-Rotation

As a starting point for all reaction pathways, **GS_A** was chosen. Free energies of all the other states are related to **GS_A**, which is set as 0,0 kcal/mol. In **GS_A**, formamide is connected with iminoacetic acid via hydrogen bonds. The bond lengths can be seen in picture 5. In order to enable amide-imidic acid tautomerism, **GS_A** has to be transformed into **GS_B** by hydrogen rotation. In transition state **TS_AB** (+22,35 kcal/mol), hydrogen H10 rotates upwards along the amide- O2, while the iminoacetic acid rotates downwards along H10 (picture 5).
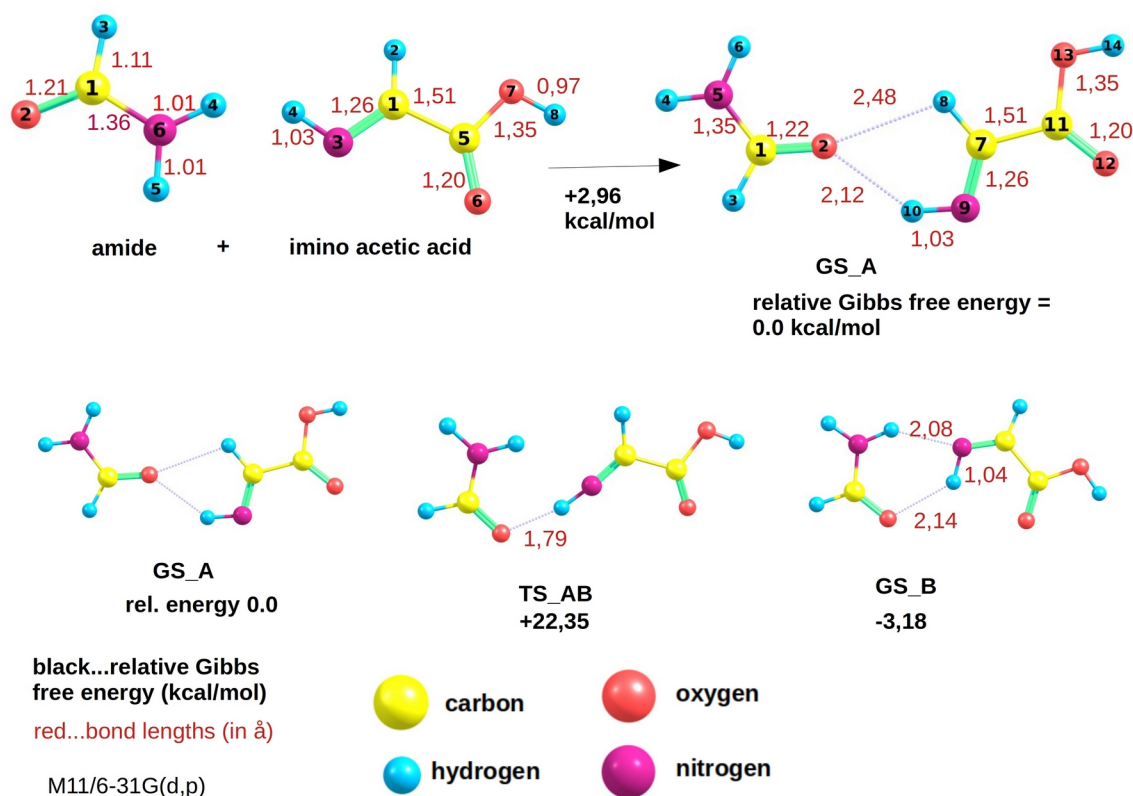
picture 5: monomer and dimer bond lengths (in angstroms), dimer formation energy, H-rotation

In the resulting **GS_B**, the imine moiety is switched from *E*- to *Z*- configuration, which is necessary for the following tautomerism reaction within the imidic acid addition-reaction pathway. The direct amide addition also has a connection to **GS_B**. Thus, both main reaction mechanisms are directly comparable. The structure of the DFT/M11/6-31-G(d,p) modeled amide is taken from lit. [13].
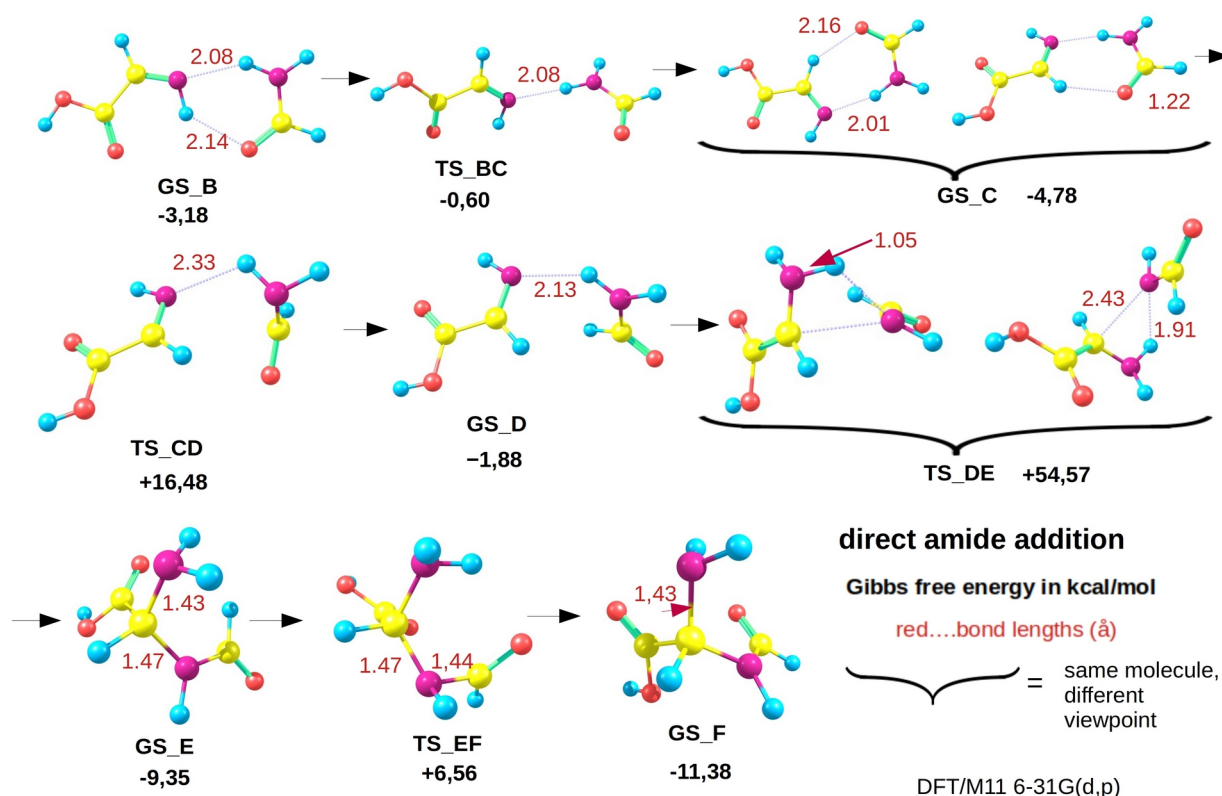
2.1.3. Direct Amide Addition

Outgoing from **GS_B**, the amide undergoes a rotation around the N-H10-N-axis in the transition state **TS_BC,** which is 2,58 kcal/mol higher in energy than **GS_B** (the atom numbering is always the same is in **GS_A**). The dihedral angle between 2-iminoacetic acid and formamide in **TS_BC** is 98,9° (measured as the dihedral angle N5-H4-N9-H10, where N5 is the amide-nitrogen, H4 is the amide hydrogen which is bound to the imine-nitrogen N9, H10 is the hydrogen of the imino group). This rotation results in **GS_C**, which exhibits an almost planar structure. The O2-H8 bond length is 2,16 angstroms, see picture 6. From **GS_B** to **GS_C**, the amide undergoes an approximate 180-degree rotation. In **TS_CD**, the O2-H8 bond is broken, since the CHO-moiety of the amide (atoms C1, O2 and H3) rotates along the amide-N-C-axis (atoms N5 and C1). This transition state exhibits a relatively high energy barrier of 21,26 kcal/mol relative to **GS_C**, since it describes an intra-molecular rotation. In the resulting **GS_D**, formamide shows only one hydrogen bond with iminoacetic acid. Its length is 2,13 angstrom, see picture 6. **GS_D** serves as starting point for the actual amide-addition transition state **TS_DE**. It is 54,57 kcal/mol higher in energy than **GS_A** and 57,75 kcal/mol higher than **GS_B**. It is thus the rate-determining step.

The resulting product **GS_E** shows lower free energy than **GS_A** and **GS_B**, respectively. This is in accordance with the experimental results of Zeng et. al [7].

Nevertheless, it seemed doubtful that a reaction mechanism which passes an energy barrier of over 50 kcal/mol is thermodynamically possible. Further electronic properties of **TS_DE** will be discussed in chapters 2.2.2 and 2.2.3.

For comparison reasons, the transformation from **GS_E** to **GS_F** was also calculated, since **GS_F** is the product of the reaction mechanism based on imidic acid, which will be discussed in the following chapter. **GS_F** is formed by rotation of the CHO-moiety (atoms C1, O2 and H3), which requires 15,91 kcal/mol (**TS_EF**). This considerably high energy can be explained by the C1-N5-bond length of 1,44 angstroms, which indicates a bond order higher than 1.

*picture 6: reaction mechanism for the direct amide addition*

Since **GS_B** and **GS_F** serve also as the educt and product states, respectively, of the alternative reaction mechanism, which will be discussed in the following, the both mechanisms are thermodynamically connected to each other and thus directly comparable.

2.1.4. Imidic Acid Addition

The alternative reaction mechanism, which involves imidic acid formation and addition, also starts with **GS_B**. By undergoing **TS_BH**, imidic acid is formed. 2-iminoacetic acid hereby serves as catalyst. Its imino group provides one hydrogen for migration to the amide-oxygen, while the amide-nitrogen donates one hydrogen to iminoacetic acid (picture 7). The latter switches thus back from *Z*- to *E*-configuration in **GS_H**. Remarkably, this tautomerism reaction bears a lower energy barrier (20,89 kcal/mol relative to

GS_A, 24,07 kcal/mol relative to **GS_B**) than the autocatalyzed formamide tautomerism, as investigated in lit [13]. On the other hand, the iminoacetic acid-catalyzed tautomerism is only possible if the imine-N is H-substituted. For other substituents, the tautomerism is supposed to be autocatalyzed as described in the latter literature, requiring an energy of 28,6 kcal/mol for overcoming the energy barrier (lit.[13]). On the other hand, it may be considered that the amide-imidic acid tautomerism requires lower energy, if a transition metal catalyst is present, as described in Wei Zeng's experimental study [7]. In the latter study, a Cu(II)-catalyst was used for the reaction of N - unsubstituted amides with substituted iminoacetic acid esters. An experimental investigation, using unsubstituted iminoacetic acid esters, would thus be elucidating.

Nevertheless, the following reaction steps may also provide information concerning Wei Zeng's experimental results [7], because the substituents used for the latter study are not directly involved in the reaction mechanism described here.
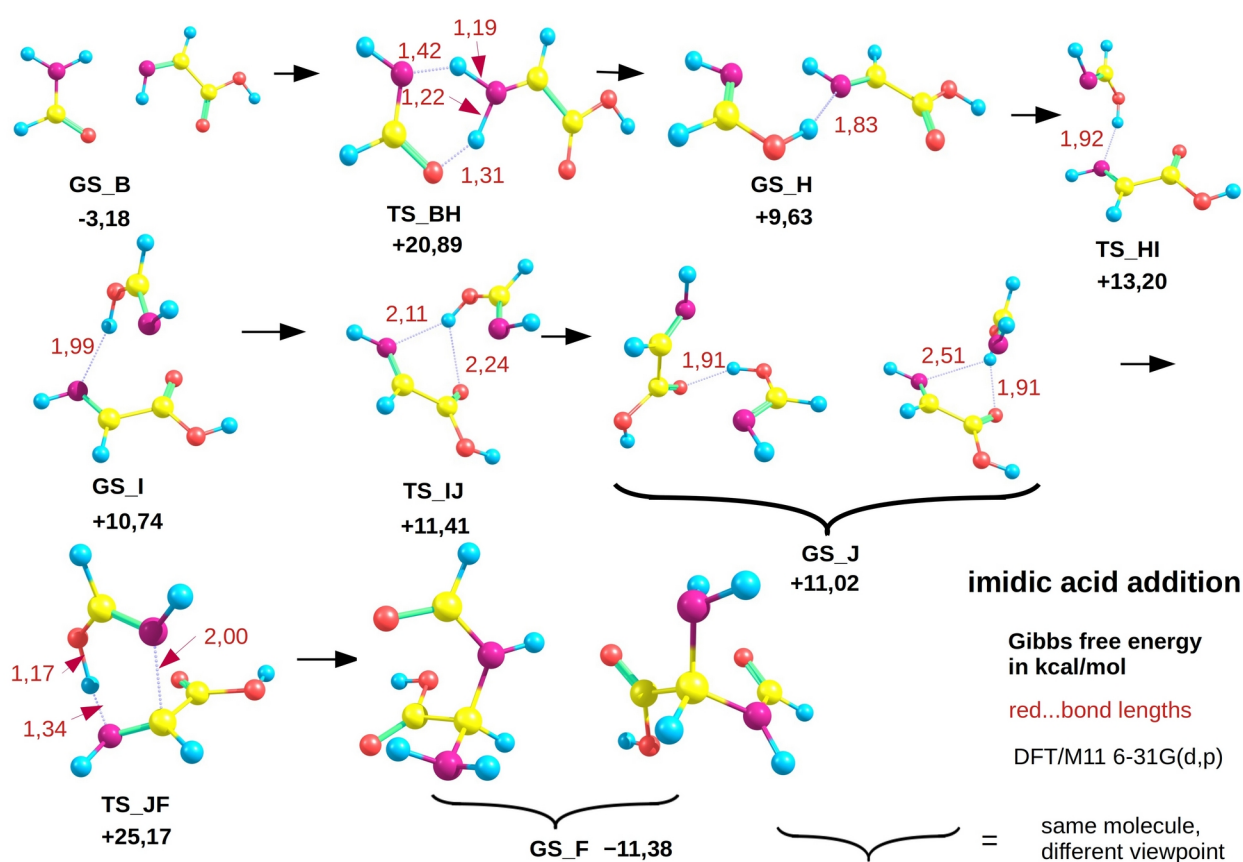
picture 7: reaction mechanism of imidic acid formation and addition

The planar geometry of **GS_H** is broken in **TS_HI**, which produces an energy barrier of 3,57 kcal/mol. The dihedral angle between imidic acid and iminoacetic acid is -79,2°, (measured as the angle between atoms C7-N9-O2-C1), and is relaxing to -39,2° in **GS_I**. The latter ground state state bears a N-H-O hydrogen bond of 1,99 angstroms length, see picture 7. For the actual addition step, the imidic acid has to undergo a further slight rearrangement step, which is described as **TS_IJ**. In this transition state, the imidic acid moves from the imine group of iminoacetic acid to the carboxyl-oxygen O12, thus elongating the N-H-O bond from 1,99 to 2,11 angstroms. A further hydrogen bond to O12 is formed with the length of 2,24 angstroms. The energy barrier of 0,67 kcal/mol is comparably small. The resulting **GS_J** shows

8

respective hydrogen bonds of 2,51 and 1,91 angstroms length. From **GS_J**, imidic acid undergoes the actual addition step via **TS_JF**. This transition state combines electrophilic hydrogen (H10; the imidic acid hydrogen bound to O2) migration and nucleophilic nitrogen addition (N5; imidic acid nitrogen), see picture 7. Atom H10 migrates to N9, the N9-H10 bond length in **TS_JF** is 1,34 angstrom, the H10-O2 bond length is 1,17 angstroms. N5 migrates to C7, the bond length is 2,00 angstroms. This transition state directly leads to the product 2-amido-2-aminoacetic acid **GS_F**, which is described in the former chapter. Bearing the largest energy barrier, **TS_JF** is the rate determining step. It is nevertheless considerably lower than the rate determining step of the amide addition, which is +54,57 kcal/mol. The imidic acid addition was therefore considered as favored reaction mechanism.

Indeed, Wei Zeng et. al. [7] reported that this reaction (using 4-methoxyphenyl as substituent for the iminoacetic ethyl-ester-N and 4-chlorophenyl as substituent for the amide-C) yielded the product "in traces", even when no Cu(II)-catalyst was present lit. This gave support to the assumption that the proposed reaction mechanism via imidic acid addition is thermodynamically possible. On the other hand, the role of the catalyst still has to be discussed, since it causes higher yields of the desired product. It was supposed that a side reaction competes with the imidic acid addition. In fact, a side reaction, which is kinetically slightly favored, was discovered. It moreover explains a decisive function of the catalyst.
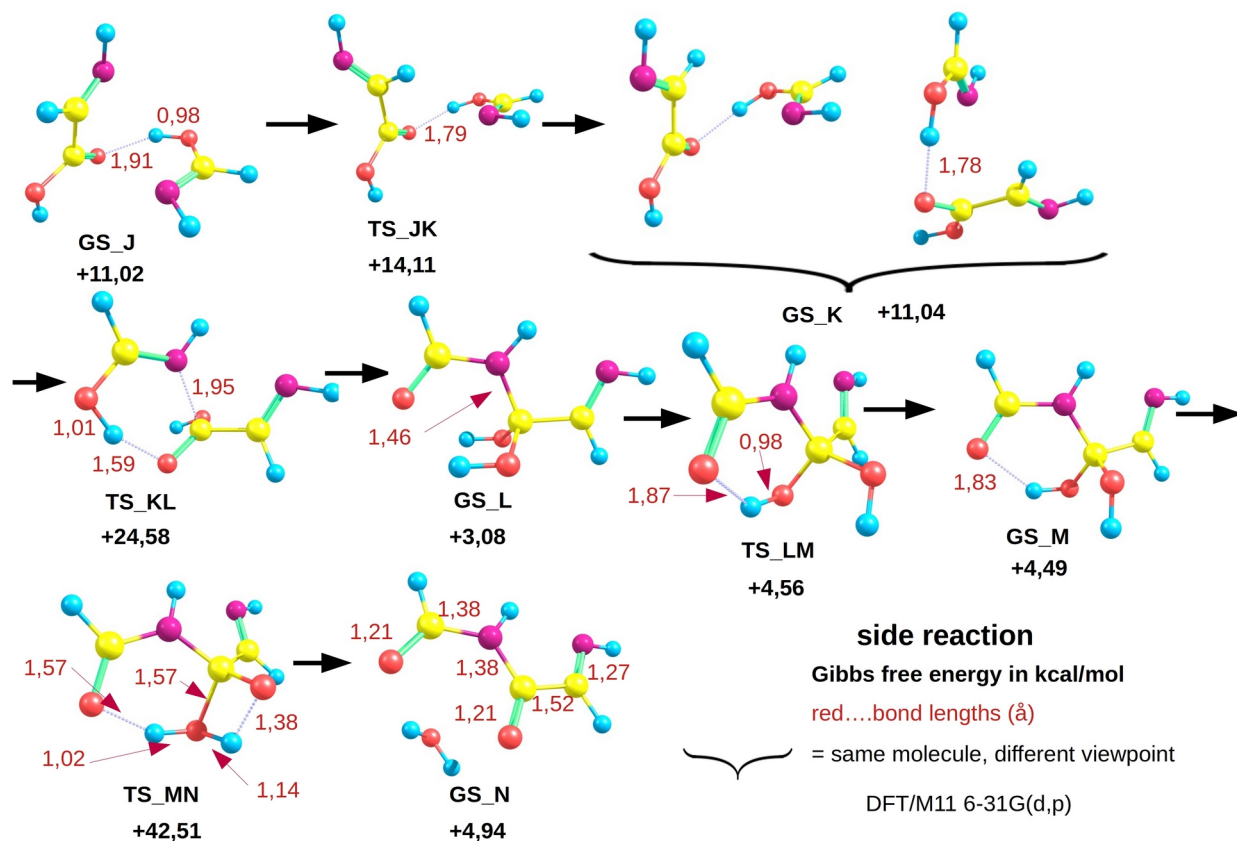
2.1.5. Side Reaction

A possible side reaction was found to be outgoing from **GS_J**, which is an intermediate state of the imidic acid addition (picture 8). In **TS_JK**, iminoacetic acid undergoes an intramolecular rotation of its imine moiety around the C7-C11- axis. The dihedral angle O12-C11-C7-N9 changes from 10,906° in **GS_J** to 24,899° in **TS_JK**. The energy barrier from **GS_J** to **TS_JK** is 3,09 kcal/mol. In the resulting ground state **GS_K**, the respective dihedral angle relaxes to 170,742°. The latter ground state is a starting point for a concerted addition reaction of imidic acid nitrogen to the α-C of iminoacetic acid, while the imidic acid H10 migrates to carboxyl-oxygen O12, undergoing transition state **TS_KL**. The side reaction thus differs from the main reaction, insofar as in the latter, the imidic acid-N adds to the β-C of iminoacetic acid and H10 migrates to N9. Noteworthy, **TS_KL** is lower in energy (+24,58 kcal/mol relative to **GS_A**) than **TS_JF**, which is +25,17 kcal/mol in free energy. Since these two steps are rate-determining for their respective reactions, it can be concluded that the side reaction is kinetically favored over imidic acid addition. On the other hand, the energy difference between both rate-determining steps is rather small (-0,59 kcal/mol), which may explain the result described in the study of Wei Zeng et. al, lit[7], that gem-diamino acid derivatives are still found "in traces".

On the other hand, the role of a transition metal catalyst is also elucidated by **TS_KL**: Iminoacetic acid (or the respective esters) are supposed to coordinate to a transition metal center via the carboxyl-atom O12 (see discussion in chapter 2.2.4). A metal-coordinated oxygen O12 should have considerably less electron-donating properties than in free state. A migration from H10 to O12, as described in **TS_KL**, would thus be thermodynamically impossible, since the metal center and H10 would have to compete for the oxygen electrons. Consequently, it can be stated that a transition metal catalyst inhibits the side reaction. **TS_KL** results in N-(1,1- dihodroxy-2-iminoethyl)amide **GS_L**, which may be considered as stable intermediate, since it is considerably lower in energy than **GS_J**, although less stable than **GS_F**. This indicates that, at high temperatures, where an energy barrier of 25,17 kcal/mol (as described by **TS_JF**) can be overcome, **GS_F** appears to be the favored product. However, this is in contradiction to the experimental results [7].

This is why it was searched for another final product of the side reaction with lower energy. A reaction mechanism, where one H_2O is eliminated from **GS_L**, was thus simulated. In **TS_LM**, a hydrogen bond between one OH-moiety and the amide-O (O2) is to be formed, 1,87 angstroms in length. It can be seen that the geminal OH-group undergoes a rotation. In **GS_M**, the rotation is accomplished and the hydrogen

bond relaxes to 1,83 angstrom. In **TS_MN**, H_2O is eliminated, requiring a considerably high energy of 42,51 kcal/mol. Unfortunately, the resulting product **GS_N** is still 16,32 kcal/mol higher in energy than **GS_F**. This indicates that a side reaction based only on one formamide and one iminoacetic acid does not explain the experimental results fully. It may thus be discussed if a second amide or a second iminoacetic acid is involved in the side-reaction mechanism. This possibility will be discussed in chapter 2.2.4.

picture 8: mechanism of a possible side reaction

2.2. Wave Function Analysis

In the following, several electronic aspects of certain reaction steps presented in this study are discussed. The reliability of the basis set 6-31G(d,p) is compared to 6-311++G(d,p) in example of the most important states. Moreover, electronic differences between the rate determining transition states **TS_JF** and **TS_DE** were elucidated. Chemical properties of the most important states were discussed based on EPS-analysis and Ruedenberg orbitals.

2.2.1. Reliability of the 6-31G(d,p) Basis Set

The reliability of the 6-31G(d,p) basis set was checked by comparison of several molecular states, which were calculated with the triple zeta basis 6-311++G(d,p) with additional diffuse functions. Imidic acid, **GS_A**, **GS_L**, **TS_KL**, **GS_N**, **TS_DE**, **TS_JF** and **GS_F** were fully re-optimized with the stronger triple-zeta basis. Their bond lengths an free energies can be seen in picture 9.
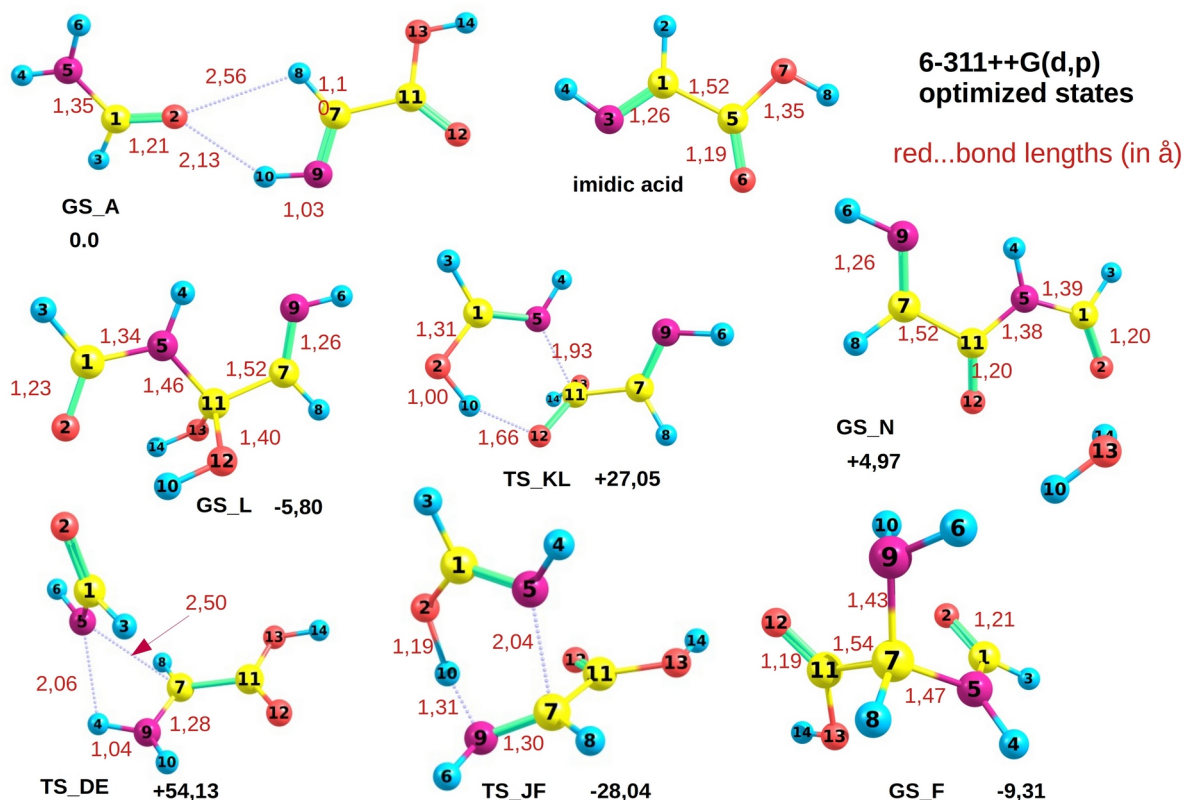
picture 9: important molecular states , re-optimized with DFT/M11, 6-311++G(d,p)

All free energies were again set in relation to **GS_A**, which was re-calculated with the triple zeta-basis. As can be seen, the energetic differences for **TS_DE** and **GS_N**, calculated with 6-31G(d,p) and 6-311++G(d,p), respectively, are relatively small. **TS_DE**, modeled with 6-311++G(d,p) is 0,44 kcal/mol lower than the same molecule calculated with 6-31G (relative to GS_A, calculated with the respective basis sets), for **GS_N**, the energy difference is 0,03 kcal/mol.

TS_JF and **GS_F** show larger deviations when calculated with those basis sets, since 6-311++G(d,p) is more capable to describe weak interactions, but they are still within an acceptable range. **TS_JF**, calculated with the triple-zeta basis set, is 2,87 kcal higher in relative energy than the same state modeled with 6-31-G(d,p), **GS_F** is 2,07 kcal/mol higher. However, **TS_JF** exhibits, for both basis sets, a higher energy than the competing **TS_KL**. It can thus be stated that 6-31G(d,p) describes the reaction kinetics sufficiently for a qualitative analysis.

TS_DE, calculated with 6-311++G(d,p) shows bigger lengths for the N5-H4 bond (2,06 angstroms) and the N5-C7 bond (2,50 angstrom) in comparison to **TS_JF** calculated with 6-31G(d,p), where the respective distances are 1,91 and 2,43 angstroms.

On the other hand, bond length deviations are smaller for **GS_F**, where the N9-C7 bond length is 1,43 angstroms for both basis sets. Also in **GS_N**, bond lengths of covalent interactions differ within in the range of 0,01 angstroms in average.

## 2.2.2. Ruedenberg Orbital Analysis of TS_DE and TS_JF

The two rate-determining steps for the main reaction, resulting in **GS_F,** are **TS_JF** for the imidic acid addition and **TS_DE** for the amide addition. Both transition states were analyzed and compared in order to gain deeper insights into their electronic structures and to find explanations, why **TS_JF** is 26,09 kcal/mol lower in energy than **TS_DE**. Ruedenberg localized molecular orbitals were calculated for both transition states, employing DFT/M11 and 6-311++G(d,p) basis set. In the following, Ruedenberg localized molecular orbitals are abbreviated as RO. RO's of **TS_JF** are abbreviated as ROa with subsequent number, RO's of **TS_DE** are abbreviated as ROb. The advantage of Ruedenberg molecular orbitals lies in their better interpretability, since respective RO's of similar molecules (as **TS_DE** and **TS_JF**) can directly be assigned to each other, and thus be compared.
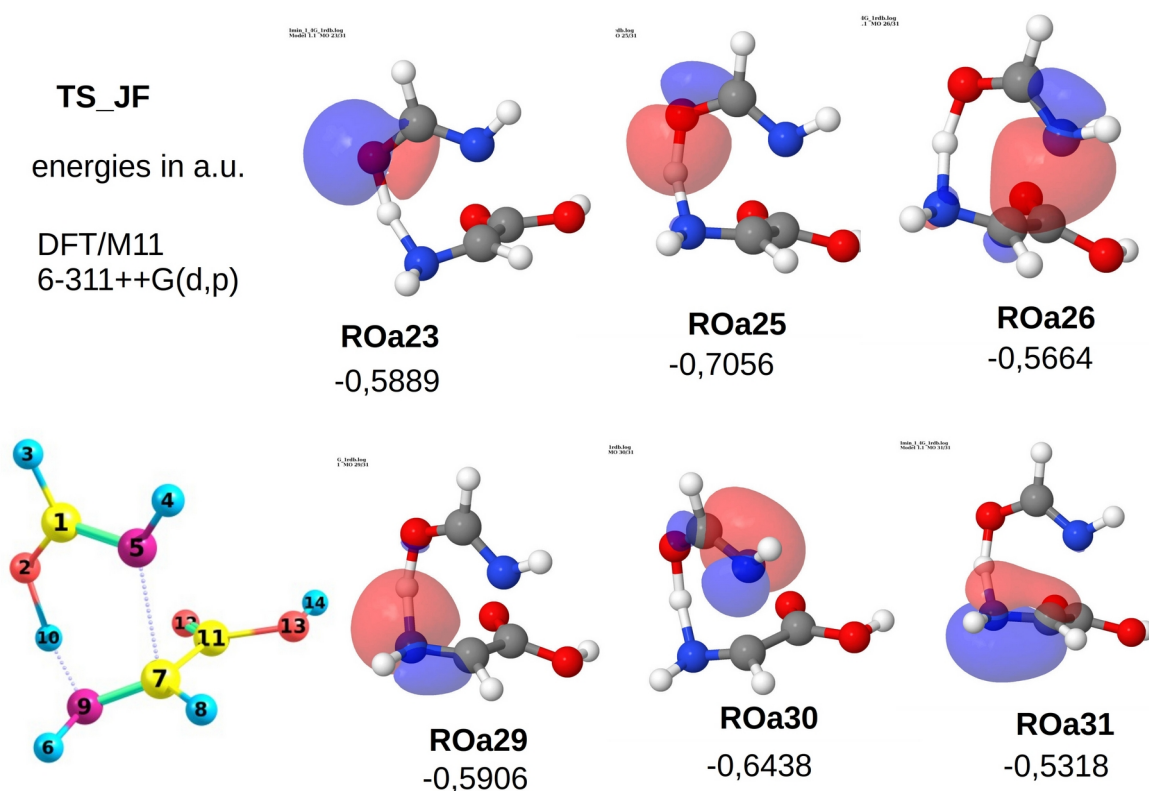
*picture 10: Ruedenberg orbitals of **TS_JF**, on the left: atom numbering of **TS_JF***

Although RO energies are not defined as clearly as MO energies, diagonal elements of Fock matrices, transformed according to Edminston-Ruedenberg algorithm, may serve as energy comparison points. They are therefore interpreted as RO-energies in the following.

It can be seen that ROa31 of **TS_JF** corresponds to ROb23 of **TS_DE**; see picture 10 and picture 11. Both RO's are π-bonds of the imine moiety (atoms C7 and N9) of iminoacetic acid. ROa31 is higher in energy, obviously due to repulsive interactions with the formic imidic acid nitrogen (atom N5). Their energies are -0,532 a.u. and -0,601 a.u., respectively.
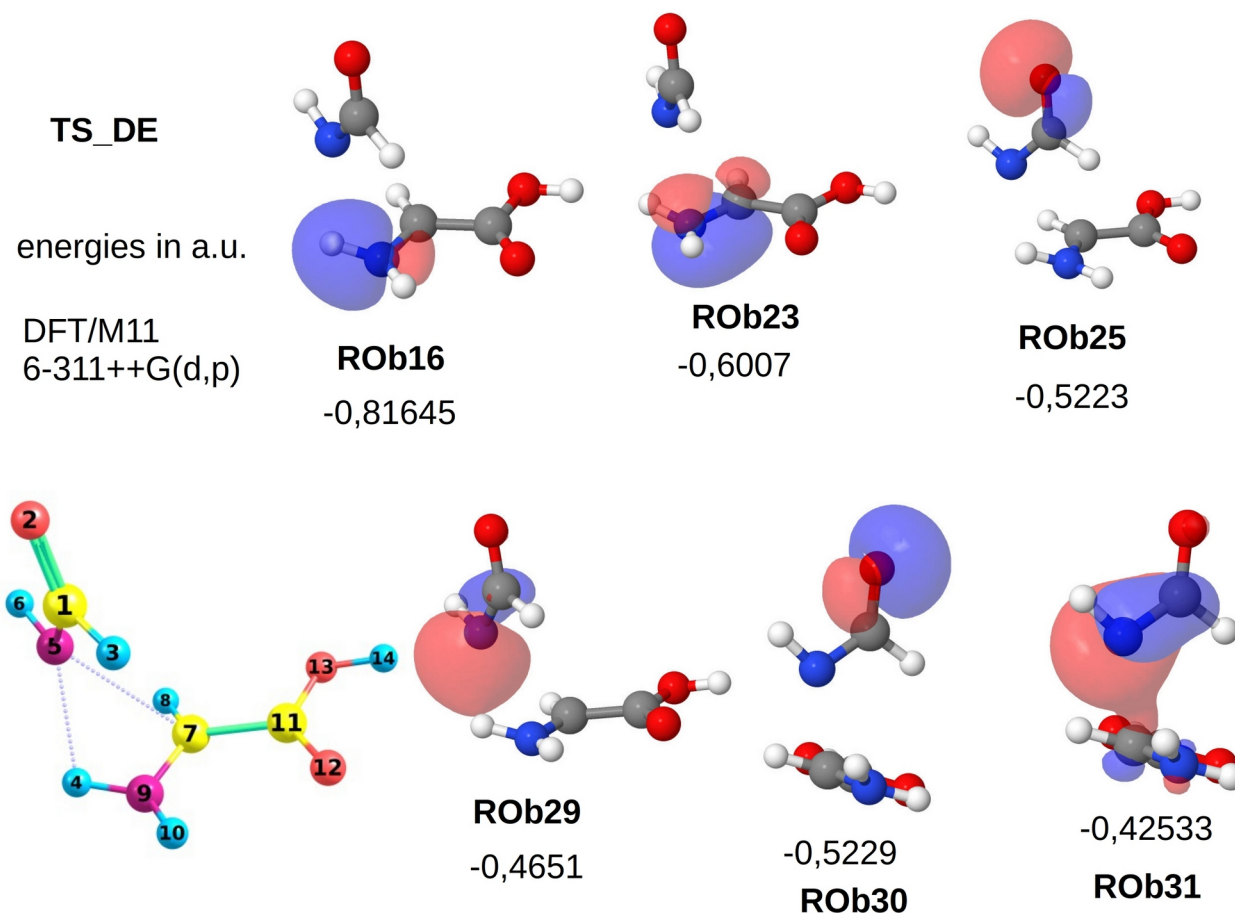
*picture 11: Ruedenberg orbitals of **TS_DE**, on the left: atom numbering of **TS_DE***

On the other hand, these repulsive interactions are compensated by lower RO energies of the formic imidic acid oxygen (atom O2 for **TS_JF** and **TS_DE**, respectively). ROa23 corresponds to ROb25. They are oxygen-p-electrons. The second corresponding p-electron pair of oxygen O2 can be found in ROa25 and ROb30, respectively.

These oxygen-p-electrons are energetically lowered in **TS_JF** due to its interactions with the migrating hydrogen H10. In sum, they are 0,249 a.u. or 156 kcal/mol lower in energy than the oxygen p-electron pairs of **TS_DE**.

Most interestingly, Ruedenberg orbitals clearly illustrate the differences of amide-N- (respective imidic acid -N; atom N5) interaction with the imino acetic acid fragment between **TS_JF** and **TS_DE**. In **TS_JF**, the imidic acid N has a lone pair (ROa26), which forms a sigma-bond to imino acetic acid-C (atom C7), while the corresponding ROb29 of **TS_DE** points to the amide-hydrogen H4 and is 63,6 kcal/mol higher in energy.

Also, the amide/imidic acid N-π-orbitals (the π-electrons of N5) are lower in energy in **TS_JF** (ROa30) than in **TS_DE** (ROb 31). One main difference between ts1 and ts2 lies thus in the fact that the amide-N in TS_JF has a strong sigma- interaction with the iminoacetic acid C (atom C7), while the corresponding N in ts2 shows π-interaction with C7, which is considerably weaker than the former.

13

These energetic disadvantages of the electronic structure **TS_DE** are partly compensated by a strong N-H bond between N9 and H4. This σ-bond is localized in ROb16 of **TS_DE**. It clearly corresponds to ROa29 of **TS_JF**. ROb16 has an energy of -0,816 a.u., and is thus 0,225 au. or 141,4 kcal/mol lower in energy than ROa29.

Formation of a strong sigma bond between N9 and H4 thus proves to be a decisive stabilization factor for **TS_DE**.

Nevertheless, ROb16 can not fully compensate the energy advantages of ROa26 and ROa30, which bind with C7. In sum, the latter two RO's are 0,3202 a.u. lower in energy than the corresponding ROb31 and ROb29.

Summing up the results of Ruedenberg orbital analysis, the imidic acid-nitrogen of **TS_JF** shows a σ-interaction with the reacting C7, while **TS_DE** has a considerably weaker π-bond with the latter atom.

2.2.3. EPS analysis, Hirshfeld Charges and Mayer Bond Orders of TS_DE and TS_JK

Ruedenberg orbital analysis left some questions open, mainly, why N5 in **TS_DE** is not approaching nearer to C7 (which is the electrophilic atom in this state), since N5-D7 is 2,5 angstroms. It was a matter of discussion, whether C7 is an insufficient electrophile, or if another factor draws responsible for the big
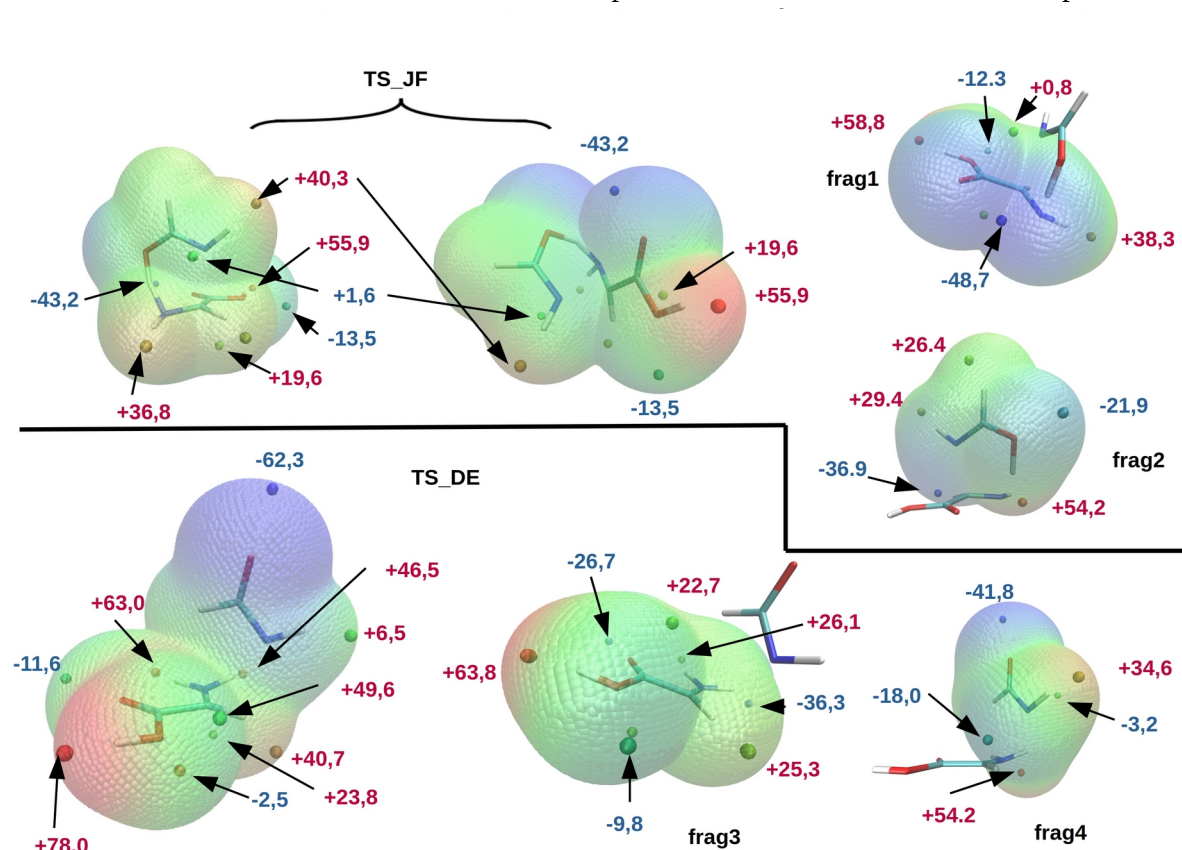

*picture 12: electrostatic potential surface of **TS_JF** and **TS_DE** and their fragments; blue: surface minima, red: surface maxima, energies in kcal/mol*

Both states were split into their iminoacetic acid- and their respective amide/ imidic acid- fragments. The fragments were calculated as single points (DFT/M11, 6-311++G(d,p)) and ESP-analyzed separately. As illustrated in picture 12, frag1 and frag2 are from **TS_JF**, frag3 and frag4 are from **TS_DE**. Note that the migrating hydrogen H4 is assigned to the amide fragment frag 4 and to the imidic acid fragment frag2, respectively. Interestingly, the global EPS maxima of frag2 and frag4 show both an energy of +54,2 kcal/mol. This indicates that the electrophility of H4 is almost the same for the amide-fragment as well as the imidic acid fragment. Moreover, it should be expected that the amide-nitrogen N5 in frag2 has a high nucleophility, since H4 is strongly dissociated, and it should thus show a low lying EPS-minimum within its sphere. But surprisingly, the minimum which was found, is higher in energy (-18,0 kcal/mol) than the N5-minimum of frag4, which has -36,9 kcal/mol. Obviously, H4 withdraws considerably electron density in frag2, thus lowering the nucleophility of N5. This may partly explain, why N5 is very loosely bound to C7 in **TS_DE**.

On the other hand, frag3 shows an EPS-maximum of +22,7 kcal/mol, which indicates high electrophility of C7 within this fragment. However, it does not appear to support the addition reaction. EPS of the whole **TS_DE** shows strong electron density migration from the aminoacetic acid fragment (a large maximum of +78,0 kcal/mol, a maximum near C7 of +49,6 kcal/mol) to the amide, especially to O2 (a large minimum of -62,3 kcal/mol). This is again surprising, since it indicates that the electrophilic attacking site of C7 in frag3 complicates the amide-addition. The low minimum at O2 (-62,3 kcal/mol) furthermore coincides well with high Ruedenberg orbital energies of O2.

In comparison to the latter, it can be seen that in **TS_JF**, electron density migration due to fragment interaction is moderate, since its EPS minima and maxima show only small alterations.
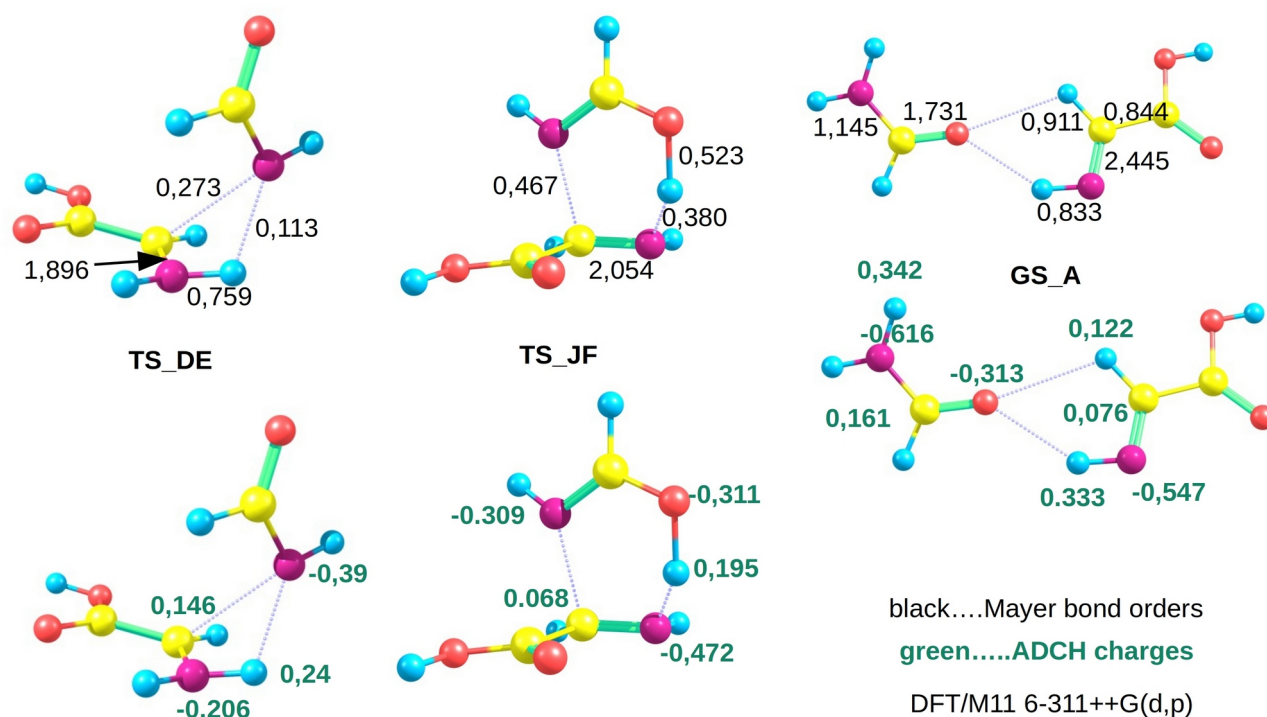

*picture 13: Mayer bond orders and ADC-Hirshfeld charges of **GS_A**, **TS_DE** and **TS_JF***

Atomic dipole-corrected Hirshfeld charges are in good accordance with the data obtained from EPS analysis, as can be seen in picture 13. Although Mayer bond orders in **TS_DE** of N5-C7 and N5-H4 are only 0,273 and 0,113, respectively, the ADCH-charge of N5, which is -0,39, is higher than the N5 in **GS_A**, where the ADCH-charge is -0,616. Together with the results of the EPS-analysis, it can be concluded that the migrating hydrogen H4 withdraws electron density from N5, thus diminishing its nucleophility.

Interestingly, the ADCH-charge of N5 in **TS_JF** (-0,39) is as the same as in **TS_DE**. The reasons therefore are different: The N5-C7-Mayer bond order is 0,467, indicating electron donation from N5 to C7.

2.2.4. EPS Analysis of Several Ground States

Electrostatic potential surfaces of several decisive ground states were calculated in order to gain information about their reactivity. All electrostatic potential surface calculations are based on the DFT/M11-6-311++G(d,p) optimized states.
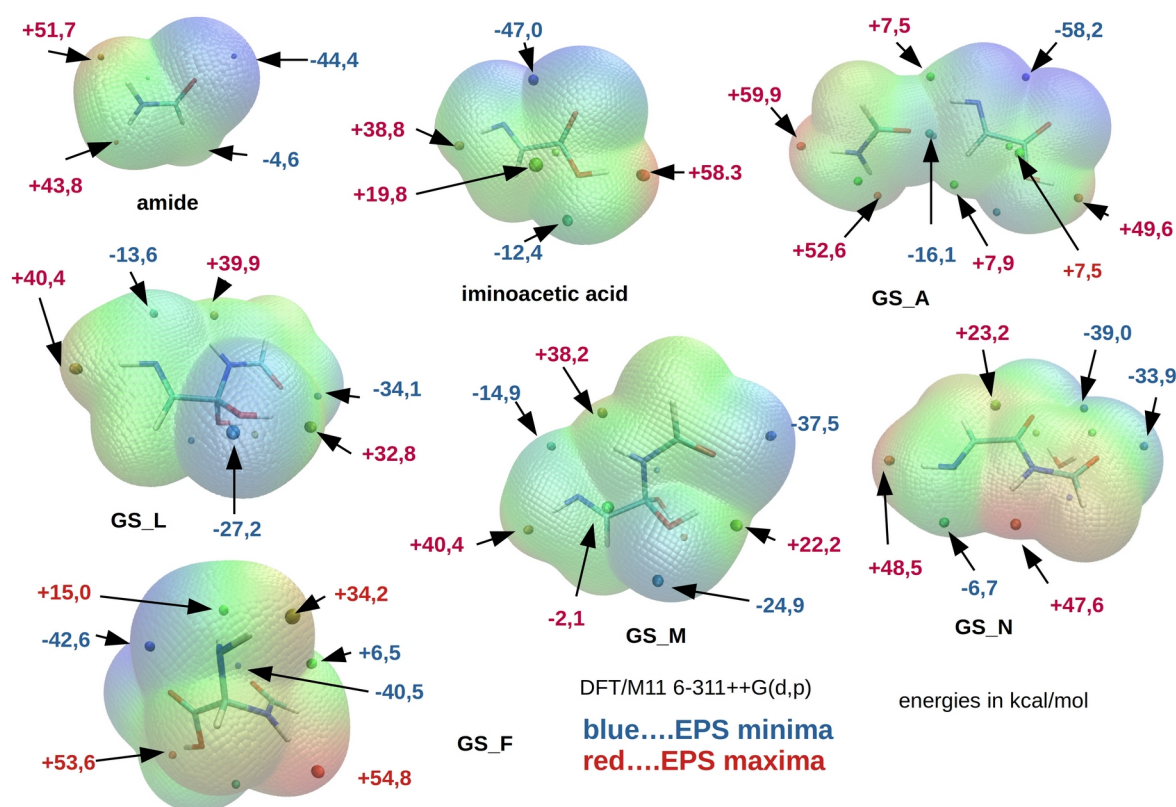
picture 14: electronic potential surfaces of several ground states

An important matter to be discussed is the question, how the side reaction could proceed outgoing from **GS_L**, **GS_M** or **GS_N**. It was supposed that another amide, imidic acid or iminoacetic acid could react with one of these states. Therefore, an EPS minimum, which could probably serve as electrophilic reaction site, was searched for. **GS_M** shows a minimum of -2,1 kcal/mol, which does not indicate big electrophility, nevertheless it could serve as an attacking point for a concerted reaction similar to the imidic acid addition. As can be seen in picture 14, N9 shows an EPS minimum of -14,9 kcal/mol, thus

16

indicating higher nucleophility than iminoacetic acid, where the corresponding minimum is -2,0 kcal/mol. The latter should ease a further H-addition. In sum, it may be considered that imidic acid addition to **GS_M** is possible.

Taking further into account that the diamide **GS_F** is 9,31 kcal/ mol lower in energy (for 6-31++G(d,p)) than **GS_A** with its imine moiety, it seems probable that adding a second amide to **GS_L**, **GS_M** or **GS_N**, respectively, may result in a molecule which is more stable than **GS_F**. Each of the three former states bear an imine moiety, an addition of a further amide should therefore again lower their energy for at about 9 kcal/mol. This could be a subject for further studies.

In **GS_N**, a higher EPS maximum (+23,2 kcal/mol) can be found and is thus supposed to be a suitable electrophilic reaction site. However, formation of **GS_N** via transition state **TS_MN** seems highly unlikely due to its high energy barrier of 42,51 kcal/mol.

Nevertheless, it should be considered that elimination of H_2O may be facilitated by a further amide serving as catalyst. This could also be a matter for further investigation.

Finally, the role of a transition metal catalyst is also a topic of interest. Coordination of iminoacetic acid on a transition metal center, such as Cu^{2+}, should inhibit side reactions. EPS-analysis of isolated iminoacetic acid (see also picture 9) shows a minimum between N3 and O6 of -47,0 kcal/mol. Although there is no minimum which is assigned only to N3 or O6, the average electrostatic surface potential for both atoms, N3 and O6, respectively, can be calculated. For N3 it is -24,5 kcal/mol, for O6, -26,1 kcal/mol is calculated. This is why a coordination to a metal center via O6 is more probable. As a consequence, the side reactions should be inhibited by a metal catalyst. Nevertheless, these EPS-based estimations can only be a starting point for further studies.

Coordination of O6 of iminoacetic acid on a metal center would most probably increase electrophility of C7, thus facilitating a nucleophilic attack of imidic acid N5. On the other hand, nucleophility of N9 would be lowered, thus complicating the migration from H4 to N9. It cannot be predicted which effect would outweigh the other, however, it seems unlikely that the big kinetic difference between amide addition an imidic acid addition can be fully compensated by a catalyst, since EPS-analysis of the present states supports the assumption that the "electronic hole" of C7 has lower influence than the nucleophility of N5. The present study might thus give an insight into the transition metal catalyzed reaction mechanism.

3. Conclusions

The formation of 2-amido-2-aminoacetic acid from 2-iminoacetic acid and formamide is supposed to proceed via imidic acid formation and and imidic acid addition. Direct amide addition shows considerably higher energy barriers. Comparison of the respective rate-determining transition states by their Ruedenberg orbitals have shown that imidic acid adds via strong σ-interaction to the electrophilic C of 2-iminoacetic acid, while formamide shows a weaker π-interaction to the respective C. Although 2-amido-2-aminoacetic acid was lower in energy than any product of a possible side reaction, the rate- determining energy barrier for the formation of N-(1,1- dihodroxy-2-iminoethyl)amide is lower than that of the imidic acid addition. The side reaction is thus supposed to be kinetically favored over the main reaction. Nevertheless, a more stable side reaction-product still has to be found. The results lead to the conclusion that involvement of a transition-metal catalyst should inhibit the side reaction.

4. Acknowledgments

[1]   Gerald B. Rowland,Haile Zhang; Emily B. Rowland; Spandan Chennamadhavuni; Yong Wang; Jon C. Antilla; "Brønsted Acid-Catalyzed Imine Amidation", 2005, J. Am. Chem. Soc., **127**, 45, 15696–15697.

[2] Yuxue Liang; Emily B. Rowland; Gerald B. Rowland; Jason A. Perman, Jon C. Antilla; "*VAPOL phosphoric acid catalysis: the highly enantioselective addition of imides to imines*", 2007, Chem. Commun., **43**, 4477-4479.

[3]   Michael Chorev; Murray Goodman; "A dozen years of retro-inverso peptidomimetics" *Acc. Chem. Res.* 1993, **26**, 5, 266–273.

[4]   Jose F. Hernandez; Jean M. Soleihac; Bernard P. Roques; Marie-Claude Fournié-Zaluski; "*Retro-inverso concept applied to the complete inhibitors of enkephalin-degrading enzymes*", J. Med. Chem., 1988, **31**, 1825-1831.

[5] Mauro Marastoni; Severo Salvadori; Fabrizio Bortolotti; Roberto Tomatis; "*HIV-1 protease inhibitors containing a novel C_2 symmetrical hydroxyalkylgem-diamino core structure*", J. Peptide Res., 1997, **49**, 538-544, https://doi.org/10.1111/j.1399-3011.1997.tb01161.x

[6]   Yoshio Nishimura; Eiki Shitara; Hayamitsu Adachi; Minako Toyoshima; Motowo Nakajima; Yoshiro Okami; Tomio Takeuchi; "*Flexible Synthesis and Biological Activity of Uronic Acid-Type gem-Diamine 1-N-Iminosugars: A New Family of Glycosidase Inhibitors*", J. Org. Chem., 2000, **65**, 2-11.

[7] Shujie Zhu; Shaomin Fu; Huanfeng Jiang; Wei Zeng; "*Cu(II)-Catalyzed Intermolecular Addition of C-Acylimine: A Convenient Access to gem-Diamino Acid Derivatives*", Org. Lett., 2011, **13**, 18, 4914-4917.

[8] Dongmei Li; Yong Wang; Keli Han;"*Recent density functional theory model calculations of drug metabolism by cytochrome P450*", Coord. Chem. Rev, 2012, **11-12**, 1137-1150.

[9] Jann Frey.; Samuel Leutwyler; "*Binding Energies of Hydrogen-Bonded cis-Amide and Nucleobase Dimers: An Evaluation of DFT Performance*", 2005, Chimia, **59**, 511-516.

[10] Pedro Salvador, P.; J.J. Dannenberg; " *Dependence upon Basis Sets of trans Hydrogen-Bond $^{13}C-^{15}N$ 3-Bond and Other Scalar J-Couplings in Amide Dimers Used as Peptide Models. A Density Functional Theory Study*", 2004, J. Phys. Chem. B, **108**, 15370-15375.

[11] Thomas Neuheuser; Bernd Hess; Christiane Reutel; Edwin Weber;" *Ab Initio Calculations of Supramolecular Recognition Modes. Cyclic versus Noncyclic Hydrogen Bonding in the Formic Acid/Formamide System*", 1994, J. Phys. Chem. , **98**, 6459-6467.

[12] Yevgeniy Podolyan; Leonid Gorb; Jerzy Leszczynski; " *Double-Proton Transfer in the Formamidine−Formamide Dimer. Post-Hartree−Fock Gas-Phase and Aqueous Solution Study*", 2002, J. Phys. Chem. A, **106**, 12103-12109.

[13] Georg Dazinger; "*Substituent Effects on the Amide - Imidic Acid Tautomerism: A DFT-Study and Wavefunction Analysis*", preprint, 2020, DOI: 10.13140/RG.2.2.25371.90405

[14] Cheok N. Tam; Petr Bour; Juergen Eckert; Frans R. Trouw; "*Inelastic Neutron Scattering Study of Hydrogen-Bonded Solid Formamide at 15 K*", 1997, J. Phys. Chem. A, **101**, 5877-5884.

[15] Hiroshi Ohtaki; Atsushi Funaki; Bernd M. Rode; Gilbert J. Reibnegger; "*The Structure of Liquid Formamide Studied by Means of X-Ray Diffraction and ab Initio LCGO-MO-SCF Calculations*", 1983, Chem. Soc. Jpn., **56**, 2116-2121.

[16] Roberto Peverati; Donald G. Truhlar*; "M11-L: A Local Density Functional That Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics". Journal of Physical Chemistry Letters. 2012, 3 (1): 117–124. doi:10.1021/jz201525m*

[17] Coleman J. Howard; Jordan D. Enyard; Gregory S. Tschumper; "*Assessing the accuracy of some popular DFT methods for computing harmonic vibrational frequencies of water clusters*" J. Chem. Phys. **143**, 214103 (2015); https://doi.org/10.1063/1.4936654

[18] Roberto Peverati, Donald G. Truhlar,*"Performance of the M11 and M11-L density functionals for calculations of electronic excitation energies by adiabatic time-dependent density functional theory", Phys. Chem. Chem. Phys.,* 2012,**14**, 11363-11370.

[19] Warren J. Hehre; Robert Ditchfield; John A. Pople; "*Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules*", J. Chem. Phys.,1972, **56**, 2257-2261.

[20] P. C. Hariharan; John A. Pople; "*The influence of polarization functions on molecular orbital hydrogenation energies*", Theor. Chim. Acta, 1973, **28**, 213-222.

[21] Raghavachari Krishnan; Stephen J. Binkley; Raymond J. Seeger; John A. Pople; "*Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions*", J. Chem. Phys., 1980, **72,** 650-654.

[22], Timothy Clark; Jayaraman Chandrasekhar; Günther W. Spitznagel; Paul Von Ragué Schleyer; "*Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li-F*", J. Comput. Chem., **4**, 1983, 294-301.

[23] , Benjamin P. Pritchard; Doaa Altarawy; Brett Didier; Tara D. Gibsom; Theresa L. Windus; "*A New Basis Set Exchange: An Open, Up-to-date Resource for the Molecular Sciences Community", J. Chem. Inf. Model.",* 2019, **59**, 4814-4820, doi:10.1021/acs.jcim.9b00725.

[24] David Feller; "*The role of databases in support of computational chemistry calculations*", J. Comput. Chem., 1996, **17**, 1571-1586, doi:10.1021/ci600510j.

[25] Karen L. Schuchardt, Brett T. Didier, Todd Elsethagen, Lisong Sun, Vidhya Gurumoorthi, Jared Chase, Jun Li, Theresa L. Windus, "*Basis Set Exchange: A Community Database for Computational Sciences*", J. Chem. Inf. Model., 2007, **47**, 1045-1052.

[26] Carlos Gonzales; Bernhard H. Schlegel; "*An improved algorithm for reaction path following*", J. Chem. Phys., 1989, **90**, 2154-2162.

[27] Samuel F. Boys; Fernando Bernardi; "*The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors*", Mol. Phys., 1970, **19**, 553-556.

[28] Frans B. van Duijneveldt; Jeanne G. C. M. van Duijneveldt-van de RijdtJ; Joop H. van Lenthe, "*State of the Art in Counterpoise Theory*" Chem. Rev., 1994, **94**, 1873-1885.

[29] Michael W. Schmidt; Kim K. Baldridge, Jerry A. Boatz, Stephen T. Elbert, Mark S. Gordon, Jan H. Jensen; Shiro Koseki; Nikita Matsunaga; Kiet A. Nguyen; Shugun J. Su; Teresa L. Windus; Michel Dupuis; John A. Montgomery; "*General Atomic and Molecular Electronic Structure System*", J. Comput. Chem. 1993, **14**, 1347-1363 doi:10.1002/jcc.540141112

[30] "*Advances in electronic structure theory: GAMESS a decade later*" Mark S.Gordon; Michael W. Schmidt, Chapter 41, 1167-1189, in
"*Theory and Applications of Computational Chemistry, the first forty years*", Clifford Dykstra;
Gernot Frenking; Kwang Kim; Gustavo Scuseria; editors, Elsevier, Amsterdam, 2005.

[31] Tian Lu; Feiwu Chen; "*Atomic Dipole corrected Hirshfeld Population Method*", J. Theor. Comput. Chem., 2012, **11**, 163-183. https://doi.org/10.1142/S0219633612500113

[32] Fred L. Hirshfeld; "*Bonded-atom fragments for describing molecular charge densities*", Theor. Chim. Acta, 1977, **44**, 129-138.

[33] Istvan Mayer; "*Charge, bond order and Valence in the AB ab initio SCF theory*", Chem. Phys.Lett., **97**, Issue 3, (1983), 270-274, doi: https://doi.org/10.1016/0009-2614(83)80005-0

[34] Adam J. Bridgeman; German Cavigliasso; Luke R. Ireland; Joanne Rothery; "*The mayer bond order as tool in inorganic chemistry*" J. Chem. Soc., Dalton Trans., 2001, 2095-2108, doi: https://doi.org/10.1039/B102094N

[35] Tian Liu, Feiwu Chen, "*Quantitative analysis of molecular surface based on improved Marching Tetrahedra algorithm*", J.Mol.Graph.Model, 2012, **38,** 314-323, https://doi.org/10.1016/j.jmgm.2012.07.004

[36] Clyde Edmiston; Klaus Ruedenberg; "*Localized Atomic and Molecular Orbitals*", Rev. Mod. Phys. 1963, **35**, 457-465.

[37] "*Jmol: an open-source Java viewer for chemical structures in 3D.*" http://www.jmol.org/

[38] Tian Lu, Feiwu Chen, "*Multiwfn: a multifunctional wavefunction analyzer*" J. Comput. Chem., 2012, **33,** 580-592. doi: https://doi.org/10.1002/jcc.22885

[39] William Humphrey, Andrew Dalke, Klaus Schulten, "*VMD - Visual Molecular Dynamics*" *J. Molec. Graphics* 1996, **14.1**, 33-38.

[40] "*Chemcraft - graphical software for visualization of quantum chemistry computations.*" https://www.chemcraftprog.com