university of groningen

A DFT study on the prebiotic formation of peptides via thiolysis

Geert-Jan de Putter

MSc Thesis

December 11, 2023

Student: Geert-Jan de Putter S3496201 g.de.putter@student.rug.nl

Supervisor UAntwerp: Prof. Dr. A. Da Cunha
First assessor RUG: Prof. Dr. M.T.C. Walvoort
Second assessor RUG: Prof. Dr. W.R. Browne

Contents

1 Abbreviations

ADF	Amsterdam Density Functional
AMBER	Assisted Model Building with Energy Refinement
AMS	Amsterdam Modelling suite
ASM	Activation Strain Model
B3LYP	Becke, 3-parameter, Lee-Yang-Parr
CHARMM22	Chemistry at Harvard Macromolecular Mechanics 22
COSMO	COnductor-like Screening MOdel
COSMO-RS	COnductor-like Screening MOdel Real Solvent
DFT	Density Functional Theory
DFTB	Density Functional based Tight Binding
disp	dispersion
DNA	DeoxyriboNucleic acid
EDA	Energy Decomposition Analysis
elstat	electrostatic
FF	Force Field
GBSA	Generalized Born model hydrophobic solvent accessible Surface Area
GFN1-xTB	Geometry, Frequency, Noncovalent, eXtended Tight Binding
GGA	Generalized Gradient Approximation
GTO	Gaussian-Type Orbital
HOMO	Highest Occupied Molecular Orbital
int	interaction
IRC	Intrinsic Reaction Coordinate
LDA	Local Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital

M06-2X	Minnesota functional 06, double the amount of local exchange with respect to M06
MBD	Many Body Dispersion
MBD@rsSCS	Many Body Dispersion, range separated Self-Consistent Screening
MEP	Minimal Energy Path
MO	Molecular Orbital
MP2	second order Møller-Plesset Perturbation theory
NEB	Nudged Elastic Band
oi	orbital interaction
PBE0	Perdew–Burke-Ernzerhof, 3:1 ratio to Hartree-Fock exchange energy
PES	Potential Energy Surface
PMF	Potential of Mean Force
PNA	Peptide Nucleic Acid
QM/MM	Quantum Mechanics, Molecular Mechanics
RC	Reaction Coordinate
RNA	RiboNucleic acid
STO	Slater-Type Orbital
SZ	Single Zeta
TIP3P	Transferable Intermolecular Potential with 3 Points
TS	Transition State
TZ2P	Triple Zeta 2 Polarizable
UFF	Universal Force Field
VdW	Van der Waals
XC	eXchange Correlation

2 Introduction

2.1 The origin of life

The origin of life is a question that has puzzled humanity since the dawn of mankind. Darwin proposed in his book 'On the Origin of Species' [1], that all the life forms are the result of evolution through natural selection from a common ancestral form of life. It is important to define what life, or a living system, actually is, before going into what the first molecules of life most likely are. Horowitz and Miller [2] defined life as following: "an organism, to be called living, must be capable of both replication and mutation: such an organism will evolve into higher forms". Replication implies the existence of DNA or RNA in a living system. However, it is not likely that either RNA or DNA were the first molecules that were used for self-replication in living systems [3]. Non-enzymatic selective formation of 3'-to-5' phosphodiester linkages present in RNA is difficult due to competing 2'-to-5', and 5'-to-5' phosphodiester linkages. Thus, it is proposed that it is more likely that simpler self-replicating molecules were responsible for self-replication, such as Peptide Nucleic Acids (PNA). This implies that peptides were most likely the first bio-polymers. From this point, the main focus will be on the prebiotic formation of amino acids and peptides.

Different theories have been proposed how these peptides were formed from amino acids, and how the first amino acids were formed in a prebiotic environment. One of these theories on the formation of precursors to amino acids, was proposed by A. Oparin in 1924 [4]. He proposed that the first compounds containing carbon were metal carbides, which were stable enough to the harsh conditions on earth in its early stages. When these metal carbides, in the form of lava, came in contact with water, primitive hydrocarbons were formed, which upon oxidation turned into carbon monoxide and other derivatives. Furthermore, the atmosphere was hypothesized to be of a reducing atmosphere, containing nitrogen gas, oxygen gas, water, methane, and primitive hydrocarbon derivatives. Reactions between primitive hydrocarbon derivatives and ammonia gave rise to amino acid precursors. It is then hypothesized that these precursors and various other organic compounds formed colloidal solutions in water, where at some point, a self replicating system was formed.

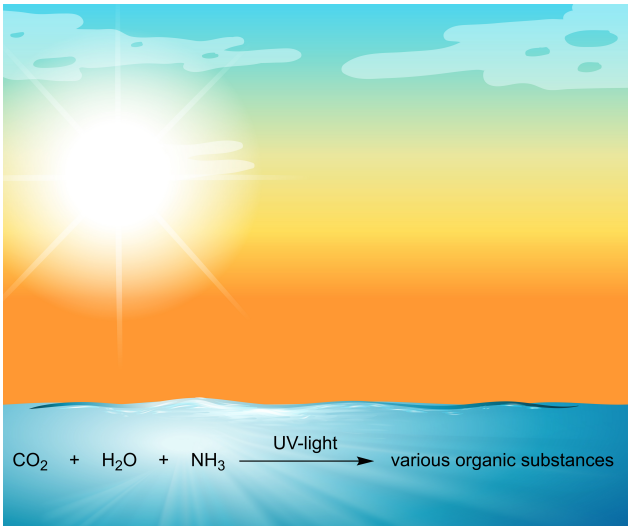


Figure 1: Representation of primordial soup as proposed by Oparin and Haldane [4, 5]

Similarly to A. Oparin, J. B. S. Haldane proposed in 1929 [5] that the earth first cooled down to a temperature above the boiling point of water, but low enough for a crust to form on earth. The atmosphere was hypothesized to contain little to no oxygen gas, and to mainly consist out of carbon dioxide. Nitrogen would mainly be found as part of metal nitrides, where ammonia was constantly formed from reacting with water. Since there was no ozone layer, the UV-light intensity was high. When UV-light acts on a mixture of water, carbon dioxide, and ammonia, a large variety of organic substances would be formed, among which amino acids. Where both Oparin's and Haldane's theories point towards a primordial soup, as presented in figure 1, in which organic substances were formed, among which amino acids.

2.2 Prebiotic formation of amino acids and peptides

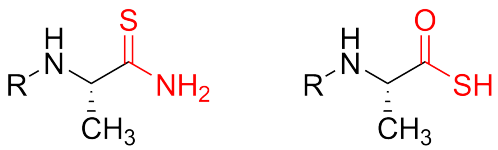
Inspired by the theory of a primordial soup, S. Miller and H. Urey performed the Miller-Urey experiment in 1952 [6]. In this experiment, carbon dioxide, ammonia, water, and hydrogen gas were placed in a vacuum system and passed along an electric discharge to simulate thunder. With use of thin layer chromatography, it was demonstrated that the amino acids glycine, α-alanine, and β-alanine were formed, together with various uncharacterized compounds. S. Miller used various set-ups for this experiment, among which was an experiment that simulated a volcanic eruption environment by using a jet nozzle to inject gasses into a spark. From this experiment, Miller identified five amino acids, together with uncharacterized compounds. Further characterization of the mixture by J. L. Bada et al. in 2008 [7] demonstrated that a mixture of 22 amino acids was

present, together with 5 different amines.

A different hypothesis was given by W. Martin et al. [8], known as the iron-sulphur hypothesis. It is hypothesized that amino acids and various other organic substrates were formed near hydrothermal vents, where iron sulfide minerals played a catalytic role. As precursors for these organic substrates, carbon monoxide, carbon dioxide, hydrogen cyanide, and hydrogen sulfide have been proposed.

Similar to iron sulfide minerals acting as catalyst, E. Anders et al. [9] proposed the use of nickel, aluminium, and clay catalysts for the formation of amino acids using carbon monoxide, nitrogen, and hydrogen gas as precursors. In this experiment, it was demonstrated that various amino acids could indeed be synthesized, albeit found at yields ranging from 0.01 to 0.1% with temperatures as high as 1000 °C.

The problem with hydrothermal vents in the ocean, is that the concentration of organic molecules will be quite low. Since the reaction rate is dependent on the concentrations of reactants, a low concentration leads to low reaction rates. Thus it might be unlikely that reactions between these organic molecules took place at these low concentrations in the ocean. A solution to this problem is given by the 'dry pond, wet pond' theory. In this theory, it is proposed that the first amino acids and peptides originate from a hydrothermal crater-lake [10]. In this hydrothermal crater-lake, it is proposed that there is a neutral pH, low salinity, and various redox states of iron and sulfur present. With the presence of minerals such as the different redox states of iron, various simple organic molecules could be formed. Furthermore, it is proposed that due to wet and dry cycles of the lake, polymerization of amino acids is facilitated, since this requires dry conditions by the absence of water [10, 11]. These wet and dry cycles can vary in time: short cycles would only last a few minutes where the water source is a geyser and the pool is small, an intermediate cycle would last hours where water comes into a pool from other hot springs, and a long cycle would last days to weeks where water is lost from a relatively large pool due to evaporation [12].



Thioamide **Thioacid**

Figure 2: Molecular structures of amino acid precursors

A different class of proposed amino acid precursors are the amino thioacids and -thioamides, which can be seen in figure 2. Patel et al. [13] have demonstrated that the amino acid Leucine is obtained from the addition of hydrogen sulfide to a nitrile. All of the precursors in their experiment are derived from simple building blocks such as hydrogen cyanide, ethyne, water, hydrogen sulfide, and simple inorganic salts such as copper (I) chloride. Key in their experiments is the use of hydrogen cyanide and hydrogen sulfide as source for carbon, nitrogen, and sulfur [13].

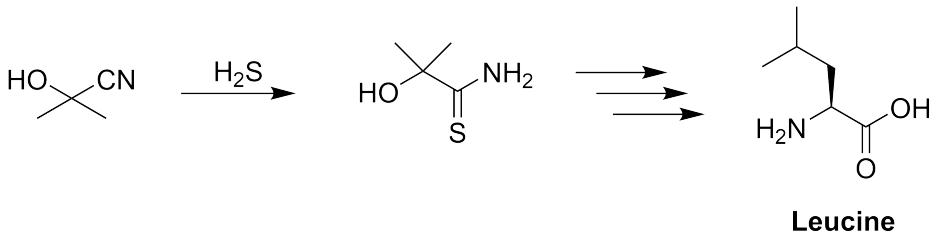


Figure 3: Synthesis of Leucine via addition of hydrogen sulfide to a nitrile [13]

The intermediate in figure 3, a thioamide, was subsequently used as key substrate for peptide ligation by Canavelli et al. [14]. Key in their reaction, was the addition of hydrogen sulfide to a nitrile, followed by hydrolysis to get a thioic acid. It was found out, that for the addition of hydrogen sulfide, the starting material **1.1** first has to be acetylated to **1.4**, followed by hydrogen sulfide addition to obtain **1.5**. If the acetylation of **1.1** was not performed, the hydrolysis of the thioamide **1.2** would not take place. The thiolysis of **1.5** to **1.6** on the other hand did take place, since the carbon as part of the thioamide is more electrophilic on **1.5**.

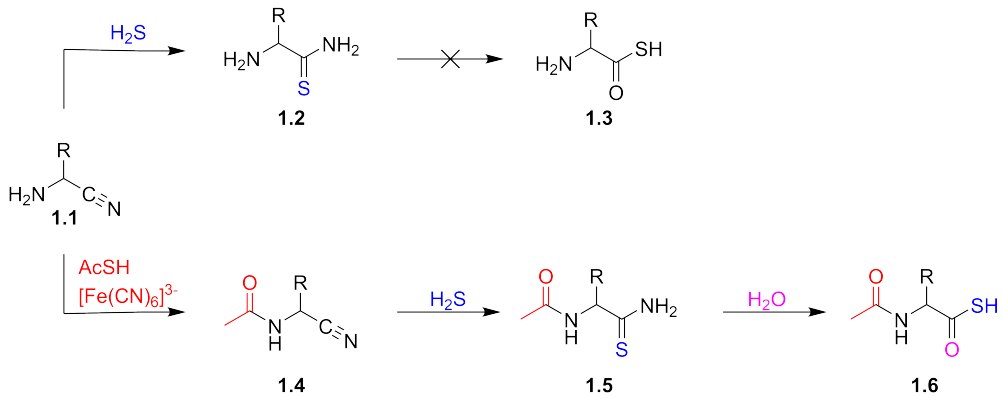


Figure 4: Synthesis of amino thioacids by Canavelli et al. [14]

With a reliable synthetic pathway from **1.1** to **1.6**, ligations were performed. In figure 5, the ligation cycle is presented, where in the first step, an addition between **1.1** and **1.9** happens to form a peptide bond, and gives **1.7**. Upon hydrogen sulfide addition, **1.8** is obtained, which gives

1.9 upon hydrolysis. Repeating this cycle will elongate the peptide. With this method, all of the 20 proteinogenic amino acid residues can be used for the ligation of amino acids to peptides. These reactions are facilitated by ferric cyanide, cupric salts, cyanoacetylene, and *N*-cyanoimidazole, and at a pH of 7, demonstrating that this reaction can take place in water pools with conditions found on the early earth. Thus it is regarded as a possibility as to how the first peptides were formed [14].

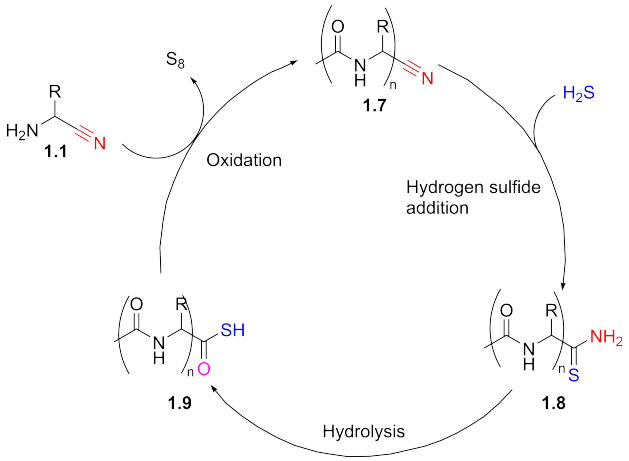


Figure 5: Ligation of amino thioacids by Canavelli et al. [14]

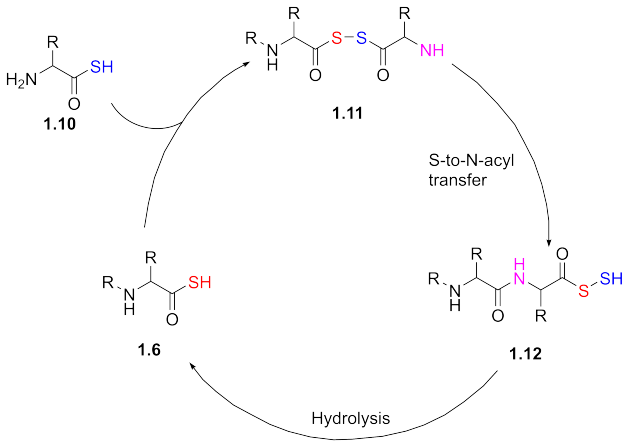


Figure 6: Ligation of amino thioacids by Okamoto et al. [15]

Similarly to Canavelli et al. [14], Okamoto et al. [15] made use of amino thioacids. In the paper of Okamoto et al., the focus was on the synthesis of polypeptides without the use of protecting groups. Key to their oligomerization reactions was the use of $FeCl_3$, and a pH $\leq$ 2.5. To simulate the conditions on the primordial earth, iron ore was used as a source of Fe^{3+}, and sulfuric acid was used as an acid, where it was hypothesized that sulfuric acid could be the result of the presence of sulfur dioxide in water. In this simulation, it was demonstrated that polypeptides up to tripeptides,

could be synthesized. in figure 6, the cycle for the synthesis of polypeptides is shown. In the first step of the reaction, **1.6** and **1.10** are oxidized to give disulfide **1.11**. Then, via an *S*-to-*N*-Acyl transfer, peptide **1.12** is obtained. Then, via hydrolysis, **1.6** is obtained, where the R-chain has now been elongated, and the cycle is repeated.

2.3 Chiral preference for L-amino acids

In nature, there is a large preference for L-amino acids instead of D-amino acids. There are various hypotheses as to why this preference can be found in nature. One of these hypotheses is the autocatalysis of a reaction, which can enhance a small initial enantiomeric excess, demonstrated by Soai et al. [16]. In this reaction, the alkylation of pyrimidine-5-carbaldehyde by diisopropylzinc is autocatalyzed by its product, where with a small enantiomeric excess of 5% of this product, an enantiomeric excess of 55% was obtained. It is hypothesized that by this process of auto catalysis, an enantiomeric excess of L-amino acids was obtained by a small initial enantiomeric excess [16]. A different method for the origin of homochirality has been proposed by Weissbuch et al. [17], where it was demonstrated that the addition of L-amino acids to floating glycine crystals in water, leads to the exclusion of the addition of D-amino acids to this crystal. It is then further proposed that upon a statistical fluctuation in the orientation of these crystals, a region of homochirality is created. However, this does imply that there will be regions of homochirality for both L- and D-amino acids, but with a small enantiomeric excess of L-amino acids, it is hypothesized that the addition of the L-amino acids to the glycine crystals will lead to chiral amplification [17]. Many other theories have been proposed as well, ranging from the interaction of polarized light with matter, extra-terrestrial origin, parity violation, and chiral fields [18].

2.4 Importance of solvent effects

In general, reactions take place in a solvent, such as an organic solvent or water. It is important to consider these solvents in computational studies, since they can have great effects on the reaction rates of reactions [19]. Thus, to accurately describe a reaction with a computational study, solvent effects will need to be included. A method to account for these solvent effects is known as QM/MM, where there is a region described by quantum mechanics, for example the region where a reaction of interest takes place, and a region described by molecular mechanics, which is used to describe the solvent explicitly. Typically, QM/MM methods are used to model a wide range of reactions, such

as enzymatic reactions. The reaction mechanism of *N*-acetyl neuraminic acid lyase was elucidated with a QM/MM method, where functionals such as B3LYP and MP2 were used in combination with the CHARMM22 forcefield [20]. Similarly, Figueiredo et al. [21], made use of QM/MM methods to describe the enzymatic synthesis of polycaprolactone, where a PMF forcefield was used to describe the MM region consisting out of water, and B3LYP to describe the QM region. With their QM/MM method, they could accurately describe the catalytic effects of the enzymes [21]. Peptide bond formation has been demonstrated by Hu et al. [22], where a lysine and asparagine substrate couple together aided via a glutamate catalyzed proton transfer in a water system. The QM region was described by the B3LYP functional, and the MM region was described by the CHARMM forcefield and the TIP3P water model to describe the water system [22]. In a different study on peptide bond formation, Świderek et al. [23] demonstrated Ribosome catalyzed peptide bond formation. In their study, the QM region was calculated with the M06-2X hybrid functional, and the MM system was described by the AMBER forcefield and TIP3P water model. The functional PBE0 has been used with success in QM/MM calculations, demonstrating its applicability for this thesis [24, 25, 26, 27].

2.5 Introduction to the research question

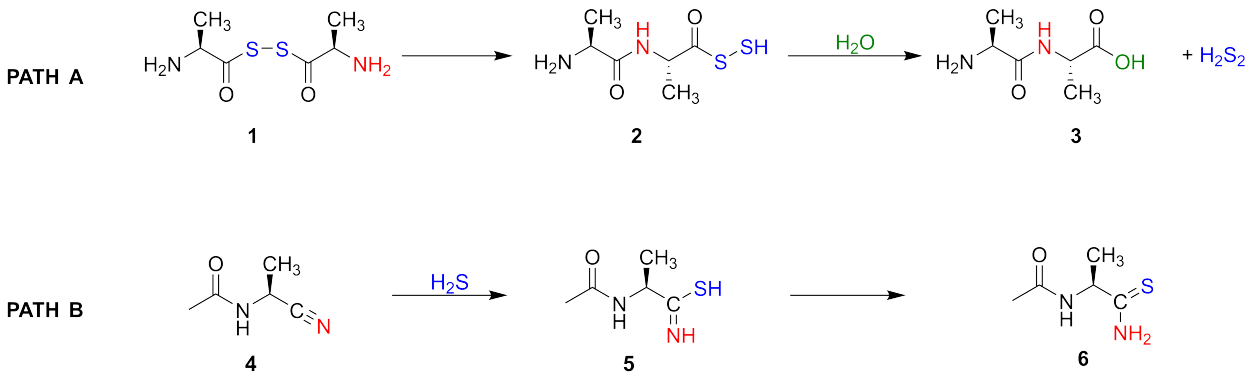


Figure 7: Path A and Path B

In this thesis, the focus will be on two paths. The first one being path A (figure 7), which is based on the research of Okamoto et al. [15]. In this path, three substrates will be considered, where we want to investigate how the *S*-to-*N*-Acyl transfer happens from **1** to **2**, and the subsequent hydrolysis to **3**. As amino acid precursor, the stereochemistry and structure of L-Alanine will be used. This amino acid is chosen, since it has one of the smallest structures, yet is chiral in comparison with L-Glycine. This will be done with use of DFTB, DFT, and QM/MM to investigate the influence

of water as solvent on the reaction energies. The details of this will be described in the Methods chapter. The second path, path B, is based on the research of Canavelli et al. [14]. Similarly to path A, DFTB, DFT, and QM/MM will be used to investigate the reaction mechanism and reaction energies, with the emphasis to investigate the effects of water as solvent on the reaction mechanism and reaction energies. The details of this will be described in the Methods chapter. On both of these mechanisms, no QM/MM study has been done yet, and by comparing the energies between path A and path B, we want to figure out whether path A or path B is more probable to happen under prebiotic circumstances. Additionally, Path B will be investigated with the stereochemistry and structure of D-Alanine, to investigate whether there is a preference of either L- or D-Alanine over one another.

3 Theoretical background

3.1 The basis of quantum chemistry

The very basis of quantum chemistry can be traced back to the Schrödinger equation, which can be considered as either a time-dependent- and time-independent equation.

$$\hat{H}\Psi(r,t) = i\hbar\frac{\partial}{\partial t}\Psi(r,t) \tag{1}$$

In equation 1, the time-dependent Schrödinger equation can be seen. In this equation, $\hat{H}$, is the time dependent Hamiltonian, which acts on the wavefunction Ψ, where the wavefunction Ψ contains both the variables r and t, where r is the position of a system and t the time.

$$\hat{H}\Psi(r) = E\Psi(r) \tag{2}$$

In equation 2, the time-independent Schrödinger equation can be seen. The biggest difference with the time-dependent Schrödinger equation is the fact that the time-independent Schrödinger equation gives back an eigenvalue, E, which is a real and observable value.

3.2 Born-Oppenheimer approximation

The Hamiltonian, as presented in equation 2, can be further separated into a kinetic- and potential energy operator, which gives: $\hat{H} = \hat{T} + \hat{V}$, where $\hat{T}$ is the kinetic energy operator and $\hat{V}$ is the potential energy operator. These kinetic- and potential energy operators can be further separated into operators specific for the nucleus and electrons [28].

$$\hat{H} = -\frac{\hbar^2}{2}\sum_{\alpha}\frac{1}{m_\alpha}\nabla_\alpha^2 - \frac{\hbar^2}{2m_e}\sum_i\nabla_i^2 + \sum_\alpha\sum_{\beta>\alpha}\frac{Z_\alpha Z_\beta e^2}{4\pi\epsilon_0 r_{\alpha\beta}} - \sum_\alpha\sum_i\frac{Z_\alpha e^2}{4\pi\epsilon_0 r_{i\alpha}} + \sum_j\sum_{i>j}\frac{e^2}{4\pi\epsilon_0 r_{ij}} \tag{3}$$

In equation 3, which is the molecular Hamiltonian, it can be seen that there five terms present. In this equation, α and β refer to the nuclei, i and j refer to the electrons [29]. The first term is the kinetic energy operator for the nuclei, the second term is the kinetic energy operator for

the electrons. The third term describes the potential energy due to repulsion between multiple nuclei, where $r_{\alpha\beta}$ is the distance between these nuclei. the fourth term describes the potential energy due to attraction between nuclei and electrons. The final term describes the potential energy due to repulsion between electrons. Within the Born-Oppenheimer approximation, the nuclear and electronic motions are separated from each other, where is approximation is based on the fact that nuclei move slower with respect to electrons, and subsequently respond slower to a change in electron distribution, whereas electrons will respond instantaneously to a change in nuclei distribution. Furthermore, the wave function can be approximated into two parts, an electronic- and nuclear wave function, which has been represented in equation 4.

$$\psi(q_i, q_\alpha) = \psi_{el}(q_i; q_\alpha)\psi_N(q_\alpha) \tag{4}$$

In equation 4, it can be seen that the wavefunction $\psi(q_i, q_\alpha)$ can be divided into an electronic wavefunction and a nuclear wavefunction. The electronic wavefunction, $\psi_{el}(q_i; q_\alpha)$, is a function of the electron coordinates, and depends parametrically on the nuclear coordinates. Furthermore, the electronic wavefunction is used to create a potential energy surface, which is a multi-dimensional surface, which is key for finding minima and transition states for a molecule.

3.3 Potential Energy Surfaces

As described before, the Born-Oppenheimer approximation leads to a PES. The number of dimensions of this PES is equal to $3N-6$ nuclear coordinates for a non-linear molecule, which thus range from $q_1, q_2, ..., q_{3N-6}$, where the energy U is a function of these coordinates [29].

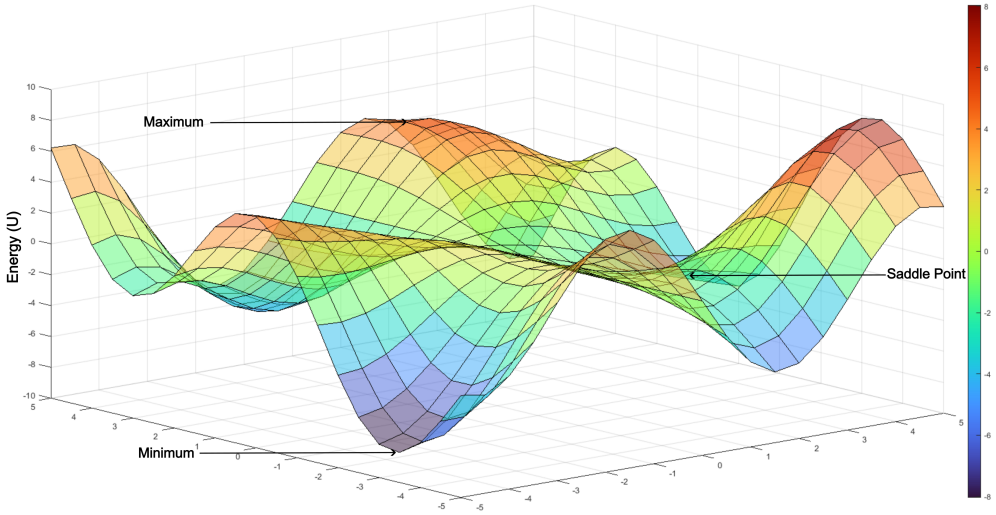


Figure 8: Example of a potential energy surface

In figure 8, a trivial PES can be seen, where the deep blue zones denote minima, yellow zones saddle points, and dark red zones maxima. All of the x and y coordinates correspond to a geometry of a molecule, where the z-axis is the energy U. From this figure, it is clear that there are multiple minima, saddle points, and maxima that can be found for a molecule, which will greatly increase in numbers with more and more nuclear coordinates present.

$$H = \begin{bmatrix} \frac{\partial^2 E}{\partial q_1^2} & \frac{\partial^2 E}{\partial (q_1, q_2)} & \cdots & \frac{\partial^2 E}{\partial (q_1, q_n)} \\ \frac{\partial^2 E}{\partial (q_2, q_1)} & \frac{\partial^2 E}{\partial q_2^2} & \cdots & \frac{\partial^2 E}{\partial (q_2, q_n)} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 E}{\partial (q_n, q_1)} & \cdots & \cdots & \frac{\partial^2 E}{\partial (q_n^2)} \end{bmatrix} \tag{5}$$

In equation 5, a Hessian matrix can be seen. With this matrix, the second derivative of the nuclear coordinates q are taken with respect to the energy E. From this matrix, eigenvalues can be calculated, which describe the normal modes of a molecule. Furthermore, these eigenvalues can be related to IR frequencies. When all the eigenvalues are positive, the set of nuclear coordinates are in a local or global minimum. With one negative eigenvalue, the nuclear coordinates are positioned on a saddle point, which characterizes a transition state. If there is more than one negative eigenvalue present, the set of nuclear coordinates are neither present in a minimum or a transition state [29].

3.3.1 Potential energy surface scans and searches

To get an idea of how the PES looks like, there are various methods available to map a PES. One of such methods is known as 'Basin Hopping', which has been represented in figure 9.

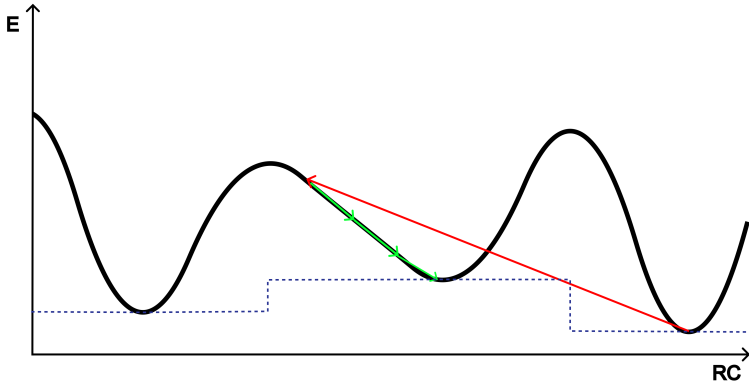


Figure 9: Graphical representation of the process of basin hopping on a PES

In figure 9, a simple representation of a PES in 2 dimensions is given as black line. In the method basin hopping, it is common to smooth the PES to remove transition states, which improves finding local minima and the hopping between minima, this has been represented by the blue dotted line [30]. However, a smoothening of the PES can lead to problems, where a minimum from the deformed PES does not necessarily correspond to a minimum on the original PES [31]. A general process of basin hopping starts from a local minimum, at the start of the red arrow. Then, via a random displacement of nuclear coordinates, a new point on the PES is reached, at the end of the red arrow. This is then followed by a minimization of energy to reach a new local minimum. Finding transition states does not work with this method, since the PES has been smoothed out, and transition states have been removed. With the method 'Nudged Elastic Band', one can find transition states if the initial and final state are known. Via interpolation between the initial and final state, a number of geometries is created, followed by a geometry optimization, where with use of added forces, a geometry is only allowed to optimize towards a structure with an energy in between its neighbours. Finally, the geometry with the highest energy is forced towards a transition state by 'nudging' it towards the transition state with use of a climbing image algorithm [32, 33, 34].

3.3.2 Optimization of geometry

A local minimum on a PES is characterized by only positive eigenvalues of the Hessian matrix. To make sure that the geometry and energy of a molecule is minimized, one can perform a geometry optimization. The first step is to provide a guess geometry, where it is attempted to construct a molecule in such a way that it is already close to an optimized structure to speed up the actual optimization. With the Hessian matrix, with dimensions of $3N-6$, which consists out of the second derivatives of energy with respect to the atomic coordinates, the aim is to find a minimum where all the Hessian eigenvalues are larger than 0.

3.3.3 Transition states

To find a transition state, a method similar to geometry optimization is used. However, instead of finding only positive eigenvalues from the Hessian, as is the case for a minimum, for a transition state one negative eigenvalue is required. The negative eigenvalue is indicative of a saddle point, which describes a transition state. One provides an initial geometry close to a transition state, where the aim of a transition state search is to find a saddle point on the PES. In contrast to a geometry optimization, where all eigen values of the Hessian should be larger than 0, now one eigenvalue should be a negative value.

3.3.4 Intrinsic reaction coordinates

When a transition state is found, an additional step is done to ensure that the found transition state corresponds to the expected reaction from reactant to product. This is typically done with an intrinsic reaction coordinate (IRC) calculation. The intrinsic reaction coordinate method was developed by Fukui, where the minimum energy path (MEP) is determined via steepest descent with use of mass-weighted Cartesian coordinates [35, 36]. This process is done by first considering equation 6, which defines the tangent t by the differential equation of x (mass weighted Cartesian coordinates) and s (path length of the MEP), where the tangent is indicative if a structure is on a minimum, saddle point, or maximum [37]. This tangent t can also be defined in terms of g, which is the eigenvector corresponding to the negative eigenvalue of the transition state. From the transition state, equation 7 is the simplest method to go from geometry x_n to geometry x_{n+1} with step size Δs, which is a steepest descent method. To illustrate how the IRC calculation will lead to

both reactant and product, the eigenvector g_n which corresponds to the eigenvalue of the transition state for g_0, is taken both as negative and positive value (equation 8).

$$\frac{dx(s)}{ds} = -\frac{g}{\|g\|} = t \tag{6}$$

$$x_{n+1} = x_n + \Delta s t(x_n) \tag{7}$$

$$x_{n+1} = x_n - \frac{\Delta s}{2}\frac{g_n}{\|g_n\|} \tag{8}$$

3.4 Basis sets

A basis set is a set of basis functions, where the basis functions as linear combinations represent the molecular orbitals.

$$\phi_i = \sum_i c_{ri}\chi_r \tag{9}$$

In equation 9, it can be seen that the basis set is represented as a sum of basis functions, where the basis functions are denoted by χ_r. The basis functions are usually taken as molecular orbitals (MO), which can be represented as a linear combination of one or more Slater-type orbitals (STO). An STO in the polar coordinates r, θ, and φ is represented as shown in equation 9 [37]. In this equation, the STO contains a normalization factor N, and $Y_{l,m}$ represents the spherical harmonic functions.

$$\chi_{\xi,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{n-1}e^{-\xi r} \tag{10}$$

The downside to the use of STOs as basis functions is that they can become computationally very expensive due to the fact that the calculation of three- and four-center two-electron integrals cannot be performed analytically [37]. To overcome this, Gaussian type orbitals (GTO) have been introduced, where the equation can be seen in equation 11.

$$\chi_{\xi,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{2n-2-l}e^{-\xi r^2} \tag{11}$$

The downside to GTOs in comparison with STOs is the fact that GTOs fail to accurately

describe the wavefunction close to the nucleus, due to the fact that an GTO has a zero slope at the nucleus. Furthermore, the tail of a wavefunction described by a GTO falls off too rapidly, leading to poor description of the tail of a wavefunction. However, a similar accuracy for GTOs can be obtained in comparison with STOs by taking three times the number of GTOs with respect to an STO [37]. This increase of GTOs is compensated by the computational efficiency of its integrals. Basis sets have different classifications, where the simplest basis set is called a minimum or single zeta (SZ) basis set. This implies that the basis set only consists of a single basis function for an orbital: 1 basis function for a 1s orbital, 1 basis function for a 2s orbital, one set of 2p basis functions, and so on. The downside to an SZ basis set is that it is inflexible: it fails to accurately describe more diffuse electron distributions, such as for π bonds. To increase the accuracy of a basis set, the number of basis functions per orbital can be increased, and polarization functions can also be added.

3.5 Density functional theory

3.5.1 Theory

Density functional theory makes use of the fact that the ground-state energy and other properties of a molecule can be calculated from the ground-state electron probability density $\rho_0(x, y, z)$. The advantage of this method, is that the number of variables is brought down to only three spatial coordinates. A regular electronic wavefunction consists out of 4N variables, where N is the number of electrons, so for larger systems, the number of variables will be large, whereas for DFT, the number of variables will always be three. The main relationship in DFT is $E_0 = E_0[\rho_0]$, where the ground-state energy E_0 is a functional of ρ_0. This can further be defined as presented in equation 12 [38, 29].

$$E[\rho] = -\frac{\hbar^2}{2m_e}\sum_{i=1}^{n}\int \psi_i^*(r_1)\nabla_1^2\psi_i(r_1)dr_1 - j_0\sum_{I=1}^{N}\frac{Z_I}{r_{I1}}\rho(r_1)dr_1 + \frac{1}{2}j_0\int\frac{\rho(r_1)\rho(r_2)}{r_{12}}dr_1r_2 + E_{xc}[\rho] \quad (12)$$

In equation 12, the first part is the kinetic energy of the electrons, the second part is the attraction between nuclei and electrons, the third term is the repulsion between electrons, and the final term is the exchange-correlation energy, a functional of the electron density, which describes

non-classical electron interactions [39]. The exchange-correlation energy cannot be described with an analytical form, so it is approximated. Different methods have been proposed to approximate the exchange-correlation energy functional E_{xc}. In figure 10, the different types of approximations can be seen on the left-hand side of the ladder, where each step on the ladder closer to heaven has a higher accuracy.

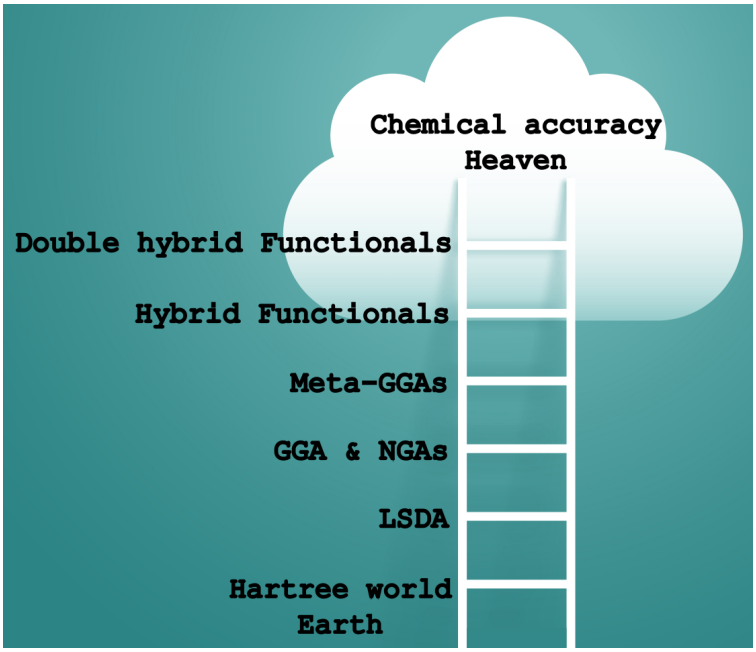


Figure 10: Ladder of DFT approximations to the exchange-correlation energy functional [40]

3.5.2 Functionals

3.5.2.1 Local Density Approximation

One of the first approximations for E_{XC} was developed by W. Kohn and L. J. Sham [41], where the definition can be seen in equation 13, which is known as the Local density approximation (LDA). This approximation is based on a homogeneous electron gas.

$$E_{xc}[\rho] = \int \rho(r)\epsilon_{xc}(\rho(r))dr \tag{13}$$

In their approximation, E_{xc} is dependent on ϵ_{xc}, where ϵ_{xc} is the exchange and correlation energy per electron in a uniform gas, which is a known value [42]. With this approximation, it was possible to calculate ground-state energies, free energies, and spin susceptibility for electronic systems [43]. A limitation of this method is that not all systems will behave like a homogeneous

electron gas, and as a result, the exchange energy is usually underestimated, and the binding energy overestimated, which is due to the fact that the gradient of the electron density is not considered, i.e. there is no gradient correction [44].

3.5.2.2 Generalized Gradient Approximations

An improvement upon LDAs, is the generalized gradient approximation (GGA), where its definition can be found in equation 14 [29].

$$E_{xc}^{GGA}[\rho^{\alpha}, \rho^{\beta}] = \int f(\rho^{\alpha}(r), \rho^{\beta}(r), \nabla\rho^{\alpha}(r), \nabla\rho^{\beta}(r))dr \tag{14}$$

In equation 14, f is a function of both the electron density and the gradient. The gradient allows for the description of a fluctuating electron density, in contrast to LDA, where a homogeneous electron gas is assumed. Usually, the exchange and correlation part are separated from each other as following: $E_{xc}^{GGA} = E_{x}^{GGA} + E_{c}^{GGA}$. Different types of exchange and correlation parameters can be combined with each other to create a GGA functional, where BLYP is an example of this, where the Becke 1988 exchange functional is combined with the Lee-Yang-Parr correlation functional [45, 46]. Limitations of GGA functionals is the ability to provide good atomization energies and also good performance for lattice parameters for solids [47]. To further improve the description of the electron density, the second derivative of the electron density is included, which results in meta-GGAs.

3.5.2.3 Hybrid functionals

Hybrid functionals, as the name implies, are a hybrid between correlation energy from DFT and exchange energy from Hartree-Fock theory. One of such functionals is the PBE0 functional, which consists out of the PBE GGA functional together with 25% exact exchange energy [48].

$$E_{xc}^{PBE0} = \frac{1}{4}E_{x}^{HF} + \frac{3}{4}E_{x}^{PBE} + E_{c}^{PBE} \tag{15}$$

The functional PBE0 is presented in equation 15, where the different percentages for exchange energy can be seen.

3.6 Density Functional based Tight Binding

Building upon DFT, Density Functional based Tight Binding (DFTB) is an approximation of DFT. DFTB only uses a minimal basis set, and only considers the valence electrons, furthermore, it neglects many integrals and uses approximations. DFTB itself is parameterized to values found from DFT calculations [37].

3.6.1 GFN1-xTB

GFN1-xTB is an abbreviation for Geometry, Frequency, Noncovalent, eXtended TB, developed by Grimme et al. [49] The parameters used in GFN1-xTB were determined by minimizing the root-mean-square-deviation between calculated and reference data with the Levenberg–Marquardt algorithm [50, 51]. Calculated data for parametrization were obtained with use of DFT functionals, PBEh-3c and PBE0-D3(BJ). Benchmarking of GFN1-xTB has shown that it has similar accuracy with respect to DFT functionals for the computation of structures. An advantage of this method, is the fact that the Hamiltonian contains physical parameters, which can be adjusted in a straightforward manner for other atoms. Furthermore, GFN1-xTB allows for efficient computations, due to its approximations. Shortcomings of GFN1-xTB include self-interaction errors, atomic orbital basis set deficiencies, and parametrization errors.

3.7 Many-Body Dispersion

To more precisely describe long range interactions between atoms and molecules, Many-body dispersion (MBD) is included. This allows for the description of Van-der-Waals effects [52]. In the MBD model, atoms are considered as a collection of three-dimensional isotropic quantum harmonic oscillators, where interactions or coupling between atoms are taken as dipole interactions between the oscillators. The MBD energy is calculated by taking the difference between the coupled and uncoupled oscillators. An improved version that builds upon MBD is MBD@rsSCS, where range separation (rs) and self-consistent screening (SCS) are included. With range separation, the electron-electron interactions are separated into short range- and long range interactions [52]. Short range interactions decay rapidly with distance, and are responsible for covalent bonding between atoms. Long range interactions on the other hand describe long range effects, such as Van-der-Waals interactions between molecules. Separation of these two types of interactions allows for

more efficient computations. Self-consistent screening (SCS) is included to account for the fact that electron-electron interactions between atoms and molecules depend on the electron density and vice versa, where SCS is an iterative method where the electron density is updated with respect to long-range interactions and vice versa. A bench-mark study has shown that including MBD@rsSCS together with the hybrid functional PBE0 allows for accurate computation of small molecular dimers and molecular crystals, however it falls short for the description of binding in supramolecular systems [52].

3.8 Implicit solvation

Including solvent effects can become quite expensive computationally if the solvent is treated explicitly. An implicit solvent model simplifies the description of a solvent by treating it as a continuous medium. In this continuous medium, it is assumed that the solvent behaves as a uniform polarizable medium with a dielectric constant [37]. One of such methods is the Generalized Born model [53], which is an approximate model of the Poisson-Boltzmann equations. The Poisson-Boltzmann equations are differential equations, which are numerically demanding.

$$\Delta G_{elec}(q) = -\left(1 - \frac{1}{\epsilon}\right)\frac{q^2}{2a} \tag{16}$$

In equation 16, the Generalized Born equation can be seen. In this equation, a spherical cavity is assumed in which a molecule is present, with q being the net charge, a being the radius of the cavity, and ϵ is the dielectric constant of the modelled implicit solvent [37]. The Generalized Born model is commonly found in combination with the Surface Area model, where the Surface Area model describes the solvation energy of an atom from an accessible surface area point of view.

$$G_elec(Q_i, Q_j) = -\left(1 - \frac{1}{\epsilon}\right)\frac{Q_i Q_j}{f_{ij}} \tag{17}$$

In equation 17, the Generalized Born model combined with the Surface Area model (GBSA) can be seen. In this model, the coulombic charges Q_i and Q_j, for the atoms i and j are included. Furthermore, the function f_{ij} depends on the internuclear distances and Born radii between two atoms i and j [37].

Another such method is the COSMO implicit solvation model [54], where originally it was

designed to be used in combination with *ab initio* methods. Similarly to the GBSA model, the COSMO model also makes use of cavities, and describes the electrostatic potential by partial atomic charges. An improvement on the COSMO model, is the COSMO-RS model, which can more accurately describe hydrogen bonding, furthermore it allows for more accurate descriptions of solvent mixtures [55].

Since implicit solvent models do not consider the solvent explicitly, there are a number of limitations. The description of tightly bound water molecules falls short, the behaviour of solvent molecules within a protein pocket falls short as well, since the solvent inside this pocket does not behave like the solvent outside of this pocket. Furthermore, solvent molecules can take part in reaction mechanisms by facilitating proton transfers for example, an implicit solvent model will fall short in describing this proton transfer accurately [56]. The COSMO-RS model also suffers from limitations, where it is known that it has difficulty accurately describing the interactions between secondary and tertiary amines with hydrogen-bonding solvents. It also falls short for chemical potential calculations, where the usual error is at least 0.8 kJ mol^{-1}. Furthermore, the description of larger networks with hydrogen bonding also falls short [57].

3.9 Activation strain model and energy decomposition analysis

The activation strain model is a method to investigate how a reaction progresses from reactant to product with use of an IRC. Within the activation strain model, the reaction strain and interaction between two fragments are investigated, where these two fragments react with each other. In equation 18, the PES $\Delta E(\xi)$ is divided into strain energy ($\Delta E_{strain}(\xi)$) and interaction energy ($\Delta E_{int}(\xi)$) along the reaction coordinate ξ. The strain energy is determined by how rigid bonds are that are broken, and deformation of bond angles. Different reaction mechanisms can also be determined based on the strain energy, as an example: an S_N2 reaction requires the breaking of one bond, whereas an E_2 reaction requires the breaking of two bonds, thus the strain energy of an E_2 reaction will be larger than that of an S_N2 reaction. Usually, the strain energy is positively valued. The interaction energy describes the interaction between the orbitals of two fragments. Upon a reaction, the interaction energy will decrease in energy, and is indicative of the stabilization between two fragments after a reaction [58, 59].

$$\Delta E(\xi) = \Delta E_{strain}(\xi) + \Delta E_{int}(\xi) \tag{18}$$

The internal energy $\Delta E_{int}(\xi)$ can further be separated with an energy decomposition analysis (EDA), as presented in equation 19. In this equation, $\Delta V_{elstat}(\xi)$ is the Coulomb interaction between the two fragments, $\Delta E_{Pauli}(\xi)$ is the Pauli repulsion between occupied orbitals present in the fragments, and $\Delta E_{oi}(\xi)$ is the orbital interaction between orbitals present in the two fragments together with polarization effects.

$$\Delta E_{int}(\xi) = \Delta V_{elstat}(\xi) + \Delta E_{Pauli}(\xi) + \Delta E_{oi}(\xi) \tag{19}$$

4 Methodology

All calculations were performed with use of AMS.2023 on the Hábrók High Performance Computing cluster (RUG), where the respective engines will be specified for each calculation performed. For the DFTB engine, the method GFN1-xTB was used. For the ADF engine, the DFT functional PBE0 was used in combination with the TZ2P basis set and Many Body Dispersion (MBD@rsSCS). For DFT, the following options were also used: a small frozen core, together with good numerical quality. Implicit solvent effects were modelled with use of the GBSA implicit solvent model in the DFTB engine and with the COSMO model in the ADF engine, where water was used as solvent. For optimizations and transition state calculations, frequencies were checked.

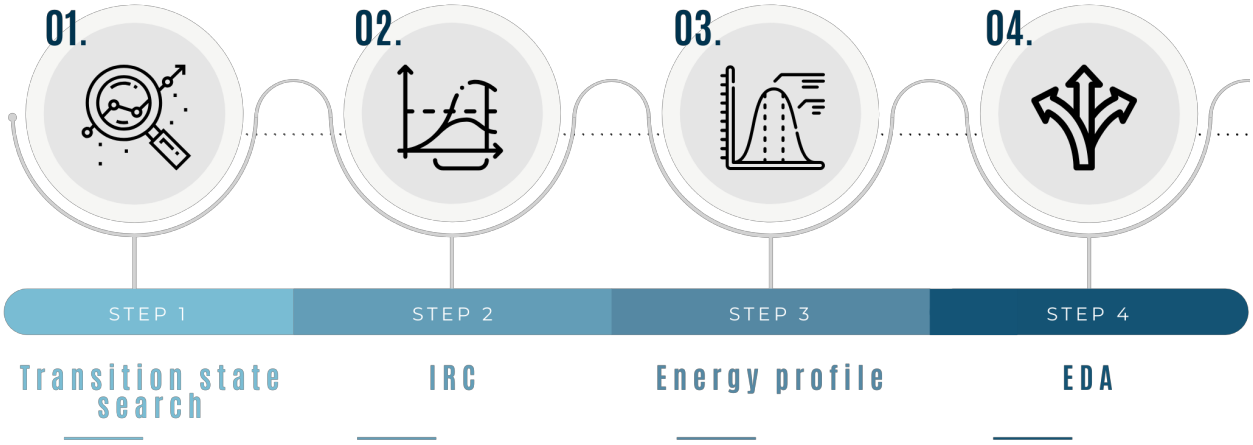


Figure 11: Workflow

4.1 Optimizations

Optimizations were performed with both the DFTB engine, and the ADF engine. Explicit water effects were modelled by adding water molecules to the input files, where the Force Field engine (UFF) was used to minimize the solvated system before using the DFTB engine or ADF engine for optimizations. Frequencies were checked to ensure a minimum was reached after optimization for each calculation.

4.2 Potential energy surface searches

Potential energy surface searches were performed with use of the DFTB engine via a saddle search to generate a set of potential transition states, which in turn could be used as input files for

transition state calculations. For this saddle search, 50 expeditions and 32 explorers were used.

4.3 Potential energy surface scans

Potential energy surface scans were performed with the use of the DFTB engine to find potential transition states. Bonds between atoms were chosen based on their expected elongation or shortening in a transition state.

4.4 Transition state calculations

Transition state calculations were performed with both the DFTB and ADF engine. The initial Hessian was calculated beforehand, and afterwards frequencies were checked for an imaginary frequency.

4.5 Intrinsic reaction coordinate calculations

IRC calculations were performed with both the DFTB and ADF engine. Optimized transition state geometries were used as input files for IRC calculations. Both forward and reverse paths were considered for IRC calculations. The start and end point of IRC calculations were used for optimizations, to compare energies between reactants, transition states, and products.

4.6 Activation strain model (ASM) and energy decomposition analysis (EDA)

The activation strain model and energy decomposition analysis were done with the PyFrag 2019 program [59]. The functional PBE0 was used, with the basis set TZ2P, and Many Body Dispersion.

5 Results

In this chapter, the results of both path A and path B will be given. For path B, transition states, IRCs, and energy decomposition analyses were performed in vacuum, implicit solvent, and explicit solvent systems to investigate the reactions. For path A, only a few transition states were found, which will be discussed in a later subchapter. Furthermore, an analysis will be given on the obtained data, together with a discussion regarding these data.

5.1 Path B

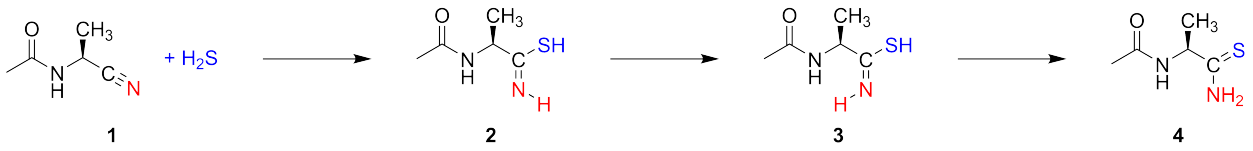


Figure 12: Reaction of Path B

In this chapter, the results obtained for Path B will be described. The three systems: vacuum, implicit solvent, and explicit solvent will be described here. Information on energetics will be given, where activation energies will be treated, together with reaction energies to give a comparison between the three systems previously described.

5.1.1 Vacuum

5.1.1.1 Optimizations

The three molecules out of the four: **1**, **2**, and **4**, as presented in figure 12, were first optimized with use of DFTB (GFN1-xTB). This ensured that for further calculations, minimized structures could be used. Optimizations of these three molecules did not lead to any problems, and lead to minimized structures, where frequencies were checked to ensure a minimum on the potential energy surface was reached. Initially only one intermediate was assumed to be present, however, after transition state searches, it was found out that there are two intermediate structures present, as can be seen in figure 12. The second intermediate structure **3** was found and taken from later IRC calculations and optimized with use of DFT (PBE0/TZ2P, MBD@rsSCS).

5.1.1.2 Potential energy surface explorations

To obtain a guess structure for a transition state calculation, structures **1**, **2**, and **4** (figure 12) were used in separate potential energy surface (PES) exploration calculations, where saddle searches were performed to find saddle points. These calculations gave multiple structures, which corresponded to saddle points on the PES. From these structures, only one structure was chosen per calculation which corresponded the closest to the assumed transition state. Since three input structures were provided for the PES exploration, three potential transition states were also taken out of these calculations.

5.1.1.3 Transition state calculations

From the PES explorations, three input structures were provided for a transition state calculation (TS). Initially, DFTB (GFN1-xTB) was used as a quick method to calculate these transition states. The structures were checked for imaginary frequencies to ensure that there was only one imaginary frequency, and that the frequency corresponded to an expected stretch. When there was indeed one negative frequency that corresponded to an expected stretch, the obtained structure was subsequently used for a transition state calculation with use of DFT (PBE0/TZ2P, MBD@rsSCS). A comparison between the transition states obtained with DFTB (GFN1-xTB) and DFT (PBE0/TZ2P, MBD@rsSCS) showed that the initial guess with DFTB was an accurate guess, since the transition states obtained with DFT showed the same stretches for the imaginary frequencies.

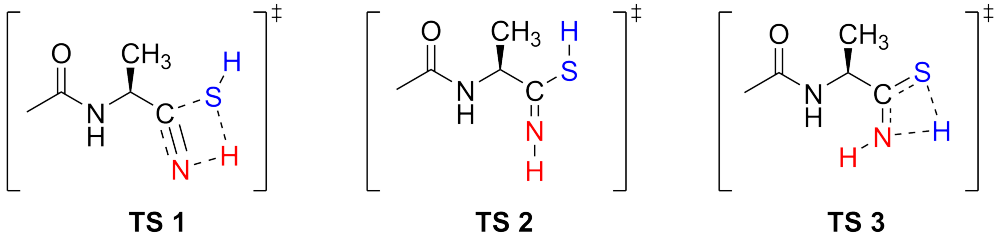


Figure 13: Transition states of Path B in a vacuum system

From these transition state calculations, the transition states **TS 1**, **TS 2**, and **TS 3** (figure 13) were found with both DFTB and DFT, with only one imaginary frequency for each. The transition states presented in figure 24 are interpretations of the imaginary frequency, where dotted lines represent bonds being broken and formed.

5.1.1.4 Intrinsic reaction coordinate calculations

To check whether the transition states were actually the correct transition states, intrinsic reaction coordinate calculations were performed at both DFTB and DFT level. If the correct transition state has been found, an intrinsic reaction coordinate calculation should lead to both the reactant and product from a transition state. In case of transition state **TS 1**, this should thus lead to both reactant **1** and product **2**, which was indeed the result as well for **TS 2** and **TS 3**. Thus it could be concluded from the intrinsic reaction coordinate calculations that the transition states **TS 1**, **TS 2**, and **TS 3** were the correct transition states. The IRCs calculated at DFTB level for transition states found with DFTB, did lead to the expected IRC graphs. Both the expected reactants and products were shown in the IRCs. Similarly, the IRCs calculated at DFT level from a transition state calculated at the same level of theory gave the same reactant and product. However, interchanging the level of theory between transition state calculations and IRC calculations did not lead to the expected reactant and product. This implies that DFTB (GFN1-xTB) is not interchangeable with DFT (PBE0, TZ2P, MBD@rsSCS), i.e. when a transition state has been calculated at a specific level of theory, this same level of theory should be used for an IRC. The theory behind this observation is most likely that the eigenvalue of the Hessian at the transition state calculated with DFT is not recognized as an eigenvalue of a transition state when DFTB is used for an IRC calculation, thus the potential energy surfaces of the methods are not the same. This was seen in the output files from the IRC calculations done at DFTB level, where the transition state was recognized as a non-stationary point on the PES. Thus the gradient of the transition state found with DFT is not considered to be zero by DFTB during an IRC calculation, where the IRC will stay on one side of the curve.

5.1.1.5 Energetics of the reaction

To accurately describe the activation energies and reaction energies of the reaction as shown in figure 4, the structures were all optimized with DFT, and their energies were then taken from those calculations. These energies were processed into a graph.

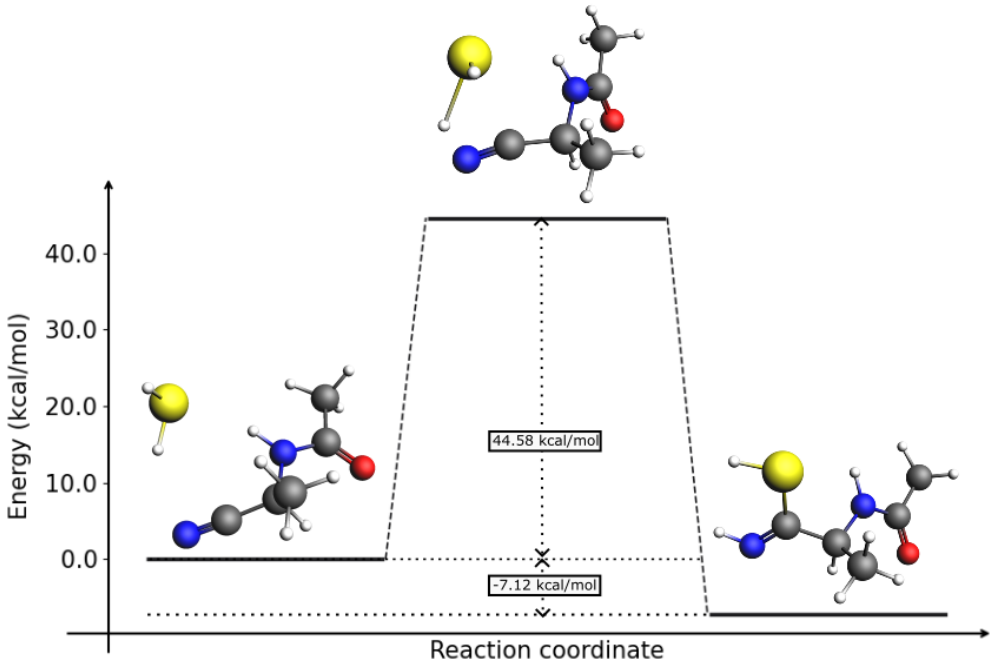


Figure 14: Energy profile for molecules **1**, **TS 1**, and **2**

From figure 14, it can be seen that there are three states, a reactant, a transition state, and a product. In this step, hydrogen sulfide performs a nucleophilic attack on the nitrile functionality, leading to an imidothioic acid. The approach angle for the hydrogen sulfide addition to the nitrile visible in the transition state is as expected. One of the lone pairs of electrons on sulfur is moving towards an anti-bonding p-orbital present on the nitrile. Subsequently, the lone pair of electrons on the nitrogen atom, which is a part of the nitrile, performs a nucleophilic attack on one of the protons present on hydrogen sulfide. It was expected that this addition would be the slowest out of the complete reaction pathway, since this is an intermolecular reaction, whereas the next reactions are intramolecular reactions. An activation energy of 44.58 kcal/mol was found from the reactant to the transition state. Furthermore, a reaction energy of -7.12 kcal/mol was found from reactant to product, making this reaction exothermic.

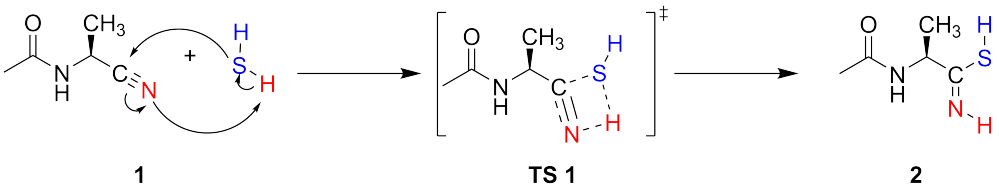


Figure 15: Electron pushing model for **1**, **TS 1**, and **2** based on the found transition state and IRC

In figure 15, an electron pushing model is proposed based on the found transition state and

IRC. In the first step, hydrogen sulfide performs a nucleophilic attack on the nitrile, which leads to transition state **TS 1**, as a square intermediate. Then upon addition and hydrogen transfer, imidothioc acid **2** is obtained.

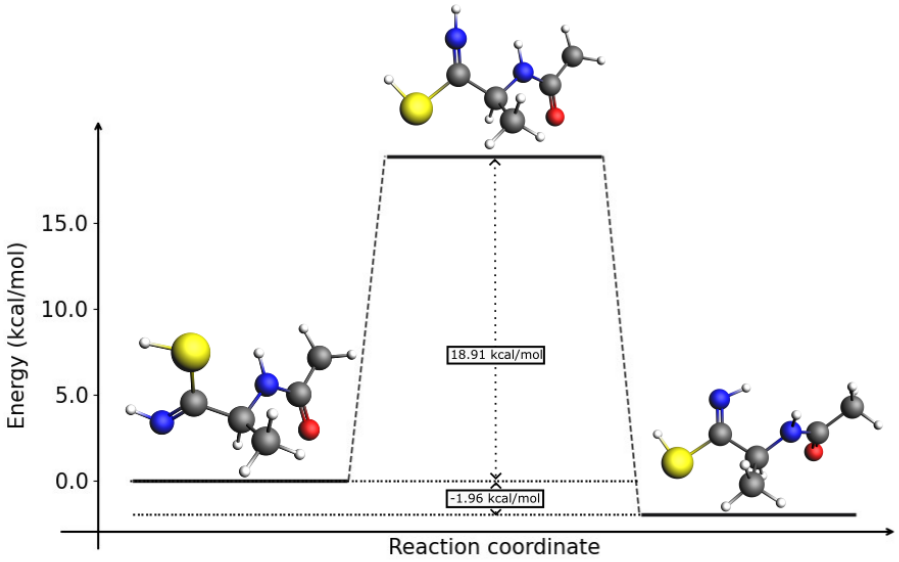


Figure 16: Energy profile for molecules **2**, **TS 2**, and **3**

In figure 16, the next step can be seen in a vacuum system. In this reaction, a proton flips from one side to the other side of the imine. This proton flip has an activation energy of 18.91 kcal/mol from reactant to the transition state. The overall reaction energy was found to be -1.96 kcal/mol, making this reaction slightly exothermic.

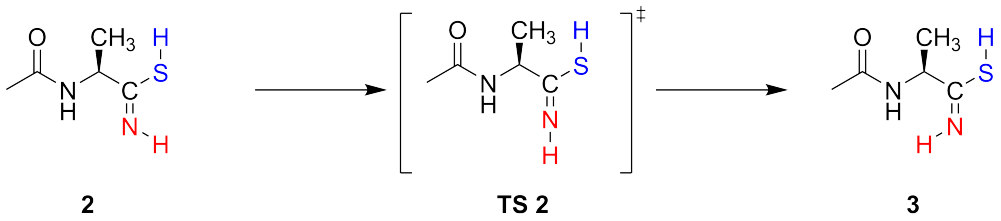


Figure 17: Electron pushing model for **2**, **TS 2**, and **3** based on the found transition state and IRC

In figure 17, the proposed mechanism for **2** to **3** can be seen via transition state **TS 2**, where the proton flip takes place. This step takes place to facilitate the proton transfer from the sulfur to the imine in the next step, without this proton flip, the lone pair on the imine is on the wrong side.

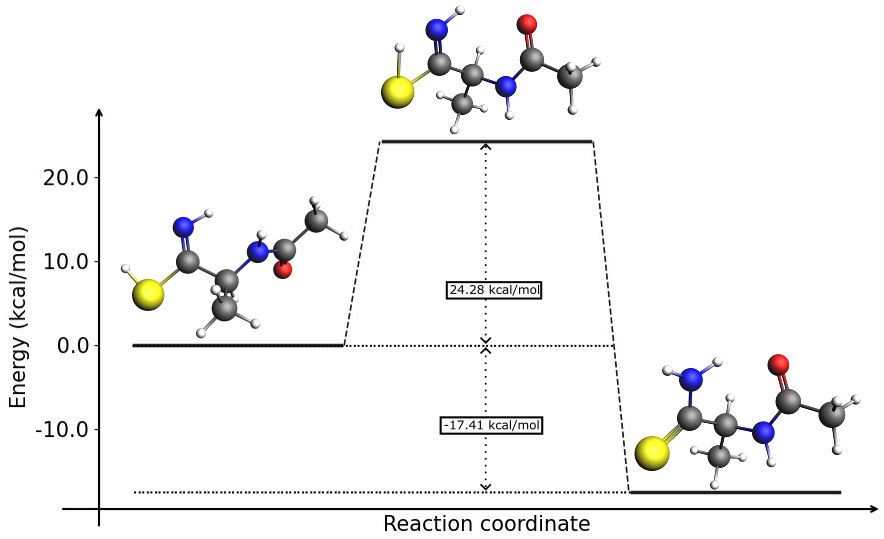


Figure 18: Energy profile for molecules **3**, **TS 3**, and **4**

From figure 18, it can be seen that from reactant to product, a proton is transferred from the thiol to the imine. An activation energy of 24.28 kcal/mol was found from reactant to transition state. Furthermore, a reaction energy of -17.41 kcal/mol was found, making this reaction an exothermic reaction.

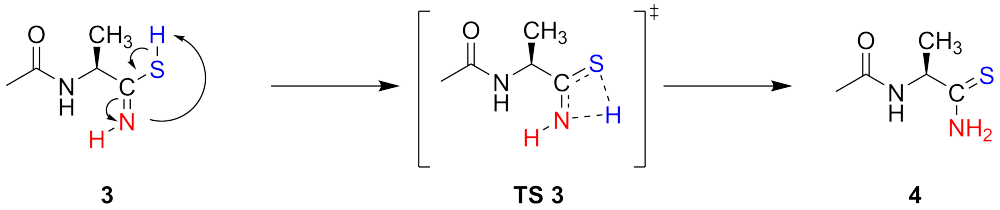


Figure 19: Electron pushing model for **3**, **TS 3**, and **4** based on the found transition state and IRC

In figure 28, the final step is shown via an electron pushing model. In the first step, the lone pair of the nitrogen attacks the proton present on the thiol, which leads to transition state **6**. From this transition state, the proton is transferred to the imine, which gives a thioamide.

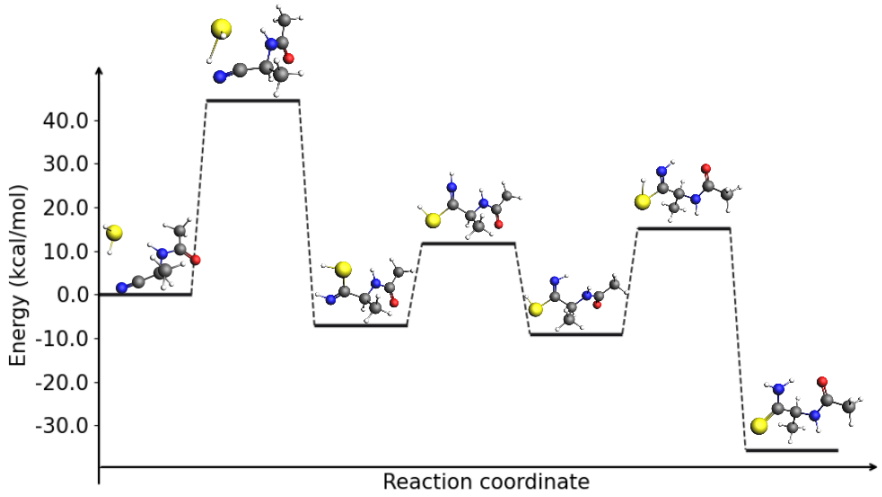


Figure 20: Energy profile for molecules **1** to **4**

In figure 20, the overall energetics can be seen for the seven steps. From the energetics, it is clear that the reaction in vacuum is an exothermic reaction, with an overall reaction energy of -26.49 kcal/mol. From figure 20, it can be seen that the first step requires the most activation energy, 44.58 kcal/mol, whereas the second and third step respectively require 18.91 kcal/mol and 24.28 kcal/mol. To investigate whether the transition states can be reached, the half life of a first order kinetic reaction is assumed. The half life is defined in equation 20, where k is the rate constant defined by the Eyring equation, which is defined in equation 21.

$$t_{1/2} = \frac{ln(2)}{k} \tag{20}$$

$$k = \frac{K_b T}{h} * \exp^{\frac{-\Delta G^{\ddagger}}{RT}} \tag{21}$$

With an energy barrier of 44.58 kcal/mol, at a temperature of 298.15K, a half life of $1.7 * 10^{12}$ years was found, implying that it is highly unlikely that this first step will take place at room temperature. If the temperature of the system is increased to 373.15K (boiling point of water at 1 atm.), where it is assumed that this reaction takes place in water at 1 atm., a half life of $3.6 * 10^5$ years is found, implying that this reaction barrier in vacuum requires too much energy to realistically be reached. The second step requires an activation energy of 18.91 kcal/mol, where a half life of 8 seconds was found for a temperature of 298.15K, with an increase of temperature

up to 373.15K, a half life of 0.01 seconds was found, implying that this second step will happen spontaneously. The third step has an activation energy of 24.28 kcal/mol, where at 298.15K, a half life of 19.4 hours was found, and at 373.15K, a half life of 14.8 seconds was found. Thus, at room temperature and an elevated temperature of 373.15K, it is unlikely that this reaction will take place in a vacuum system, since the first step, nucleophilic addition of hydrogen sulfide, requires 44.58 kcal/mol, which lead to a half life of at least $3.6 * 10^5$ years.

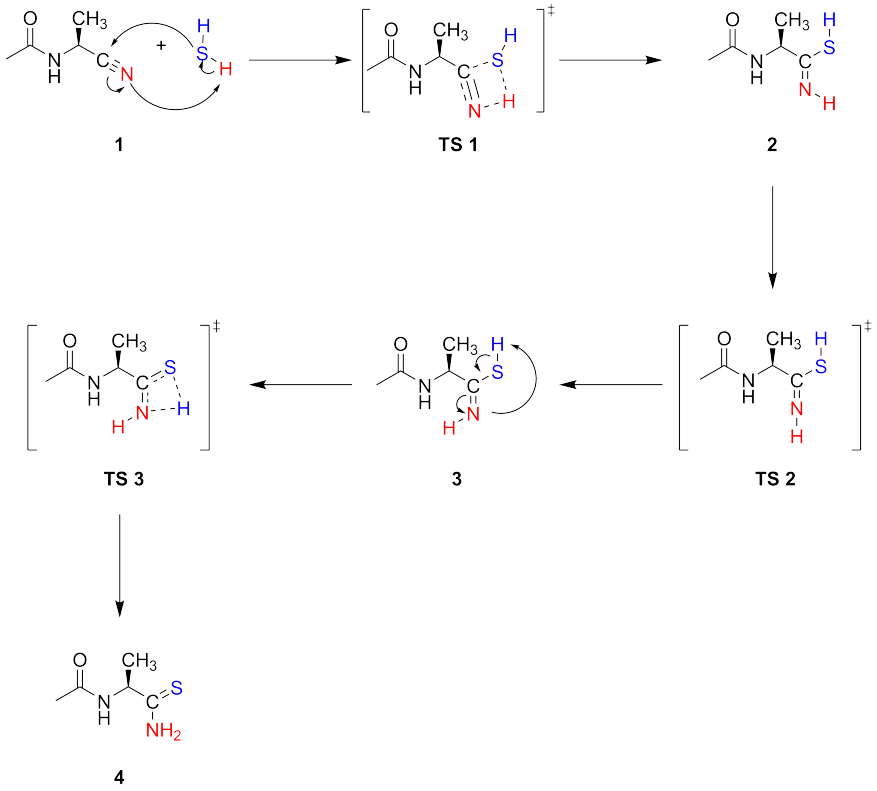


Figure 21: Proposed reaction mechanism with transition states included for Path B in a vacuum system

Putting all the pieces together, leads us to a complete electron pushing model in figure 21.

5.1.1.6 Activation Strain Model and Energy Decomposition Analysis

To further investigate the reactions as presented in figure 21, an ASM and EDA were performed on the IRC paths for **1** to **2**, **2** to **3**, and **3** to **4**. The fragments used for ASM and EDA can be seen in figure 22.

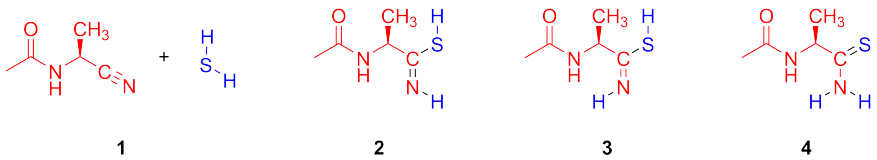


Figure 22: Fragments as used in the ASM, where the two fragments are distinguished based on the colors blue and red

From this, it was figured out which kind of forces dominate the reactant, transition state, and product. These forces are the interaction- and strain energy, where the interaction energy can be further decomposed into electrostatic-, orbital interaction-, Pauli-, and dispersion energy. Typically, electrostatic- and orbital interaction energies have negative values. To relate the electrostatic-, orbital interaction-, and Pauli energies to the total interaction energy, the absolute values were taken and converted to percentages relative to the total interaction energy as sum of the absolute values. Similarly, the strain energy was also taken as absolute value, from where the contribution of both the interaction- and strain energies could be calculated in percentages.

	$\mathbf{E}_{int}$	$\mathbf{E}_{strain}$
1	0	0
TS 1	55.8	44.2
2	50.9	49.1

Table 1: Contributions of the interaction- and strain energy towards the total energy in percentages from ASM for the molecules **1** to **2**

In table 1, it can be seen that the interaction energy and strain energy in percentages are very similar to each other for the addition of hydrogen sulfide to molecule **1** to get imidothioic acid **2**. It is important to notice that the interaction energy decreases in energy, and the strain interaction increases from **1** to **2**, which is not surprising, since the addition of hydrogen sulfide should indeed lead to an energy decrease in interaction energy, and an energy increase in strain energy. For transition state **TS 1**, the contribution of the interaction energy is larger than the contribution of the strain energy, which suggests that for the transition state, the interaction energy plays a more important role for bringing the two reactants together.

	$\mathbf{E}_{elstat}$	$\mathbf{E}_{oi}$	$\mathbf{E}_{Pauli}$	$\mathbf{E}_{disp}$
1	31.1	20.1	43.5	5.3
TS 1	22.0	24.7	52.8	0.5
2	14.7	40.8	44.4	0.1

Table 2: Contributions of electrostatic-, orbital interaction-, pauli-, and dispersion energy in percentages to the interaction energy from EDA for the molecules **1** to **2**

In table 2, the interaction energy has been decomposed into its components, where these components are presented in percentages. It is important to notice that the Pauli interaction energy increases in energy, whereas the orbital interaction-, electrostatic, and dispersion energy decrease in energy overall. It is clear that for transition state **TS 1**, the Pauli interaction energy dominates, meaning that the Pauli interaction energy is the dominant factor for the interaction energy. The Pauli interaction energy is a description of the repulsion between electrons with the same spin, in this case the electrons present in the orbitals of hydrogen sulfide and the orbitals on the nitrile of the amino acid precursor in **1**. Overall, it is not surprising that the orbital interaction energy increases, since a new bond is formed between both substrates present in **1**. Furthermore, the electrostatic energy contribution decreases, since there are now less hydrogen bonds present when comparing **1** with **2**. This is further demonstrated by a decrease in dispersion energy, since the amount of molecules decreases from 2 to 1.

	$\mathbf{E}_{int}$	$\mathbf{E}_{strain}$
2	0	0
TS 2	48.7	51.3
3	46.0	54.0

Table 3: Contributions of the interaction- and strain energy towards the total energy in percentages from ASM for the molecules **2** to **3**

The reaction from **2** to **3** is the proton flip of the imine. From table 3, it can be seen that the interaction- and strain energy contributions are very similar for transition state **TS 2**. It is important to notice that the interaction energy first increases in energy, before it decreases again after the transition state. This is not surprising, since the proton in transition state **TS 4** basically has an angle of 180°, when considering the angle between the proton, nitrogen, and carbon atom part of the imine.

	$\mathbf{E}_{elstat}$	$\mathbf{E}_{oi}$	$\mathbf{E}_{Pauli}$	$\mathbf{E}_{disp}$
2	14.7	40.8	44.4	0.1
TS 2	19.1	37.0	43.8	0.1
3	22.6	35.8	41.4	0.2

Table 4: Contributions of electrostatic-, orbital interaction-, pauli-, and dispersion energy in percentages to the interaction energy from EDA for the molecules **2** to **3**

Based on table 4, both the orbital interaction- and Pauli repulsion energy decrease in contribution from **2** to **3**, whereas the electrostatic interaction energy increases in contribution. The decrease in Pauli repulsion energy suggests that this proton flip leads to a more favorable position for the proton on the imine on **3** in comparison with **2**. The decrease in orbital interaction energy is most likely due to the fact that the proton in **2** could form a hydrogen bride with the sulfur atom, whereas in **3** this is no longer the case.

	$\mathbf{E}_{int}$	$\mathbf{E}_{strain}$
3	0.0	100.0
TS 3	3.6	96.4
4	53.7	46.3

Table 5: Contributions of the interaction- and strain energy towards the total energy in percentages from ASM for the molecules **3** to **4**

In table 5, the interaction- and strain energy contributions can be seen for **3** to **4**. For transition state **TS 3**, it is clear that the strain energy has a very big contribution of 96.4% on this reaction. This suggests that transition state **TS 3** is quite a strained system, which can be explained by its strained square geometry.

	$\mathbf{E}_{elstat}$	$\mathbf{E}_{oi}$	$\mathbf{E}_{Pauli}$	$\mathbf{E}_{disp}$
3	22.7	35.6	41.6	0.2
TS 3	20.3	35.7	43.9	0.1
4	15.3	40.5	44.1	0.1

Table 6: Contributions of electrostatic-, orbital interaction-, Pauli-, and dispersion energy in percentages to the interaction energy from EDA for the molecules **3** to **4**

The decomposition of the interaction energy in table 6 shows that the Pauli repulsion energy has the biggest contribution towards the interaction energy. This is most likely due to the nucleophilic

attack of the imine on the thiol, where a proton is transferred. During the nucleophilic attack, the lone pair on the imine will get closer to the proton and its bond with the sulfur atom, where electron-electron repulsion will take place.

5.1.2 Implicit solvent

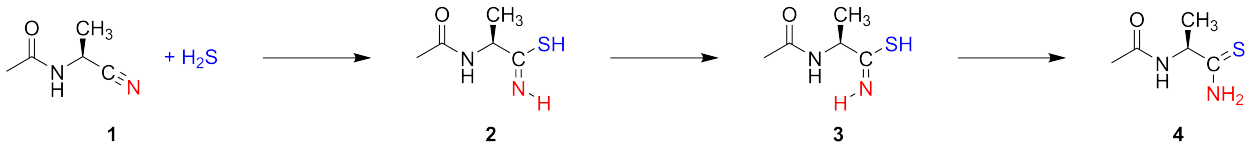


Figure 23: Reaction of Path B

5.1.2.1 Optimizations

Just as for the molecules in vacuum, the molecules present in 23 were first optimized with use of DFTB (GFN1-xTB), with water as implicit solvent. Input structures were taken from the vacuum system as initial guess structures. Optimizations of these three molecules did not result into any problems, and lead to minimized structures, where frequencies were checked to ensure a minimum on the potential energy surface was reached. Once again, it was found out after transition state searches, that there were four molecules in total, with two intermediates. The two intermediate structures **2** and **3**, as presented in 23 were taken from IRC calculations and optimized with use of DFT (PBE0/TZ2P, MBD@rsSCS).

5.1.2.2 Potential energy surface explorations

In contrast to the system in vacuum, no Potential energy surface explorations were done, since transition states could be taken from the system in vacuum.

5.1.2.3 Transition state calculations

From the system in vacuum, transition states could be taken as good initial guess, where an implicit solvent was added. Initially, DFTB (GFN1-xTB) was used as a quick method to calculate these transition states. The structures were checked for imaginary frequencies to ensure that there was only one imaginary frequency, and that the frequency corresponded to an expected stretch. When there was indeed one negative frequency that corresponded to an expected stretch, the obtained

structure was subsequently used for a transition state calculation with use of DFT (PBE0/TZ2P, MBD@rsSCS). A comparison between the transition states obtained with DFTB (GFN1-xTB) and DFT (PBE0/TZ2P, MBD@rsSCS) showed that the initial guess with DFTB was an accurate guess for two of the transition states, since the transition states obtained with DFT showed the same stretches for the imaginary frequencies.

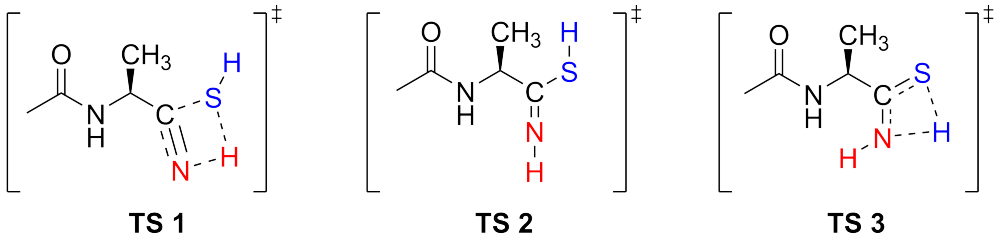


Figure 24: Transition states of Path B in the implicit solvent system

From these transition state calculations, the transition states **TS 2**, and **TS 3** (figure 24) were found with both DFTB and DFT, with only one imaginary frequency for each structure. The transition state **TS 1** appears to be tough to find, in contrast to the vacuum system. At DFTB level, a transition state similar to the transition state found for the vacuum system was found, with only one imaginary frequency present. However, when this same transition state search was done with DFT (PBE0, TZ2P, MBD@rsSCS), no transition state was found. The difference between the DFTB and DFT transition state calculations can be found in the different implicit solvation models used. Within DFTB, the solvation model GBSA was used, whereas for DFT, the solvation model COSMO was used. These are two different solvation models, where GBSA makes use of a polarized continuum.

5.1.2.4 Intrinsic reaction coordinate calculations

To check whether the transition states were actually the correct transition states, intrinsic reaction coordinate calculations were performed at DFTB level. From this it could be concluded from the intrinsic reaction coordinate calculations that the transition states **TS 2** and **TS 3** were the correct transition states.

5.1.2.5 Energetics of the reaction

To accurately describe the activation energies and reaction energies of the reaction as shown in figure 4, the structures were all optimized with DFT, and the energies were taken from those calculations. These energies were then processed into a graph.

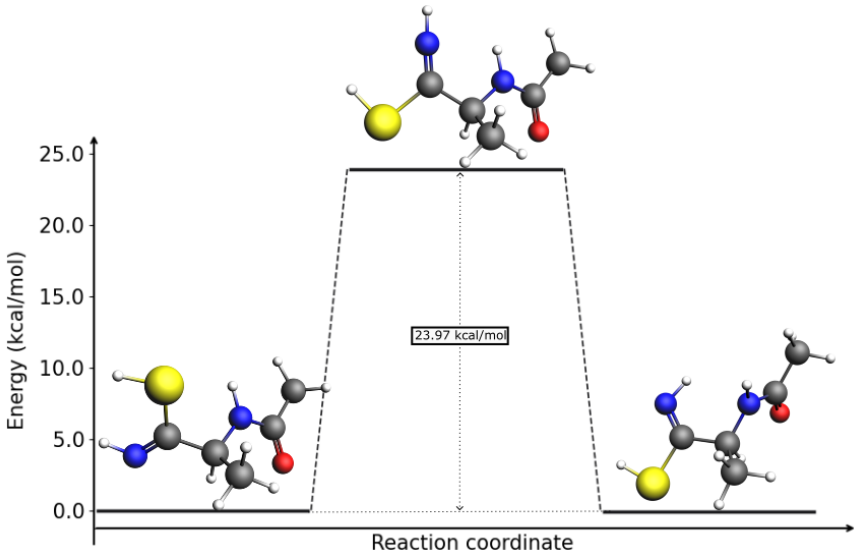


Figure 25: Reaction of Path B step 2, from **2** to **3**

In figure 25, the second step can be seen. From reactant to the transition state, an activation energy of 23.97 kcal/mol was found. A reaction energy of only 0.06 kcal/mol was found between reactant and product, implying that there is not much difference in internal energy for both reactant and product after a proton flip.

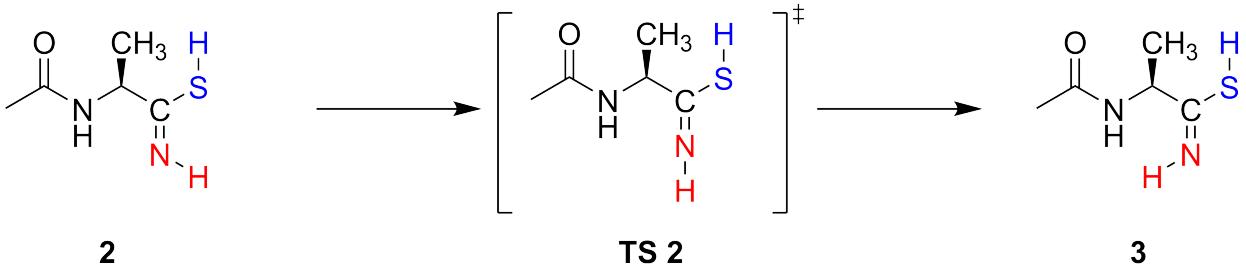


Figure 26: Electron pushing model for **2**, **TS 2**, and **3** based on the found transition state and IRC

In figure 26, the proposed reaction mechanism can be seen for this step. There are no electron pushing arrows visible, since the reaction from **2** to **3** is a proton shift on the imine.

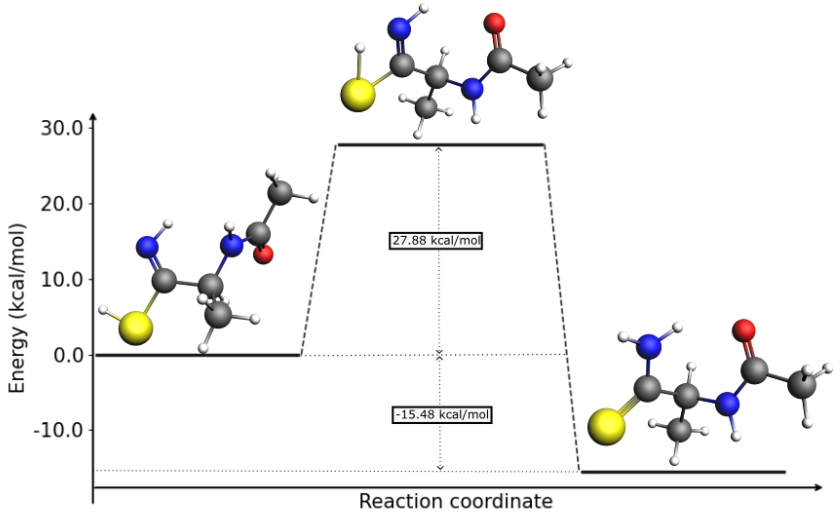


Figure 27: Reaction of Path B step 3. from **3** to **4**

In figure 27, the energetics of step 3 can be seen. From reactant to the transition state, an activation energy of 27.88 kcal/mol was found, and from reactant to product, a reaction energy of -15.48 kcal/mol was found.

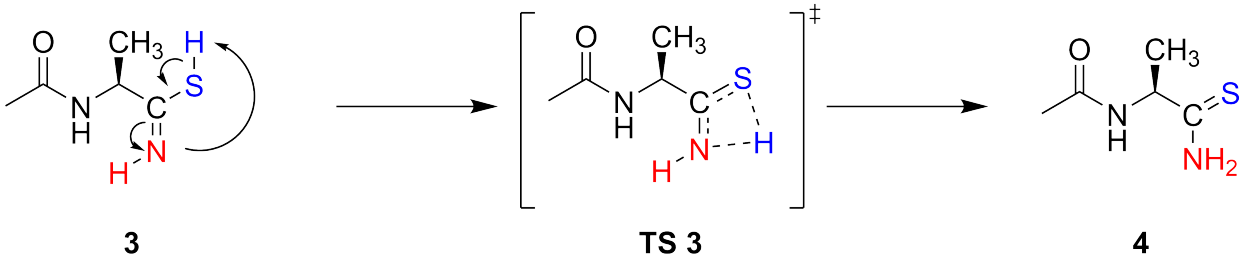


Figure 28: Electron pushing model for **3**, **TS 3**, and **4** based on the found transition state and IRC

In figure 28, the mechanism for the proton transfer from the thiol to the imine can be seen, which results in thioamide **4** from imidothioic acid **3**.

The energy barrier for the proton flip from structure **2** to structure **3** via transition state **TS 2**, was found to be 23.97 kcal/mol. This gives a half life of 11.5 hours at room temperature (298.15K). With an increased temperature to 373.15K, a half life of 9.7 seconds was found. This implies that the proton flip will most likely take place spontaneously at elevated temperatures up to the boiling point of water. The next step is the proton transfer from sulfur to the imine from structures **3** to **4** via transition state **TS 3** with an energy barrier of 27.88 kcal/mol. At room temperature, a half life of 353 days was found. At an elevated temperature of 373.15K, a half life of 32 minutes was

found. This implies that this reaction is unlikely to take place at room temperature, but it will take place at temperatures close to the boiling point of water.

Since it is not yet clear whether the first step will take place, due to difficulty of finding the relevant transition state, no complete reaction mechanism can yet be given.

5.1.3 Explicit solvent

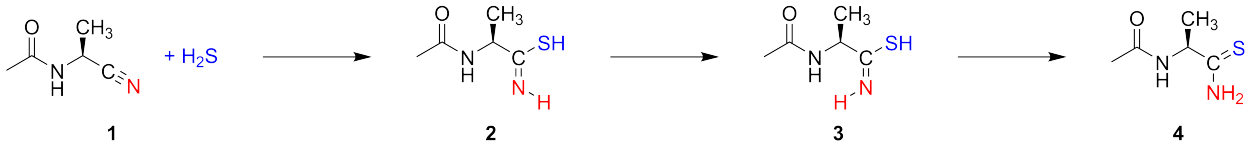


Figure 29: Reaction of Path B

Before an explicit solvent system of more than 50 water molecules was considered, only four water molecules were placed around the molecules presented in figure 29. This was done to make it easier to find transition states, where water would play a role for proton transfers. If an explicit solvent system of water was used directly, it would take more time to find these transition states, since larger systems will take a longer time to compute. Two intermediate structures were assumed, since it was not clear yet how water would influence the transition states, and if it would decrease or increase this number.

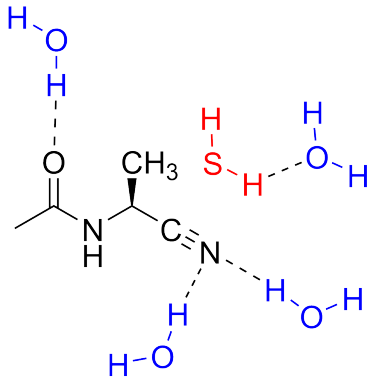


Figure 30: Initial orientation of water molecules around **start**

In figure 30, a graphical representation is given to show how water molecules were placed. It was chosen to place the water molecules in such a way that hydrogen bonding was possible, indicated by dotted lines.

5.1.3.1 Optimizations

The molecules, as presented in 29, could be taken as optimized structures from the vacuum system. Then, four water molecules were placed around them in such a way that hydrogen bonding was possible. After which the structures were optimized using DFTB (GFN1-xTB), followed up by an optimization using DFT (PBE0, TZ2P, MBD@rsSCS). This lead to optimized structures, where indeed hydrogen bonding between the water molecules and the substrates was observed by the orientation of the water molecules towards the substrates and the distances. This gave optimized structures with four water molecules around them, which were then used for transition state calculations. After IRC calculations (DFT), the reactant, intermediate, and product were optimized again to ensure that the four water molecules were orientated in such a way around the substrates, that it related to the minima corresponding to the transition states.

5.1.3.2 Transition states

Finding transition states appeared to be a bit harder in comparison with the vacuum system without any additional water molecules. Water molecules had to be orientated in such a way around the substrates, that proton transfers could take place. Finding transition states was further facilitated by potential energy surface scans, where protons which were assumed to be moving in a transition state, were investigated on bond lengths and energies.

5.1.3.3 Intrinsic reaction coordinates

Intrinsic reaction coordinates were calculated to ensure that transition states lead to the correct reactant and product. This was done with both DFTB (GFN1-xTB) and DFT (PBE0, TZ2P, MBD@rsSCS). Once again, it was observed that transition states calculated by DFT, followed by an IRC with DFTB, did not lead to the expected IRC.

5.1.3.4 Energetics of the reaction

The energetics of the reaction are described in this part, where both the activation- and reaction energy are of interest.

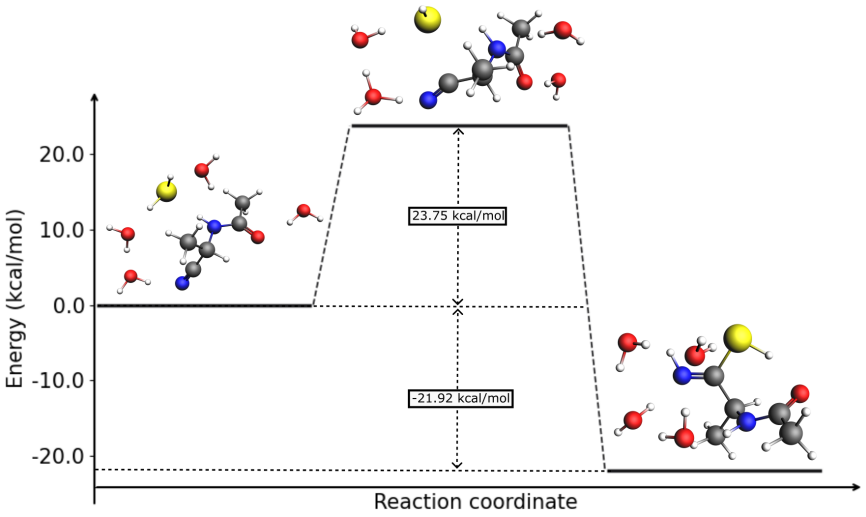


Figure 31: Reaction of Path B step 1, from **1** to **5**

The reaction from **1** to imidothioic acid **5** can be seen in 31, where both the activation- and reaction energy can be seen. From reactant **1** to transition state **TS 4**, an activation energy of 23.75 kcal/mol was found. A reaction energy of -21.92 kcal/mol was found from **1** to **5**, making this reaction exothermic.

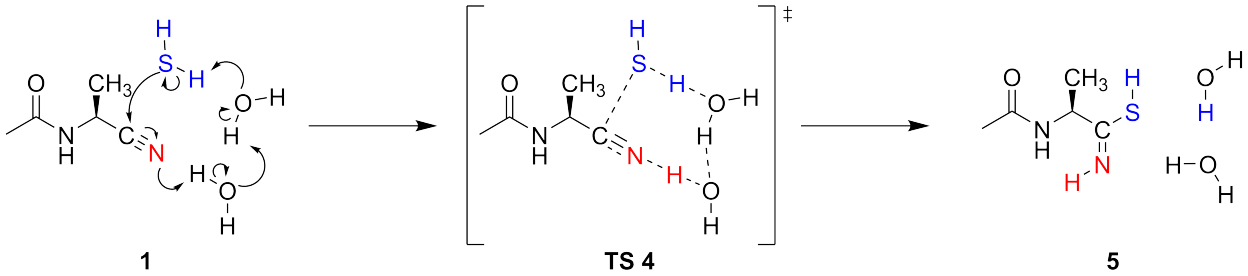


Figure 32: Reaction of Path B step 1

Based on the found transition state **TS 4** for the reaction from **1** to **5**, the mechanism in 32 is proposed. From the transition state, it is clear that only two out of the four water molecules actually take part in the reaction mechanism by facilitating proton transfers from hydrogen sulfide to the nitrile, where this leads to an octagonal transition state. Both the proton abstraction from the hydrogen sulfide and proton addition to the nitrile happen simultaneously.

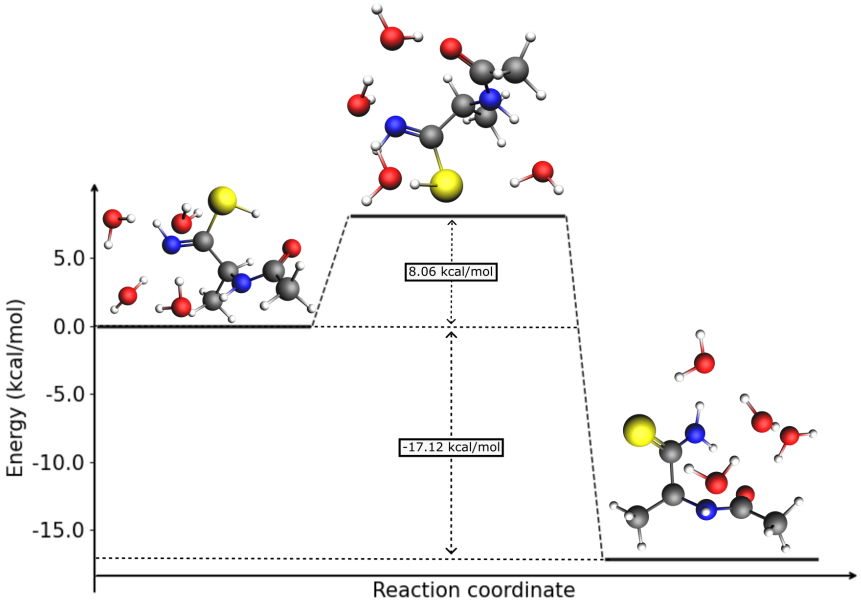


Figure 33: Reaction of Path B step 2

The second step, the tautomerization from imidothioic acid **5** to thioamide **6**, can be seen in 33. An activation energy of 8.06 kcal/mol was found and a reaction energy of -17.12 kcal/mol was found. This low activation energy barrier suggests that this reaction will happen spontaneously, furthermore, it is an exothermic reaction.

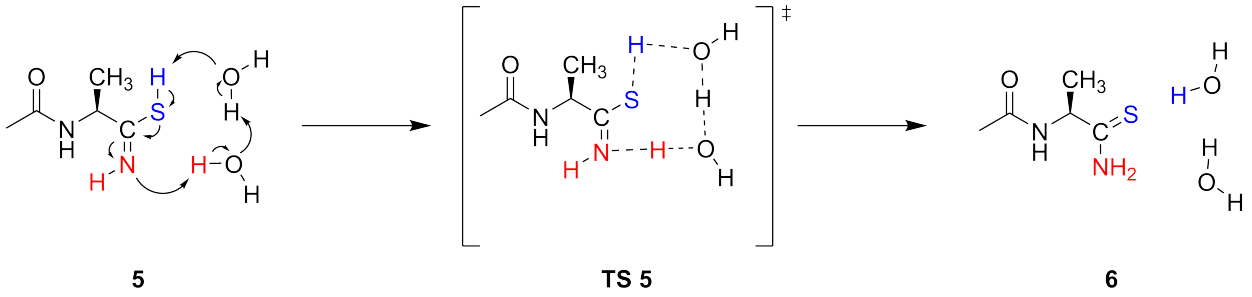


Figure 34: Reaction of Path B step 2

From transition state **TS 5**, reactant **5**, and product **6**, the reaction mechanism in figure 34 is proposed. Once again, only two water molecules actually take part in the proton transfer. A similar transition state **TS5** to the previous step **TS 4** was found, where two water molecules aid the proton transfer. A proton is abstracted from the sulfur atom, and a proton addition to the imine takes place simultaneously.

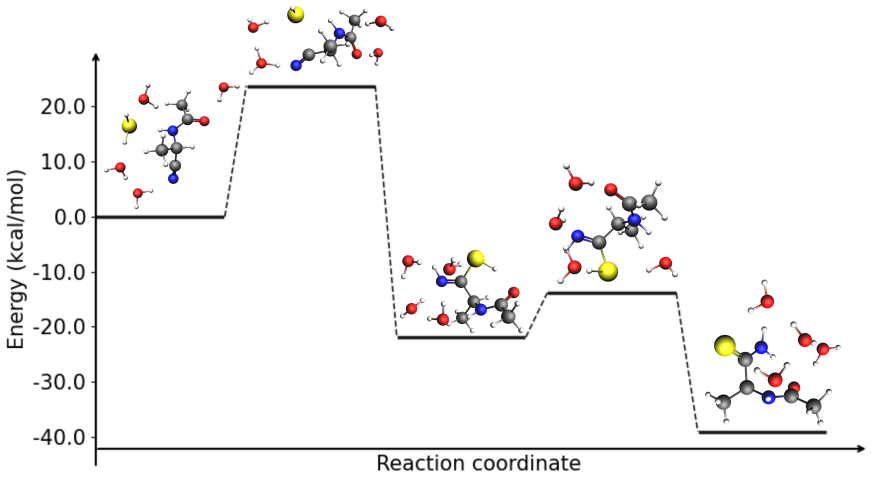


Figure 35: Energy profile for molecules **1** to **6**

From reactant **1** to transition state **TS 4** required an activation energy of 23.75 kcal/mol. At a temperature of 298K, this gives a half life of 8.11 hours. Where with a temperature of 373K, a half life of 7.34 seconds is obtained. This suggests that the addition of hydrogen sulfide aided with explicit water will take place at room temperature, and will spontaneously take place at the boiling point of water. For the second step, from **5** to transition state **TS 6**, a half life of 90.9 nanoseconds is obtained at 298K, and at 373K, a half life of 4.7 nanoseconds is obtained. This suggests that the proton inversion will spontaneously take place at room temperature. In comparison with the vacuum system, where an activation barrier of 44.58 kcal/mol was found, it is clear that the addition of water decreases the activation energy drastically, by more than 20 kcal/mol. Furthermore, inversion of the proton on the imine is not required, since the proton transfer aided by water already leads to the desired structure where the tautomerization of imidothioic acid takes place. Thus, the total number of steps is decreased from three steps to two steps. The addition of more than two water does not seem to lead to different transition states, this is demonstrated by the fact that four water molecules were included explicitly, and in each transition state, only two water molecules took part in the proton transfer.

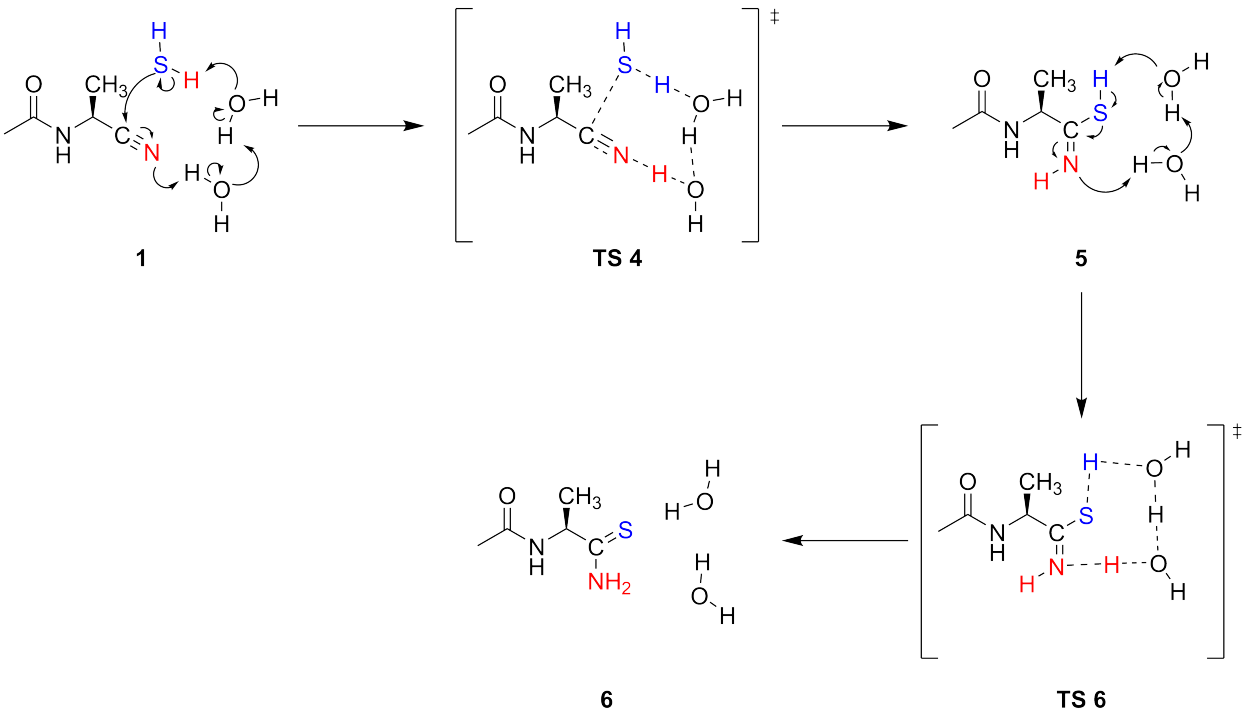


Figure 36: Proposed reaction pathway of Path B with 2 explicit water molecules taking part in the reaction

Combining all the data, leads to the proposed mechanism in figure 36. Only two water molecules have been drawn, since the other two water molecules do not take part in the mechanism. In the first step, the addition of hydrogen sulfide to the nitrile takes place, which leads to transition state **TS 4**. Upon proton transfers from hydrogen sulfide to the nitrile, imidothioic acid **5** is obtained. This subsequently tautomerizes to **6** via transition state **TS 6**, where once again proton transfer is aided by water molecules.

5.2 Path A

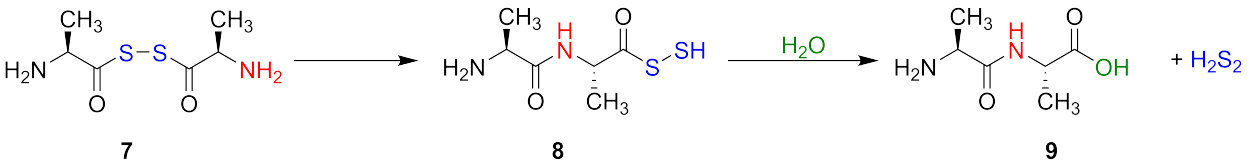


Figure 37: Reaction of Path A

In this subchapter, the results obtained for Path A will be treated, where 3 subsections can be found, which correspond to the three systems investigated, namely: vacuum, implicit solvent, and explicit solvent.

5.2.1 Vacuum

5.2.1.1 Optimizations

Optimization of the three structures found in figure 1 was done with DFTB, GFN1-xTB, which resulted in optimized structures, which in turn were checked for frequencies to ensure a minimum was found.

5.2.1.2 PES exploration

PES explorations for the optimized structures were done to find structures close to an expected transition state.

5.2.1.3 Transition states

Finding transition states in a vacuum system appeared to be tough. The transition state from intermediate to product was found, with nucleophilic addition of water to the dithioperoxy acid. The transition state from start to intermediate was not found, where the *S*-to-*N*-Acyl transfer takes place. The transition states that were found with an imaginary frequency from **7** to **8** would simply show a rocking mode of the proton attached to the sulfur atom.

5.2.1.4 IRCs

The intrinsic reaction coordinate from intermediate to product was calculated to ensure that the correct transition state was found, which indeed was confirmed to be the case. Since no other transition states were found, the energetics of the reaction pathway were not investigated.

5.2.2 Implicit solvent

5.2.2.1 Optimizations

Optimized structures could be taken from the vacuum system, and the system was re-optimized with implicit solvation accounted for.

5.2.2.2 Transition states

Similarly to the vacuum system, only the transition state from intermediate to product was found. During the attempt of finding a transition state from start to intermediate, it was observed that the transition states that were found, had an imaginary frequency of a proton transfer that would supposedly take place after the Acyl transfer.

5.2.2.3 IRCs

The intrinsic reaction coordinate from intermediate to product was calculated to ensure that the correct transition state was found, which indeed was confirmed to be the case. Since no other transition states were found, the energetics of the reaction pathway were not investigated.

5.2.3 Explicit solvent

5.2.3.1 Optimizations

Similarly to the strategy applied for Path B, the initial optimizations were done with 4 explicit water molecules included in a vacuum system. The starting geometries were taken as geometries optimized in a vacuum system without any explicit water molecules.

5.2.3.2 Transition states

Once again, finding transition states appeared to be tough. Interestingly, the transition state from **8** to **9** could be found using DFTB (GF1N-xTB), however, when DFT (PBE0, TZ2P, MBD@rsSCS) was used, it could not be found. Since this transition state could not be found with DFT, subsequent IRCs were not taken. It was expected that the transition states with four explicit water molecules would look different to the transition states found in vacuum and implicit solvent, however with the absence of those transition states, no comparison can be given.

6 Discussion

6.1 Path B

6.1.1 Vacuum

From TS searches and IRC calculations, the mechanism as proposed in figure 38 was made. Three transition states are proposed from nitrile **1** to thioamide **4** in a vacuum system.

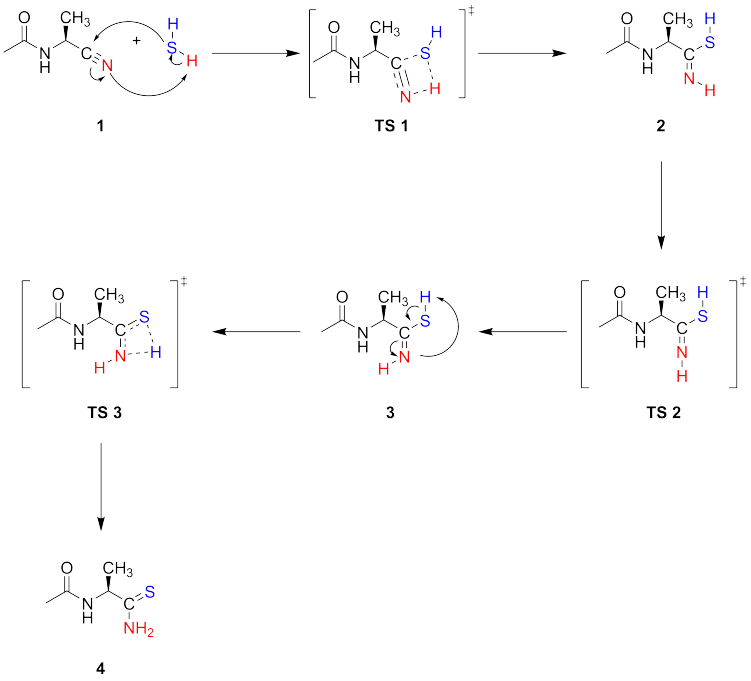


Figure 38: Proposed reaction mechanism with transition states included for Path B in a vacuum system

In the original experiment of Canavelli et al. [14], the thiolysis of nitrile **1** with hydrogen sulfide took place in an aqueous environment with a pH of 9. Furthermore, in their thiolysis, sodium hydrosulfide was used instead of hydrogen sulfide gas, where the hydrosulfide anion is a stronger nucleophile than hydrogen sulfide gas. This suggests that the vacuum system does not correlate to the original experiment [14], since a system in vacuum was considered, with no solvent effects, and thus also no pH. If hydrogen sulfide gas was used in the experiment of Canavelli et al. [14], the thiolysis from **1** to **2** still would not be correct, since at a pH of 9, hydrogen sulfide would be deprotonated and exist as the hydrosulfide anion, since hydrogen sulfide has a pKa of 7 [60]. However, modelling this reaction in vacuum without any solvent effects does provide valuable information about the energetics between each structure and transition state, which in turn are

used to compare with energetics obtained from a QM/MM system. The importance of a pH of 9 and water as solvent are demonstrated in the first step, where the thiolysis of **1** to **2** takes place. An energy barrier of 44.58 kcal/mol was found, suggesting that the addition of hydrogen sulfide in this system does not take place. In other words, hydrogen sulfide is not nucleophilic enough, and deprotonation of hydrogen sulfide is required to create a strong enough nucleophile for the thiolysis. Thus, this high energy barrier leads to the conclusion that some form of aid is required, be it solvent effects or a stronger nucleophile, for this thiolysis to lower this activation energy.

From a chemical point of view, transition state **TS 1** does make sense. The alternative would be a non-concerted transition state, where hydrogen sulfide would add to the nitrile, but the proton transfer would happen in a subsequent step. However, this would lead to a negatively charged nitrogen atom, which has a high basicity, and would be quite unstable in a vacuum system without any stabilization from a solvent. In the absence of water, it would indeed make the most sense that concerted transition state **TS 1** would take place.

The proton flip in transition state **TS 2** can be described as either rotation or inversion of the C=N imine bond. To distinguish between a rotation or inversion mechanism, one can investigate the C=N-H bond angle during the transition state.

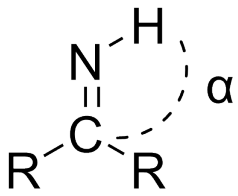


Figure 39: Angle α of the C=N-H bond

In figure 39, the bond angle of C=N-H is represented by α, where the bond angle at the transition state is the most relevant. If the bond angle α is close to 180°, the mechanism is described as inversion, whereas a bond angle closer to 113° describes a rotation mechanism [61]. This bond angle was found to be 175.6°, indicating that this transition state goes via an inversion mechanism.

In transition state **TS 3**, the tautomerization of imidothioic acid **3** takes place. This tautomerization has been investigated with a computational study by Würmel et al. [62], where it was found out that an imidothioic acid with C_2H_5 as R-group has an activation energy of 24.86 kcal/mol (MP2/6-311++G(2d,p)), where at room temperature, this tautomerization will take place

spontaneously. These values are in agreement with what was found in this study.

6.1.2 Implicit solvent

With the inclusion of implicit solvent effects, where water is simulated, it is demonstrated how water plays a role explicitly in the reaction mechanism of Path B. The structures and transition states found are very similar to the ones presented in figure 38. However, transition state **TS 1** appeared to be problematic to find. With DFTB, this transition state was found, and this was confirmed with an IRC calculation. When DFT was used on the other hand for this same transition state calculation, no transition state could be found which resembles **TS 1**. The functional PBE0 is known to perform well for finding transition states [63], implying that either the difference in implicit solvent model is the cause or the difference in method used. It has been demonstrated that GFN1-xTB has shortcomings when it comes to calculating geometries accurately [49, 64], which can explain the discrepancy between DFTB and DFT for finding transition state **2**. Additionally, the PES obtained with DFTB is different from the PES obtained with DFT, which was noticed with IRC calculations on transition states. However, in the vacuum system, transition state **TS 1** was found for both DFTB and DFT, suggesting that this difference is due to the different implicit solvent models. It was observed that upon a transition state calculation with DFT, hydrogen sulfide would move away from the nitrile functionality, whereas with DFTB, a transition state with the expected imaginary eigenvalue was obtained. So apparently the combination of PBE0, COSMO leads to different results in comparison with GFN1-xTB, GBSA.

6.1.3 Explicit solvent

The addition of four water molecules lead to different results with respect to the previous systems: vacuum, and implicit solvation. The number of transition states was decreased by one, since the imine inversion did not have to take place for the tautomerization of the imidothioic acid. From the transition states, it was observed that two out of the four water molecules take place in the mechanism, thus it is expected that the addition of more water molecules will not lead to significant differences in the mechanisms.

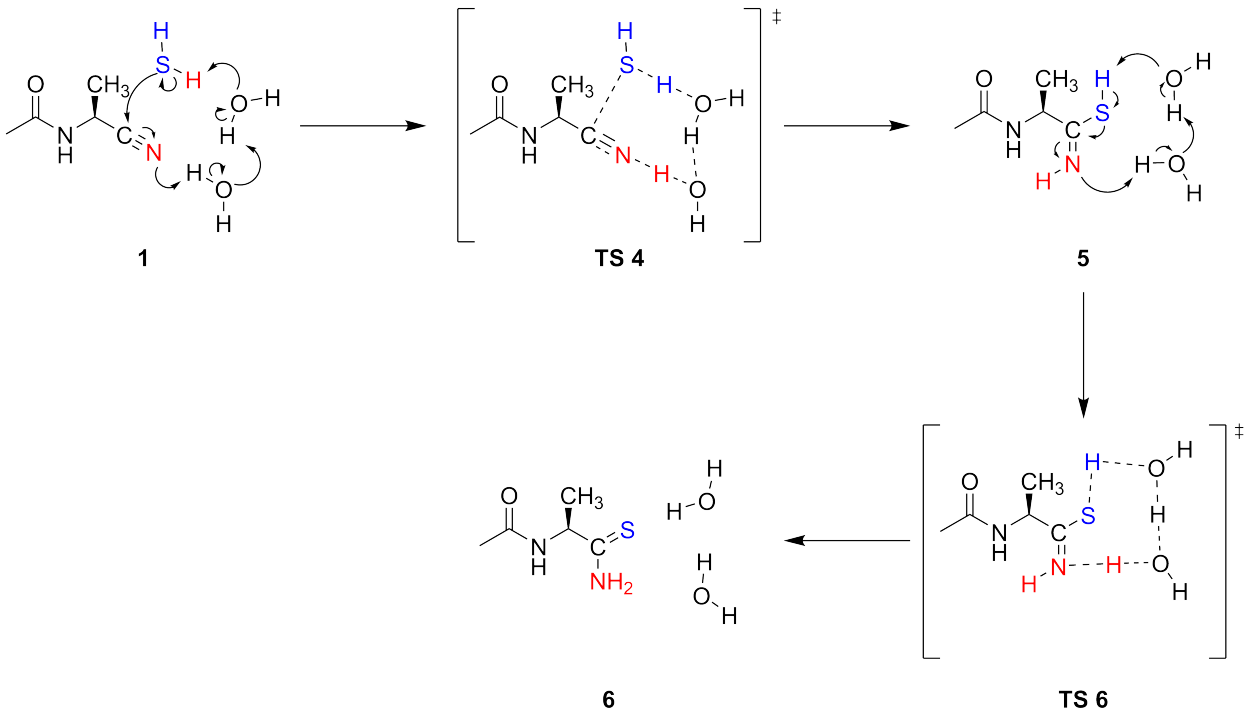


Figure 40: Proposed reaction pathway of Path B with 2 explicit water molecules taking part in the reaction

These two water molecules do facilitate the proton transfers (figure 40) in such a way that the activation energy of the hydrogen sulfide addition in **1** is decreased by 20 kcal/mol with respect to the vacuum system. Furthermore, the tautomerization of the imidothioic acid **5** readily takes place with the aid of two water molecules, with an activation energy of 8.06 kcal/mol. These activation energies are realistic at both room temperature (298K) and the boiling point of water (373K) to be overcome. This demonstrates that water as an explicit solvent is of great importance for this reaction path.

The conclusion that water promotes the tautomerization of imidothioic acid is shared by Freeman et al. [65], where in the gas phase, activation energies of roughly 30 kcal/mol were found. The addition of one and two water molecules lead to a decrease of 20 kcal/mol of the activation energy, where this observation is shared in this thesis [65].

It was observed that only two water molecules out of the four take part in the found transition states, however, the other two water molecules did form hydrogen bonds with the carbonyl oxygen and secondary amine. It is expected that the addition of more water molecules will lead to the formation of more hydrogen bonds. In a study by Y. Mo [66], it was observed in a QM/MM simulation with an imidothioic acid placed in a water box of 750 molecules, that three to five

water molecules will form hydrogen bonds with the sulfur atom, and two water molecules will form hydrogen bonds with the imine. Thus, it is expected that in a QM/MM system, more hydrogen bonds will form, but no significant changes to the mechanism will happen.

Typically, the thiolysis of a nitrile with hydrogen sulfide takes place in a basic medium [67], which is also the case for the experiment performed by Canavelli et al. [14]. The basic medium ensures deprotonation of hydrogen sulfide, which transforms it into a stronger nucleophile. Similarly to the vacuum- and implicit solvent systems, pH was not accounted for, and instead of hydrogen sulfide, sodium hydrosulfide was used as nucleophile [14]. This implicates that to accurately model the reaction, sodium hydrosulfide should be used as nucleophile, and a pH of 9 should be modeled as well by adding hydroxide ions to the explicit water molecules.

In this research, the focus was on the precursor of the amino acid L-Alanine. Thus it can only be concluded for L-Alanine that the thiolysis and subsequent imidothioic acid tautomerization of the L-Alanine precursor is enhanced by the presence of water due to a lowering of activation energies. However, it is likely that other amino acids will have very similar albeit similar mechanisms for thiolysis and tautomerization. Other amino acid precursors were used for thiolysis as well in the experiments of Canavelli et al. [14], however, it was chosen to investigate L-Alanine since it already incorporates stereochemistry in comparison to Glycine, yet does not introduce any sterically or electronically influential groups.

In the three systems investigated, it was observed that there is a difference in the PES between DFT and DFTB. The same level of theory should be used when calculating an IRC from a transition state. Interestingly, in literature there are examples where single point calculations at a higher level of theory are performed on a DFTB optimized geometry. Investigated systems range from organic molecules to perovskites, where different properties were investigated [68, 69, 70, 71, 72]. With the observation that the PES is different between DFT and DFTB, one wonders how justified and accurate a single point energy calculation with a higher level of theory is on a DFTB optimized geometry. In this case, it would be better to optimize the geometry of a molecule with DFTB initially to obtain a good guess geometry, and reoptimize this geometry at the desired level of theory, e.g. with a DFT functional.

6.2 Path A

Path A appeared to be tough (figure 41). It was difficult to find transition states for the first step, the *S*-to-*N*-Acyl transfer. Neither in the vacuum system nor implicit solvent system nor the explicit solvent system it could be found.

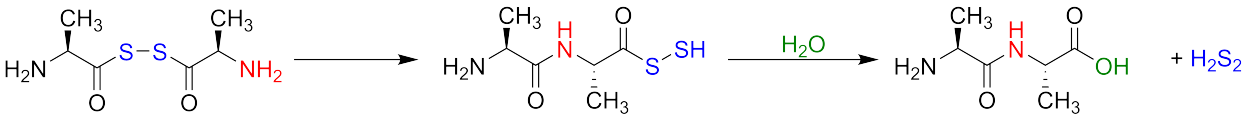


Figure 41: Reaction of Path A

In literature, *S*-to-*N*-Acyl transfers are described as quick processes, where computational studies show that the activation energy is small [73, 74, 75]. This implies that the potential energy surface is quite flat at this point, making it tough to find a transition state. It is possible to investigate flat potential energy surfaces, however, the computational methods do appear to become quite expensive where CCSD(T) in combination with cc-pV5Z or cc-pVQZ as basis sets have been used [76].

On the other hand, in the prebiotic experiment of Okamoto et al. [15], an acidic environment with a pH of 1.2 was used together with $FeCl_3$ as oxidizing agent. These effects and compounds were not simulated in this thesis, which suggests that both the pH and inorganic salt might be of importance for the *S*-to-*N*-Acyl transfer.

7 Conclusion and future perspectives

For path B, three systems have been investigated: vacuum, implicit solvation, and explicit solvation. It was observed that upon adding four explicit water molecules, the activation barrier for thiolysis of the nitrile by hydrogen sulfide was lowered by 20 kcal/mol. Furthermore, the subsequent tautomerization of imidothioic acid readily takes place. With the inclusion of explicit water molecules, the total number of transition states is decreased by one, since inversion of the imine does not take place. Both at room temperature and elevated temperatures up to 373K, the reaction in path B will take place.

It was observed that geometries at a local minimum found with DFTB are not necessarily corresponding to a local minimum found with DFT, which was also the case for transition states. This suggests that there are differences between the potential energy surfaces of DFT and DFTB. This implies that one should be careful by claiming that geometries found with DFTB are the same as geometries found with DFT.

Unfortunately there was not enough time to get all the relevant data for a QM/MM system, which would have allowed for a more accurate description of solvation effects instead of four explicit water molecules. As a future perspective, it would be interesting to further research such a QM/MM system, and figure out whether this affects the reaction energetics.

Initially, only one enantiomer of Alanine was considered, namely L-Alanine. I was wondering whether D-Alanine would actually affect the reaction energetics, and if so, if this could be a reason as to why in nature only L-amino acids are found in proteins. Unfortunately, time was the limiting factor. For the future, looking into D-Alanine in a QM/MM system and comparing it to its L-enantiomer is something to further investigate. However, simply changing from L-Alanine to D-Alanine is not expected to change the reaction energetics, since there is no external chiral influence present to induce a chiral preference for either enantiomer. This will simply lead to the same Hamiltonian, since it is mirrored, which leads to the same energies. Thus, as experimentally done by Weissbuch et al. [17], a chiral imbalance should be introduced to create a preference for either L-Alanine or D-Alanine.

Path A appeared to be tough to investigate, possibly due to a flat potential energy surface. With an increase in level of theory, up to CCSD(T) in combination with a quadruple or pentuple basis set, it might be possible to find relevant transition states, since it has been demonstrated that

this method can find transition states with flat potential energy surfaces [76].

In both Path A and Path B, both pH and a catalyst seem to play an important role in literature [14, 15]. To more accurately describe both these reactions, including pH effects and a catalyst should be done for further research.

8 Acknowledgements

I want to thank Prof. Dr. Ana Da Cunha for giving me the chance to perform this research under her supervision, and providing me with day-to-day guidance.

I want to thank Prof. Dr. Remco Havenith for giving me access to the AMS.2023 package, and for his advice to do this research in Antwerp.

I want to thank the Center for Information Technology of the University of Groningen for their support and for providing access to the Hábrók high performance computing cluster.

I want to thank the Structural Chemistry Group of the University of Antwerpen for the kind and social environment.

I am grateful for the Erasmus grant that gave me the financial means to live in Antwerp, and subsequently allowed me to do this research.

References

(1) Darwin, C., *On the Origin of Species*; John Murray: London, 1859.

(2) Horowitz, N. H.; Miller, S. L. In *Prog. Chem. Org. Nat. Prod.* Springer Vienna: Vienna, 1962, pp 423–459.

(3) Alberts, B.; Johnson, A.; Lewis. J. In *Molecular Biology of the Cell*, 4th ed.; Garland Science: New York, 2002.

(4) Oparin, A. I., *The Origin of Life*, Moscow, 1924.

(5) Haldane, J. B. S., *The Origin of Life*; Rationalist Annual: 1929.

(6) Miller, S. L. *Science* **1953**, *117*, 528–529.

(7) Johnson, A. P.; Cleaves, H. J.; Dworkin, J. P.; Glavin, D. P.; Lazcano, A.; Bada, J. L. *Science* **2008**, *322*, 404–404.

(8) Martin, W.; Russell, M. J. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **2007**, *362*, 1887–1926.

(9) Yoshino, D.; Hayatsu, K.; Anders, E. *Geochim. Cosmochim. Acta* **1971**, *35*, 927–938.

(10) Chatterjee, S. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20033–20046.

(11) Kitadai, N.; Maruyama, S. *Geosci. Front.* **2018**, *9*, 1117–1153.

(12) Damer, B.; Deamer, D. *Astrobiology* **2020**, *20*, 429–452.

(13) Patel, B. H.; Percivalle, C.; Ritson, D. J.; Duffy, C. D.; Sutherland, J. D. *Nat. Chem.* **2015**, *7*, 301–307.

(14) Canavelli, P.; Islam, S.; Powner, M. W. *Nature* **2019**, *571*, 546–549.

(15) Okamoto, R.; Haraguchi, T.; Nomura, K.; Maki, Y.; Izumi, M.; Kajihara, Y. *Biochemistry* **2019**, *58*, 1672–1678.

(16) Soai, K.; Shibata, T.; Morioka, H.; Choji, K. *Nature* **1995**, *378*, 767–768.

(17) Weissbuch, I.; Addadi, L.; Berkovitch-Yellin, Z.; Gati, E.; Lahav, M.; Leiserowitz, L. *Nature* **1984**, *310*, 161–164.

(18) Sallembien, Q.; Bouteiller, L.; Crassous, J.; Raynal, M. *Chem. Soc. Rev.* **2022**, *51*, 3436–3476.

(19) Hughes, E. D.; Ingold, C. K. *J. Chem. Soc.* **1935**, 244.

(20) Daniels, A. D.; Campeotto, I.; van der Kamp, M. W.; Bolt, A. H.; Trinh, C. H.; Phillips, S. E. V.; Pearson, A. R.; Nelson, A.; Mulholland, A. J.; Berry, A. *ACS Chem. Biol.* **2014**, *9*, 1025–1032.

(21) Figueiredo, P. R.; Almeida, B. C.; Dourado, D. F. A. R.; Sousa, A. F.; Silvestre, A. J. D.; Carvalho, A. T. P. *ChemCatChem* **2020**, *12*, 4845–4852.

(22) Hu, X.; Hu, H.; Melvin, J. A.; Clancy, K. W.; McCafferty, D. G.; Yang, W. *J. Am. Chem. Soc.* **2011**, *133*, 478–485.

(23) Świderek, K.; Marti, S.; Tuñón, I.; Moliner, V.; Bertran, J. *J. Am. Chem. Soc.* **2015**, *137*, 12024–12034.

(24) Lambros, E.; Lipparini, F.; Cisneros, G. A.; Paesani, F. *J. Chem. Theory Comput.* **2020**, *16*, 7462–7472.

(25) Vasilevskaya, T.; Khrenova, M. G.; Nemukhin, A. V.; Thiel, W. *J. Comput. Chem.* **2015**, *36*, 1621–1630.

(26) Giudetti, G.; Polyakov, I.; Grigorenko, B. L.; Faraji, S.; Nemukhin, A. V.; Krylov, A. I. *J. Chem. Theory Comput.* **2022**, *18*, 5056–5067.

(27) Bravaya, K. B.; Khrenova, M. G.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. *J. Phys. Chem. B* **2011**, *115*, 8296–8303.

(28) Born, M.; Oppenheimer, R. *Ann. Phys. (Berl.)* **1927**, *389*, 457–484.

(29) Levine, I. N., *Quantum Chemistry*, 7th ed.; Pearson Education: 2013.

(30) Pillardy, J.; Piela, L. *J. Phys. Chem.* **1995**, *99*, 11805–11812.

(31) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A* **1997**, *101*, 5111–5116.

(32) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9901–9904.

(33) Ásgeirsson, V.; Jónsson, H. In *Handbook of Materials Modeling*; Springer International Publishing: Cham, 2018, pp 1–26.

(34) Jay, A.; Gunde, M.; Salles, N.; Poberžnik, M.; Martin-Samos, L.; Richard, N.; Gironcoli, S. d.; Mousseau, N.; Hémeryck, A. *Comput. Mater. Sci.* **2022**, *209*, 111363.

(35) Fukui, K. *J. Phys. Chem.* **1970**, *74*, 4161–4163.

(36) Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363–368.

(37) Jensen, F., *Introduction to Computational Chemistry*, 3rd ed.; John Wiley & Sons: New Jersey, 2017.

(38) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

(39) Atkins, P.; Friedman, R., *Molecular Quantum Mechanics*, 4th ed.; Oxford University Press: Oxford, 2005.

(40) Liao, X.; Lu, R.; Xia, L.; Liu, Q.; Wang, H.; Zhao, K.; Wang, Z.; Zhao, Y. *Energy Environ. Mater.* **2022**, *5*, 157–185.

(41) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(42) Pines, D. *Phys. Rev.* **1953**, *92*, 626–636.

(43) Zhang, R.; Zhao, C.; Huo, Y.; Han, Y.; Hong, J.; Liu, Y.; Zhang, A.; Guo, R.; Ai, Y. *Emerging Nanomaterials for Recovery of Toxic and Radioactive Metal Ions from Environmental Media* **2022**, 313–379.

(44) Sham, L. J.; Reut, Z.; Wilkins, J. W.; Heine, V.; Falicov, L. M.; Edwards, D. M. *Philos. Trans. Royal Soc. A* **1991**, *334*, 481–490.

(45) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(46) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(47) Peverati, R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, *8*, 2310–2319.

(48) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(49) Grimme, S.; Bannwarth, C.; Shushkov, P. *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.

(50) Levenberg, K. *Q. Appl. Math.* **1944**, *2*, 164–168.

(51) Marquardt, D. W. *SIAM J. Appl. Math.* **1963**, *11*, 431–441.

(52) Tkatchenko, A.; DiStasio, R. A.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012**, *108*, 236402.

(53) Born, M. *Z. Phys.* **1920**, *1*, 45–48.

(54) Klamt, A.; Schüürmann, G. *J. Chem. Soc.* **1993**, 799–805.

(55) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. *J. Phys. Chem. A* **1998**, *102*, 5074–5085.

(56) Onufriev, A. *Annu. Rep. Comput. Chem.* **2008**, 125–137.

(57) Klamt, A.; Eckert, F.; Arlt, W. *Annu. Rev. Chem. Biomol.* **2010**, *1*, 101–122.

(58) Bickelhaupt, F. M.; Houk, K. N. *Angew. Chem. Int. Ed.* **2017**, *56*, 10070–10086.

(59) Sun, X.; Soini, T. M.; Poater, J.; Hamlin, T. A.; Bickelhaupt, F. M. *J. Comput. Chem.* **2019**, *40*, 2227–2233.

(60) Chen, K. Y.; Morris, J. C. *Environ. Sci. Technol.* **1972**, *6*, 529–537.

(61) Gálvez, J.; Guirado, A. *J. Comput. Chem.* **2010**, *31*, 520–531.

(62) Würmel, J.; Simmie, J. M. *Int. J. Chem. Kinet.* **2023**, *55*, 381–391.

(63) Xu, X.; Alecu, I. M.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 1667–1676.

(64) Kromann, J. C.; Welford, A.; Christensen, A. S.; Jensen, J. H. *ACS Omega* **2018**, *3*, 4372–4377.

(65) Freeman, F.; Po, H. N. *J. Phys. Chem. A* **2006**, *110*, 7904–7912.

(66) Mo, Y. *J. Chem. Phys.* **2007**, *126*.

(67) Walter, W.; Bode, K.-D. *Angew. Chem. Int. Ed.* **1966**, *5*, 447–461.

(68) Vicent-Luna, J. M.; Apergi, S.; Tao, S. *J. Chem. Inf. Model.* **2021**, *61*, 4415–4424.

(69) Bursch, M.; Hansen, A.; Pracht, P.; Kohn, J. T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2021**, *23*, 287–299.

(70) Spicher, S.; Grimme, S. *J. Chem. Theory Comput.* **2021**, *17*, 1701–1714.

(71) Elstner, M. *Theor. Chem. Acc.* **2006**, *116*, 316–325.

(72) Nénon, S.; Champagne, B. *J. Chem. Phys.* **2013**, *138*.

(73) Devaraj, N. K.; Perrin, C. L. *Chem. Sci.* **2018**, *9*, 1789–1794.

(74) Liu, R.; Orgel, L. E. *Nature* **1997**, *389*, 52–54.

(75) Asiimwe, N.; Al Mazid, M. F.; Murale, D. P.; Kim, Y. K.; Lee, J.-S. *Pept. Sci.* **2022**, *114*.

(76) Kim, S.; Wheeler, S. E.; DeYonker, N. J.; Schaefer, H. F. *J. Chem. Phys.* **2005**, *122*.

9 Supplementary information

9.1 Path B xyz coordinates, frequencies, and IRCs

9.1.1 Vacuum

1			
atom	x	y	z
C	2.425898	0.749192	-0.75161
N	1.215977	0.450223	-0.19789
C	1.062103	-0.74172	0.60966
C	-0.27742	-1.30456	0.385836
C	1.319792	-0.49935	2.102441
H	1.802555	-1.46283	0.240211
O	3.405496	0.042157	-0.58549
H	2.952804	1.80695	-2.52051
H	0.411792	1.045209	-0.35964
N	-1.34782	-1.70509	0.225929
H	1.215744	-1.42921	2.667982
H	0.617266	0.237366	2.501034
H	2.339716	-0.12535	2.21695
S	-1.94534	1.837556	-0.89588
H	-2.22685	0.53446	-0.71472
H	-2.34463	2.151724	0.346501
C	2.464033	2.022382	-1.56732
H	1.47293	2.446088	-1.74845
H	3.080265	2.755923	-1.03864

Table 1: xyz coordinates molecule 1

1			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	19.9084	26	1087.2833
2	39.5383	27	1117.4889
3	57.4264	28	1181.2903
4	65.4426	29	1205.6088
5	87.2242	30	1273.8304
6	97.771	31	1321.1647
7	100.8185	32	1344.745
8	159.8178	33	1383.9353
9	184.8658	34	1389.4017
10	191.5328	35	1452.7563
11	236.6461	36	1469.0961
12	294.5285	37	1474.8032
13	305.419	38	1477.1774
14	373.8205	39	1553.077
15	420.0333	40	1776.924
16	557.9138	41	2370.2685
17	562.3584	42	2700.0849
18	582.9222	43	2731.4373
19	614.3405	44	3056.7457
20	662.5969	45	3058.298
21	843.3807	46	3067.4352
22	923.9493	47	3138.5865
23	995.5041	48	3141.9024
24	1008.2242	49	3151.4878
25	1044.8938	50	3152.569
		51	3542.423

Table 2: Frequencies molecule 1

TS 1			
atom	x	y	z
C	2.316459	0.994184	-1.09265
N	1.303141	0.804428	-0.19549
C	0.999268	-0.51366	0.293624
C	-0.40336	-0.82881	0.004574
C	1.317087	-0.71312	1.777293
H	1.601319	-1.20567	-0.32066
O	3.031274	0.077878	-1.46019
H	1.814213	3.12208	-1.09578
H	0.633653	1.530504	0.035669
N	-1.41815	-1.34614	-0.31497
H	1.080972	-1.73211	2.095669
H	0.744556	-0.00455	2.381271
H	2.384154	-0.53174	1.923624
S	-1.88699	1.25429	0.660266
H	-2.16733	-0.51818	-0.04832
H	-2.01789	0.87057	1.941191
C	2.485478	2.412157	-1.58418
H	3.523016	2.71455	-1.42188
H	2.303432	2.424457	-2.66267

Table 3: xyz coordinates molecule TS 1

TS 1 Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-1279.776	26	1044.4818
2	30.2947	27	1071.0779
3	38.0453	28	1122.5488
4	48.8886	29	1186.066
5	90.8543	30	1237.9584
6	121.0264	31	1298.7141
7	174.0646	32	1309.9764
8	189.1428	33	1384.1569
9	213.589	34	1390.019
10	244.0408	35	1452.8638
11	270.7161	36	1464.6963
12	358.6946	37	1470.4392
13	409.5747	38	1478.5333
14	430.7128	39	1521.4218
15	468.0415	40	1775.3188
16	573.156	41	1793.7045
17	590.3833	42	2158.0905
18	601.0284	43	2714.6835
19	630.4662	44	3007.3023
20	668.6521	45	3059.5345
21	835.7884	46	3062.3179
22	918.083	47	3142.3704
23	987.0055	48	3146.6947
24	999.8537	49	3155.1705
25	1009.4785	50	3155.338
		51	3550.4539

Table 4: Frequencies molecule TS 1

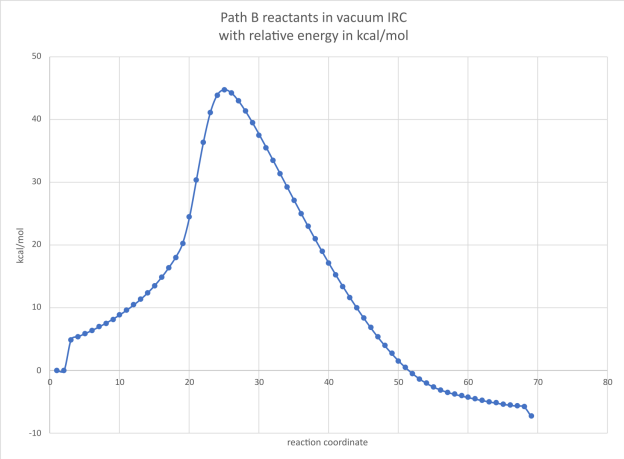


Figure 1: IRC of TS 1

1 #IRC	E_{elstat}	EnergyTotal	E_{int}	E_{oi}	E_{Pauli}	Steric	StrainTotal	frag1Strain	frag2Strain
1	-10.734	-3102.99	-4.461	-6.93	15.03	4.296	-3098.53	-326.778	-2771.75
2	-15.672	-3097.92	0.259	-7.046	24.988	9.316	-3098.18	-326.775	-2771.41
3	-16.085	-3097.42	0.691	-7.179	25.965	9.881	-3098.11	-326.771	-2771.34
4	-16.518	-3096.91	1.15	-7.324	27	10.482	-3098.06	-326.776	-2771.28
5	-16.982	-3096.34	1.631	-7.486	28.105	11.123	-3097.97	-326.771	-2771.2
6	-17.482	-3095.76	2.134	-7.673	29.292	11.811	-3097.89	-326.771	-2771.12
7	-18.01	-3095.14	2.662	-7.881	30.553	12.543	-3097.8	-326.782	-2771.02
8	-18.579	-3094.47	3.224	-8.11	31.91	13.331	-3097.7	-326.774	-2770.92
9	-19.217	-3093.76	3.819	-8.388	33.419	14.202	-3097.58	-326.775	-2770.8
10	-19.911	-3093.01	4.446	-8.701	35.049	15.138	-3097.46	-326.781	-2770.67
11	-20.677	-3092.2	5.099	-9.066	36.83	16.153	-3097.3	-326.774	-2770.52
12	-21.508	-3091.37	5.79	-9.473	38.755	17.247	-3097.16	-326.783	-2770.37
13	-22.431	-3090.46	6.516	-9.943	40.872	18.44	-3096.97	-326.777	-2770.2
14	-23.443	-3089.51	7.29	-10.478	43.187	19.745	-3096.8	-326.786	-2770.02
15	-24.561	-3088.49	8.102	-11.085	45.722	21.161	-3096.6	-326.779	-2769.82
16	-25.792	-3087.43	8.966	-11.774	48.503	22.71	-3096.4	-326.79	-2769.61
17	-27.148	-3086.29	9.873	-12.55	51.538	24.39	-3096.16	-326.778	-2769.39
18	-28.636	-3085.12	10.839	-13.426	54.863	26.228	-3095.96	-326.8	-2769.16
19	-30.29	-3083.83	11.865	-14.42	58.535	28.244	-3095.7	-326.787	-2768.91
20	-32.085	-3082.5	12.95	-15.527	62.516	30.432	-3095.44	-326.789	-2768.66
21	-34.037	-3081.08	14.097	-16.76	66.845	32.808	-3095.18	-326.79	-2768.39
22	-36.184	-3079.58	15.316	-18.146	71.595	35.41	-3094.9	-326.79	-2768.11
23	-38.551	-3077.99	16.608	-19.71	76.815	38.263	-3094.6	-326.789	-2767.81
24	-41.13	-3076.31	17.975	-21.458	82.506	41.375	-3094.29	-326.789	-2767.5
25	-43.937	-3074.53	19.423	-23.416	88.715	44.778	-3093.96	-326.788	-2767.17
26	-47.012	-3072.64	20.964	-25.632	95.545	48.532	-3093.6	-326.779	-2766.82
27	-50.393	-3070.6	22.615	-28.175	103.117	52.724	-3093.22	-326.773	-2766.45
28	-54.045	-3068.42	24.384	-31.066	111.429	57.384	-3092.8	-326.748	-2766.05
29	-58.038	-3066.04	26.325	-34.468	120.763	62.725	-3092.37	-326.731	-2765.64
30	-62.324	-3063.41	28.45	-38.495	131.201	68.877	-3091.86	-326.685	-2765.18
31	-67.104	-3060.53	30.798	-43.686	143.522	76.418	-3091.33	-326.717	-2764.61
32	-72.8	-3057.52	33.001	-51.442	159.182	86.382	-3090.52	-326.678	-2763.84
33	-79.486	-3054.28	33.698	-63.568	178.697	99.211	-3087.98	-325.054	-2762.92
34	-86.929	-3051.62	31.223	-81.441	201.545	114.616	-3082.85	-320.947	-2761.9
35	-94.908	-3051.33	24.37	-106.569	227.806	132.897	-3075.7	-314.569	-2761.13
36	-103.593	-3054.28	12.861	-140.367	258.782	155.189	-3067.14	-306.871	-2760.27
37	-112.728	-3059.91	-2.393	-182.423	294.72	181.992	-3057.51	-298.682	-2758.83
38	-121.3	-3066.62	-19.411	-229.979	333.825	212.525	-3047.21	-289.944	-2757.26
39	-128.506	-3072.4	-35.882	-274.584	369.156	240.651	-3036.52	-281.037	-2755.48
40	-134.797	-3076.88	-50.48	-305.523	391.777	256.98	-3026.4	-272.804	-2753.6
41	-141.055	-3080.63	-63.183	-325.67	405.47	264.415	-3017.45	-265.67	-2751.78
42	-147.527	-3083.92	-74.309	-343.204	418.344	270.817	-3009.61	-259.542	-2750.07
43	-154.171	-3086.89	-84.153	-360.722	432.658	278.487	-3002.73	-254.384	-2748.35
44	-160.845	-3089.61	-92.812	-377.606	447.553	286.708	-2996.79	-250.152	-2746.64
45	-167.586	-3092.12	-100.471	-394.656	463.685	296.099	-2991.65	-246.684	-2744.96
46	-174.386	-3094.41	-107.307	-411.401	480.393	306.007	-2987.11	-243.768	-2743.34
47	-181.285	-3096.51	-113.418	-428.294	498.076	316.791	-2983.09	-241.264	-2741.83
48	-188.231	-3098.38	-118.952	-444.658	515.855	327.623	-2979.42	-238.992	-2740.43
49	-195.19	-3100.03	-123.898	-461.12	534.335	339.144	-2976.14	-236.926	-2739.21
50	-201.973	-3101.46	-128.355	-476.839	552.384	350.411	-2973.1	-235	-2738.1
51	-208.367	-3102.65	-132.387	-491.33	569.241	360.875	-2970.26	-233.049	-2737.21
52	-214.092	-3103.62	-136.022	-504.598	584.607	370.515	-2967.6	-231.138	-2736.46
53	-218.832	-3104.39	-139.266	-516.335	597.846	379.014	-2965.13	-229.304	-2735.82
54	-222.221	-3104.98	-142.211	-525.513	607.475	385.254	-2962.77	-227.43	-2735.34
55	-224.158	-3105.44	-144.725	-532.174	613.567	389.409	-2960.71	-225.757	-2734.95
56	-224.87	-3105.78	-146.727	-536.213	616.325	391.455	-2959.05	-224.278	-2734.77
57	-225.029	-3106.02	-148.174	-538.677	617.509	392.48	-2957.85	-223.147	-2734.7
58	-225.029	-3106.18	-149.05	-539.792	617.759	392.73	-2957.13	-222.347	-2734.78
59	-221.858	-3109.13	-169.659	-615.034	669.375	447.517	-2939.47	-206.74	-2732.73

Table 5: ASM data based on the TS of molecule 1 to 2 with energies in kcal/mol

2			
atom	x	y	z
C	-3.02346	-0.75409	0.191379
N	-1.68947	-0.53803	0.358313
C	-1.01648	0.564703	-0.308
C	0.461254	0.248677	-0.4715
C	-1.24058	1.899966	0.408107
H	-1.42176	0.633621	-1.32101
O	-3.71406	-0.04303	-0.51879
H	-2.83078	-2.58062	1.392818
H	-1.15786	-1.1406	0.96541
N	0.980551	0.227172	-1.62236
H	-0.74512	2.710474	-0.13438
H	-0.85334	1.873992	1.431809
H	-2.31385	2.101623	0.440563
S	1.287405	-0.02749	1.094959
H	1.982738	0.046644	-1.61668
H	2.435473	-0.47829	0.56662
C	-3.59437	-1.9297	0.958408
H	-4.23161	-1.54442	1.760478
H	-4.22786	-2.50806	0.282036

Table 6: xyz coordinates molecule 2

2			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	36.0975	26	1096.937
2	48.8858	27	1120.944
3	52.5046	28	1174.056
4	99.7617	29	1252.592
5	140.0216	30	1264.682
6	208.2466	31	1313.702
7	217.8934	32	1365.389
8	245.1802	33	1382.545
9	269.4884	34	1387.412
10	363.0765	35	1450.18
11	383.1235	36	1464.136
12	422.6338	37	1472.31
13	461.081	38	1483.157
14	525.8636	39	1534.108
15	582.1276	40	1725.467
16	629.3729	41	1777.198
17	650.428	42	2720.564
18	673.32	43	3049.07
19	862.8512	44	3056.841
20	875.2609	45	3107.629
21	911.1199	46	3131.204
22	961.6213	47	3134.246
23	990.6691	48	3147.312
24	1010.237	49	3151.325
25	1042.092	50	3519.439
		51	3660.195

Table 7: Frequencies molecule 2

TS 2			
atom	x	y	z
C	-2.66472	-1.25813	-0.11437
N	-1.52636	-0.63951	0.289878
C	-1.13774	0.653428	-0.21085
C	0.374184	0.816842	-0.01159
C	-1.90775	1.791355	0.471852
H	-1.35408	0.686516	-1.28631
O	-3.44962	-0.73954	-0.89274
H	-2.99783	-3.34816	-0.35889
H	-0.86098	-1.10534	0.888035
N	1.077426	0.140442	0.734834
H	-2.97677	1.629301	0.316022
H	-1.63213	2.761181	0.050038
H	-1.69622	1.801085	1.545731
S	0.98567	2.179999	-1.03214
H	1.582965	-0.44033	1.357079
H	2.241279	2.095762	-0.56969
C	-2.90174	-2.63745	0.466385
H	-2.10885	-2.97023	1.141404
H	-3.8536	-2.62832	1.004359

Table 8: xyz coordinates molecule TS 2

TS 2 Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-1020.86	26	1043.278
2	29.1951	27	1067.639
3	50.132	28	1113.715
4	62.7864	29	1203.25
5	101.1557	30	1241.449
6	138.0826	31	1308.773
7	222.8692	32	1362.458
8	238.967	33	1381.937
9	264.3931	34	1386.556
10	309.386	35	1452.183
11	316.8018	36	1461.806
12	390.7321	37	1473.954
13	418.9887	38	1483.561
14	451.041	39	1533.856
15	483.9773	40	1769.985
16	498.7989	41	1832.283
17	619.7818	42	2731.353
18	627.0084	43	3051.756
19	659.779	44	3057.803
20	745.9822	45	3061.142
21	822.0366	46	3136.447
22	912.052	47	3140.712
23	986.5312	48	3144.19
24	999.6409	49	3150.448
25	1007.845	50	3633.878
		51	3941.29

Table 9: Frequencies molecule TS 2

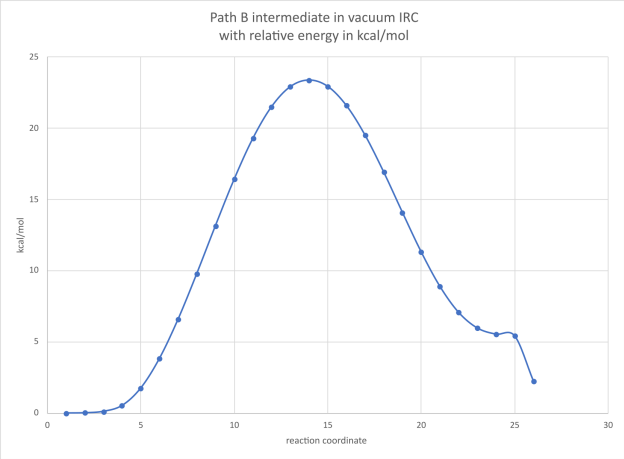


Figure 2: IRC of TS 2

2 #IRC	E_{elstat}	EnergyTotal	E_{int}	E_{oi}	E_{Pauli}	Steric	StrainTotal	frag1Strain	frag2Strain
1	-220.345	-3114.06	-168.009	-613.428	667.484	447.139	-2946.05	-213.197	-2732.86
2	-218.465	-3114.15	-168.917	-611.588	662.853	444.389	-2945.23	-212.512	-2732.72
3	-218.85	-3113.91	-171.293	-611.234	660.5	441.65	-2942.62	-210.38	-2732.24
4	-222.273	-3112.8	-173.041	-607.561	658.496	436.223	-2939.76	-207.02	-2732.74
5	-226.026	-3110.75	-173.014	-599.306	654.021	427.995	-2937.74	-203.949	-2733.79
6	-229.625	-3107.95	-171.593	-587.646	647.385	417.761	-2936.36	-201.321	-2735.04
7	-232.926	-3104.64	-169.216	-573.28	638.704	405.779	-2935.42	-199.089	-2736.33
8	-235.802	-3101.1	-166.294	-556.923	628.155	392.353	-2934.81	-197.206	-2737.6
9	-238.029	-3097.64	-163.221	-538.841	615.386	377.357	-2934.42	-195.612	-2738.81
10	-239.56	-3094.56	-160.371	-520.013	600.952	361.391	-2934.19	-194.285	-2739.9
11	-240.163	-3092.12	-158.065	-500.47	584.331	344.168	-2934.06	-193.175	-2740.88
12	-239.849	-3090.55	-156.561	-481.112	566.177	326.328	-2933.99	-192.272	-2741.72
13	-238.899	-3089.98	-156.285	-463.567	547.971	309.073	-2933.7	-191.502	-2742.2
14	-237.259	-3090.39	-156.984	-447.095	529.174	291.914	-2933.41	-190.897	-2742.51
15	-234.869	-3091.78	-158.452	-431.021	509.252	274.383	-2933.33	-190.483	-2742.85
16	-232.233	-3093.98	-160.774	-416.794	490.077	257.843	-2933.2	-190.21	-2742.99
17	-229.642	-3096.73	-163.723	-404.568	472.32	242.679	-2933.01	-190.086	-2742.92
18	-227.408	-3099.75	-166.974	-394.366	456.642	229.234	-2932.77	-190.113	-2742.66
19	-225.929	-3102.7	-170.19	-386.26	443.849	217.92	-2932.51	-190.3	-2742.21
20	-225.664	-3105.28	-173.04	-380.386	434.87	209.205	-2932.24	-190.659	-2741.58
21	-227.217	-3107.22	-175.245	-377.147	430.989	203.772	-2931.98	-191.2	-2740.78
22	-231.348	-3108.32	-176.65	-377.55	434.128	202.78	-2931.67	-191.89	-2739.78
23	-235.781	-3108.63	-177.316	-380.194	440.546	204.765	-2931.31	-192.373	-2738.94
24	-240.221	-3111.37	-183.476	-380.289	438.99	198.769	-2927.89	-192.581	-2735.31

Table 10: ASM data based on the TS of molecule 2 to 3 with energies in kcal/mol

3 atom	x	y	z
C	-2.33772	-1.44487	-0.27626
N	-1.70177	-0.49965	0.483387
C	-1.25097	0.755865	-0.0652
C	0.280597	0.797299	-0.15843
C	-1.83431	1.934735	0.716332
H	-1.62614	0.776549	-1.09618
O	-2.49095	-1.32439	-1.47691
H	-2.47647	-3.55375	-0.01056
H	-1.67672	-0.62073	1.483215
N	1.070959	-0.16036	0.081452
H	-2.9253	1.867178	0.719504
H	-1.55203	2.889778	0.266542
H	-1.47581	1.935019	1.751743
S	0.907882	2.359354	-0.70201
H	0.540204	-0.98897	0.354889
H	2.17546	1.921812	-0.7759
C	-2.845	-2.65223	0.484446
H	-2.5511	-2.66671	1.537249
H	-3.93702	-2.66522	0.417214

Table 11: xyz coordinates molecule 3

3			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	28.1591	26	1082.266
2	49.0979	27	1129.514
3	55.3542	28	1193.354
4	84.6829	29	1248.528
5	167.5813	30	1278.933
6	210.0671	31	1313.811
7	240.4973	32	1350.951
8	289.5699	33	1382.311
9	330.0674	34	1400.269
10	360.4351	35	1451.09
11	406.4684	36	1468.686
12	451.781	37	1475.993
13	474.6289	38	1478.223
14	501.6252	39	1533.504
15	606.1622	40	1706.016
16	623.7256	41	1790.208
17	634.6106	42	2716.414
18	743.6162	43	3047.897
19	844.2124	44	3058.812
20	886.3923	45	3071.632
21	904.1949	46	3126.873
22	964.9786	47	3138.542
23	989.0393	48	3143.379
24	1007.728	49	3150.009
25	1041.456	50	3482.14
		51	3658.372

Table 12: Frequencies molecule 3

TS 3			
atom	x	y	z
C	-1.21825	-1.50172	0.304917
N	-1.36253	-0.16083	0.48437
C	-0.52754	0.797458	-0.21894
C	0.809278	0.912284	0.494238
C	-1.24723	2.126397	-0.3763
H	-0.32152	0.356909	-1.2012
O	-0.40357	-1.96706	-0.47774
H	-1.51444	-2.92314	1.860024
H	-1.97474	0.181515	1.207008
N	1.753026	0.027612	0.428554
H	-0.6188	2.834822	-0.92076
H	-1.4682	2.566703	0.602081
H	-2.18178	1.986099	-0.92646
S	1.314261	2.159065	1.565074
H	1.597319	-0.8177	-0.1227
H	2.347476	0.919742	1.333509
C	-2.12786	-2.37879	1.135575
H	-2.9021	-1.8224	1.669935
H	-2.59472	-3.11561	0.477922

Table 13: xyz coordinates molecule TS 3

TS 3 Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-1669.37	26	1106.296
2	25.4416	27	1144.525
3	53.3636	28	1193.305
4	81.5138	29	1265.434
5	98.2514	30	1312.22
6	180.8137	31	1337.833
7	210.8617	32	1356.619
8	233.4988	33	1384.032
9	258.8688	34	1405.258
10	334.7614	35	1451.462
11	408.8949	36	1468.317
12	445.8832	37	1473.673
13	496.7105	38	1479.414
14	551.2903	39	1534.896
15	613.4098	40	1610.188
16	636.277	41	1765.536
17	668.599	42	1798.906
18	719.227	43	3049.876
19	770.6456	44	3059.758
20	905.5363	45	3077.218
21	920.239	46	3128.271
22	957.4123	47	3138.774
23	992.5465	48	3147.43
24	1017.878	49	3152.958
25	1044.165	50	3485.729
		51	3666.646

Table 14: Frequencies molecule TS 3

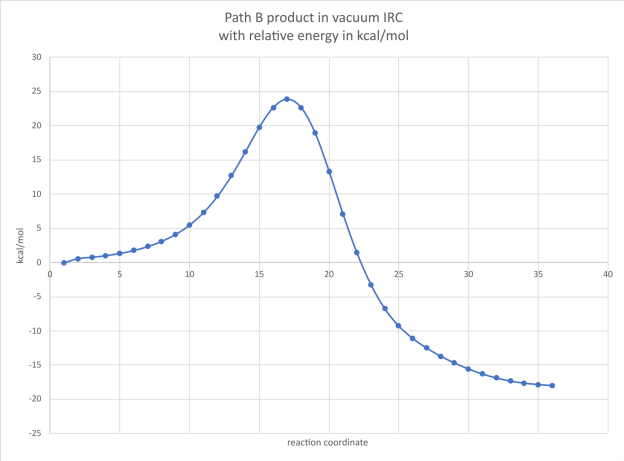


Figure 3: IRC of TS 3

3 #IRC	E_{elstat}	EnergyTotal	E_{int}	E_{oi}	E_{Pauli}	Steric	StrainTotal	frag1Strain	frag2Strain
1	-245.189	-3110.42	-182.299	-384.343	449.337	204.147	-2928.12	-192.463	-2735.66
2	-248.267	-3109.85	-181.907	-388.249	456.694	208.427	-2927.95	-192.597	-2735.35
3	-248.5	-3109.69	-181.85	-388.613	457.341	208.841	-2927.84	-192.652	-2735.19
4	-248.786	-3109.46	-181.75	-388.778	457.888	209.102	-2927.71	-192.688	-2735.03
5	-249.173	-3109.19	-181.435	-388.985	458.79	209.617	-2927.75	-192.785	-2734.97
6	-249.786	-3108.83	-180.968	-389.142	460.02	210.234	-2927.86	-192.903	-2734.95
7	-250.771	-3108.37	-180.236	-389.384	461.972	211.201	-2928.13	-193.083	-2735.05
8	-252.239	-3107.79	-179.237	-389.811	464.858	212.619	-2928.55	-193.306	-2735.25
9	-254.272	-3107.06	-177.985	-390.615	468.94	214.668	-2929.08	-193.586	-2735.49
10	-256.704	-3106.16	-176.507	-391.707	473.935	217.231	-2929.66	-193.869	-2735.79
11	-259.383	-3105.03	-174.945	-393.431	479.893	220.51	-2930.09	-194.105	-2735.98
12	-262.146	-3103.63	-173.471	-396.128	486.821	224.675	-2930.16	-194.212	-2735.95
13	-264.955	-3101.83	-172.269	-400.429	495.129	230.174	-2929.56	-194.12	-2735.44
14	-267.965	-3099.55	-171.3	-406.867	505.544	237.579	-2928.25	-193.832	-2734.42
15	-271.489	-3096.76	-170.339	-416.098	519.262	247.772	-2926.42	-193.417	-2733
16	-275.955	-3093.69	-169.47	-429.064	537.568	261.613	-2924.22	-192.794	-2731.42
17	-281.481	-3090.59	-169.448	-446.378	560.441	278.96	-2921.15	-191.228	-2729.92
18	-288	-3087.85	-171.311	-468.885	587.617	299.617	-2916.54	-187.959	-2728.58
19	-295.335	-3086.26	-175.951	-497.126	618.567	323.232	-2910.31	-182.984	-2727.32
20	-303.754	-3086.67	-183.305	-534.212	656.73	352.976	-2903.36	-176.848	-2726.52
21	-313.447	-3089.59	-193.828	-581.011	702.708	389.261	-2895.76	-170.06	-2725.7
22	-324.116	-3094.54	-207.043	-636.31	755.469	431.353	-2887.5	-163.09	-2724.41
23	-335.425	-3100.36	-220.894	-699.601	816.22	480.795	-2879.47	-156.404	-2723.06
24	-346.41	-3105.78	-233.944	-762.673	877.223	530.813	-2871.84	-150.284	-2721.56
25	-356.324	-3110.42	-245.511	-815.2	928.09	571.767	-2864.91	-144.98	-2719.93
26	-364.854	-3114.15	-255.222	-855.944	967.649	602.795	-2858.93	-140.657	-2718.28
27	-372.173	-3116.88	-262.858	-889.829	1001.214	629.041	-2854.02	-137.432	-2716.59
28	-378.223	-3118.82	-268.711	-918.276	1029.86	651.638	-2850.11	-135.183	-2714.93
29	-382.609	-3120.36	-273.456	-939.652	1050.882	668.274	-2846.9	-133.46	-2713.44
30	-385.41	-3121.72	-277.46	-955.282	1065.317	679.908	-2844.26	-131.986	-2712.27
31	-386.912	-3122.98	-280.865	-967.115	1075.257	688.345	-2842.11	-130.676	-2711.44
32	-387.5	-3124.12	-283.779	-976.702	1082.527	695.027	-2840.34	-129.514	-2710.83
33	-387.585	-3125.14	-286.313	-985.019	1088.403	700.818	-2838.82	-128.486	-2710.34
34	-387.472	-3126.01	-288.549	-992.639	1093.682	706.211	-2837.46	-127.574	-2709.89
35	-387.338	-3126.74	-290.55	-999.681	1098.598	711.26	-2836.19	-126.747	-2709.45
36	-387.32	-3127.33	-292.258	-1006.94	1104.137	716.817	-2835.08	-126.077	-2709
37	-387.216	-3127.79	-293.845	-1012.96	1108.473	721.257	-2833.94	-125.43	-2708.51
38	-387.169	-3128.09	-295.067	-1018.27	1112.528	725.359	-2833.03	-124.982	-2708.05
39	-385.834	-3128.4	-296.683	-1019.81	1111.142	725.308	-2831.71	-125.358	-2706.36

Table 15: ASM data based on the TS of molecule 3 to 4 with energies in kcal/mol

4			
atom	x	y	z
C	-1.11022	-1.46111	0.3417
N	-1.29652	-0.13486	0.539942
C	-0.51426	0.886806	-0.14371
C	0.819178	1.094448	0.599107
C	-1.33887	2.148784	-0.31387
H	-0.26875	0.469317	-1.12763
O	-0.24223	-1.89593	-0.41042
H	-1.4355	-3.03451	1.737346
H	-1.96443	0.173078	1.228339
N	1.666732	0.080125	0.397979
H	-0.76513	2.905052	-0.85305
H	-1.60243	2.576236	0.657963
H	-2.24862	1.920144	-0.87687
S	1.141825	2.396444	1.554096
H	1.344007	-0.7523	-0.08977
H	2.541596	0.071544	0.897878
C	-2.03826	-2.3838	1.09764
H	-2.77581	-1.85726	1.708286
H	-2.55422	-3.02087	0.374159

Table 16: xyz coordinates molecule 4

4			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	44.582	26	1146.338
2	65.315	27	1179.488
3	96.6364	28	1271.986
4	136.2587	29	1316.441
5	196.1183	30	1332.027
6	237.7588	31	1343.852
7	252.7585	32	1384.731
8	266.0409	33	1400.079
9	334.3408	34	1450.91
10	416.272	35	1457.262
11	447.6779	36	1465.235
12	484.5869	37	1467.367
13	518.458	38	1477.252
14	577.0705	39	1543.316
15	613.0017	40	1635.184
16	637.9468	41	1742.632
17	680.3171	42	3056.415
18	718.5477	43	3060.305
19	747.917	44	3067.225
20	911.6096	45	3134.971
21	959.204	46	3141.316
22	994.6519	47	3149.222
23	1018.76	48	3151.773
24	1046.038	49	3464.62
25	1065.66	50	3664.455
		51	3686.318

Table 17: Frequencies molecule 4

9.1.2 Implicit solvent

1			
atom	x	y	z
C	2.403859	0.650508	-0.93685
N	1.340244	0.394547	-0.14419
C	1.265945	-0.77328	0.702422
C	0.385591	-1.7931	0.095671
C	0.799689	-0.42911	2.118421
H	2.269119	-1.21107	0.730913
O	3.38034	-0.1001	-0.99481
H	2.432812	1.669783	-2.80021
H	0.540358	1.022106	-0.16149
N	-0.31487	-2.5817	-0.37223
H	0.747774	-1.33198	2.730769
H	-0.18864	0.038655	2.098258
H	1.514623	0.265708	2.564294
S	-1.53646	2.355264	-0.28757
H	-2.38691	1.377673	-0.63806
H	-1.90001	2.305855	1.003734
C	2.326872	1.922292	-1.74148
H	1.394021	2.468307	-1.589
H	3.169944	2.560759	-1.4622

Table 18: xyz coordinates molecule 1

1			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	64.6655	26	1077.23
2	24.2282	27	1127.505
3	17.4553	28	1179.676
4	41.8434	29	1182.155
5	73.5337	30	1286.505
6	85.0045	31	1307.718
7	93.269	32	1347.265
8	117.3774	33	1375.701
9	173.5526	34	1392.495
10	183.8173	35	1438.758
11	200.6332	36	1456.87
12	252.9984	37	1461.017
13	306.8351	38	1470.778
14	367.5323	39	1569.322
15	418.5599	40	1693.785
16	568.5735	41	2362.539
17	577.3437	42	2721.118
18	603.7772	43	2735.734
19	641.1923	44	3060.417
20	688.4034	45	3061.288
21	833.837	46	3097.823
22	914.8056	47	3138.958
23	993.8445	48	3145.713
24	1015.908	49	3154.707
25	1043.021	50	3158.563
		51	3479.801

Table 19: Frequencies molecule 1

2			
atom	x	y	z
C	-2.62601	-1.15945	0.02845
N	-1.98075	-0.12134	0.6015
C	-1.18258	0.814611	-0.1603
C	0.161991	0.229211	-0.59315
C	-0.95821	2.085118	0.658346
H	-1.71496	1.058554	-1.08479
O	-2.5402	-1.3847	-1.18111
H	-3.46573	-1.65793	1.976228
H	-2.10016	0.040667	1.589493
N	0.728752	0.688976	-1.63102
H	-0.37537	2.80249	0.077839
H	-0.41634	1.861055	1.583159
H	-1.92129	2.535588	0.913931
S	0.773117	-1.00984	0.51202
H	1.639681	0.269772	-1.80866
H	1.973505	-1.1261	-0.07791
C	-3.43912	-2.02946	0.950219
H	-4.45825	-2.09798	0.56042
H	-3.01125	-3.03667	0.943502

Table 20: xyz coordinates molecule 2

2			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	37.1402	26	1080.808
2	49.9993	27	1125.173
3	63.2842	28	1200.167
4	84.6475	29	1271.567
5	161.4736	30	1279.11
6	215.7714	31	1308.734
7	237.7748	32	1348.714
8	261.2523	33	1374.972
9	362.5524	34	1385.211
10	373.0885	35	1437.632
11	456.9142	36	1453.792
12	470.578	37	1462.02
13	496.7094	38	1467.646
14	539.8922	39	1542.519
15	599.346	40	1682.677
16	627.2808	41	1696.315
17	654.5861	42	2718.471
18	693.5858	43	3051.559
19	842.4327	44	3059.387
20	885.9901	45	3092.228
21	930.6797	46	3130.643
22	973.9133	47	3135.007
23	999.2198	48	3153.941
24	1021.698	49	3156.555
25	1041.625	50	3523.054
		51	3652.368

Table 21: Frequencies molecule 2

TS 2			
atom	x	y	z
C	-2.65102	-1.27652	-0.14082
N	-1.54662	-0.63287	0.268381
C	-1.15392	0.667528	-0.21828
C	0.363066	0.809148	-0.02179
C	-1.89939	1.798692	0.501144
H	-1.37416	0.718835	-1.29055
O	-3.43511	-0.78785	-0.96383
H	-2.96207	-3.37089	-0.34721
H	-0.91062	-1.07622	0.917462
N	1.042694	0.143785	0.750879
H	-2.97542	1.663139	0.366361
H	-1.61608	2.773081	0.095294
H	-1.67012	1.780639	1.571282
S	1.002933	2.136584	-1.08662
H	1.553037	-0.42807	1.384734
H	2.252855	2.078449	-0.60398
C	-2.88642	-2.63932	0.462316
H	-2.09681	-2.9428	1.152559
H	-3.84369	-2.62644	0.991692

Table 22: xyz coordinates molecule TS 2

TS 2 Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-1048.65	26	1042.266
2	18.5935	27	1065.677
3	44.7772	28	1107.188
4	59.248	29	1189.795
5	102.5434	30	1255.211
6	136.9814	31	1306.861
7	220.8823	32	1356.472
8	232.0087	33	1373.972
9	261.6184	34	1378.669
10	285.6786	35	1438.747
11	312.1197	36	1446.49
12	391.6863	37	1460.617
13	431.4112	38	1463.239
14	441.7383	39	1548.065
15	482.2765	40	1675.78
16	529.6034	41	1810.281
17	618.8859	42	2727.107
18	629.4056	43	3050.706
19	660.0586	44	3059.061
20	731.0557	45	3075.684
21	805.0168	46	3134.444
22	905.5623	47	3136.065
23	968.8788	48	3144.014
24	999.8168	49	3154.221
25	1014.629	50	3604.133
		51	3867.321

Table 23: Frequencies molecule TS 2

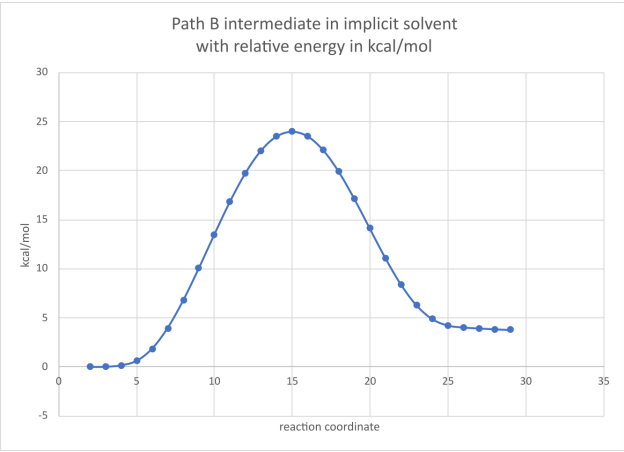


Figure 4: IRC of TS 2

3			
atom	x	y	z
C	-2.14776	-1.47768	-0.30891
N	-1.89476	-0.4143	0.485636
C	-1.36903	0.829651	-0.02526
C	0.145082	0.743736	-0.25402
C	-1.71096	1.967941	0.936911
H	-1.8338	1.020131	-0.99948
O	-1.88255	-1.46405	-1.5126
H	-2.12559	-3.54467	0.18757
H	-2.16106	-0.4588	1.457188
N	0.924751	-0.1186	0.254377
H	-2.79595	2.049165	1.041119
H	-1.32426	2.917747	0.563061
H	-1.27141	1.780473	1.922159
S	0.733854	2.016419	-1.31923
H	0.388914	-0.77446	0.827483
H	2.028829	1.679413	-1.21293
C	-2.76714	-2.67647	0.362423
H	-2.90854	-2.53945	1.436125
H	-3.73483	-2.87548	-0.10709

Table 24: xyz coordinates molecule 3

3			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	37.4848	26	1068.686
2	50.0207	27	1129.338
3	69.555	28	1203.052
4	95.6059	29	1262.149
5	160.9148	30	1277.758
6	205.3662	31	1316.318
7	226.4972	32	1372.875
8	281.0394	33	1381.052
9	308.7608	34	1396.415
10	340.7053	35	1438.61
11	376.1363	36	1453.654
12	447.6209	37	1463.154
13	503.5495	38	1472.356
14	527.6557	39	1541.13
15	621.0422	40	1680.249
16	638.4057	41	1700.006
17	648.695	42	2720.125
18	754.6891	43	3051.35
19	843.3904	44	3060.555
20	865.9785	45	3070.828
21	915.2325	46	3131.659
22	957.5858	47	3137.704
23	992.3371	48	3153.044
24	1016.72	49	3156.107
25	1043.689	50	3466.78
		51	3652.621

Table 25: Frequencies molecule 3

TS 3			
atom	x	y	z
C	-1.20973	-1.58085	0.656279
N	-1.42621	-0.25506	0.537293
C	-0.56827	0.593314	-0.25746
C	0.806004	0.750406	0.352692
C	-1.21618	1.967799	-0.43698
H	-0.42551	0.126924	-1.23966
O	-0.25691	-2.12849	0.095406
H	-1.63696	-2.83116	2.31971
H	-2.20363	0.165846	1.022151
N	1.90673	0.921657	-0.30527
H	-0.58652	2.603308	-1.06295
H	-1.35817	2.458209	0.530865
H	-2.18833	1.848737	-0.92266
S	1.193499	0.824695	2.037375
H	1.907268	0.900618	-1.32512
H	2.422021	1.020027	0.989985
C	-2.18787	-2.34883	1.506886
H	-2.97555	-1.72005	1.925904
H	-2.63759	-3.13575	0.894655

Table 26: xyz coordinates molecule TS 3

TS 3 Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-1772.55	26	1092.187
2	46.0077	27	1122.467
3	49.1106	28	1207.936
4	62.952	29	1271.427
5	82.3974	30	1300.362
6	158.244	31	1308.466
7	208.643	32	1364.823
8	225.504	33	1377.804
9	254.0504	34	1388.722
10	350.9157	35	1438.674
11	378.775	36	1452.923
12	454.8445	37	1463.195
13	528.7933	38	1468.566
14	562.1374	39	1542.377
15	594.7349	40	1592.033
16	640.754	41	1693.262
17	659.4288	42	1791.619
18	691.3362	43	3056.608
19	718.006	44	3059.727
20	890.8077	45	3066.169
21	908.7471	46	3136.527
22	952.6429	47	3137.782
23	998.389	48	3156.731
24	1023.882	49	3159.215
25	1043.358	50	3502.901
		51	3652.917

Table 27: Frequencies molecule TS 3

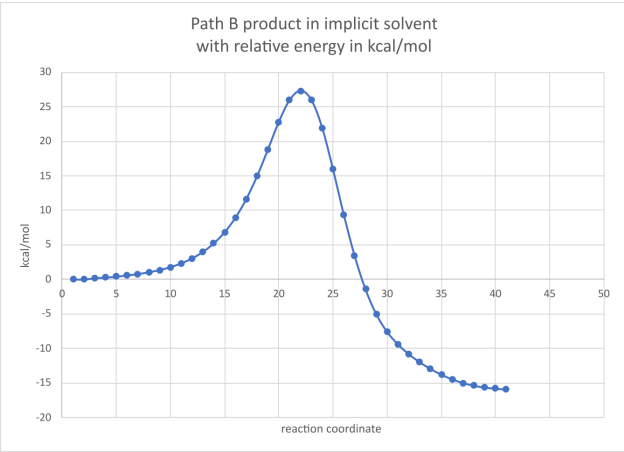


Figure 5: IRC of TS 3

4			
atom	x	y	z
C	-1.27577	-1.57545	0.557099
N	-1.45122	-0.24155	0.485499
C	-0.47806	0.625365	-0.12779
C	0.81041	0.74722	0.694044
C	-1.07911	2.014261	-0.34038
H	-0.20702	0.191799	-1.09748
O	-0.2845	-2.12402	0.066642
H	-1.91655	-2.79685	2.171938
H	-2.23384	0.17678	0.963129
N	1.909351	0.817322	-0.03431
H	-0.35434	2.673614	-0.82302
H	-1.3651	2.453632	0.620303
H	-1.96401	1.938989	-0.97805
S	0.798212	0.862011	2.35638
H	1.881784	0.716563	-1.03846
H	2.807506	0.965346	0.401135
C	-2.35056	-2.35406	1.269831
H	-3.20792	-1.73908	1.550115
H	-2.68117	-3.17053	0.622485

Table 28: xyz coordinates molecule 4

4			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	42.5966	26	1123.32
2	48.6355	27	1214.389
3	54.2574	28	1273.987
4	78.7091	29	1305.198
5	162.6637	30	1306.462
6	227.2811	31	1365.563
7	237.0587	32	1377.233
8	262.3813	33	1387.022
9	346.8341	34	1437.063
10	374.45	35	1453.32
11	447.6709	36	1457.418
12	461.5186	37	1459.682
13	533.8647	38	1476.264
14	563.102	39	1547.768
15	614.2463	40	1631.762
16	638.0336	41	1692.294
17	657.0876	42	3052.219
18	685.0822	43	3058.631
19	714.9585	44	3067.377
20	904.0345	45	3134.251
21	940.8097	46	3135.335
22	988.9184	47	3147.793
23	1013.828	48	3155.398
24	1041.909	49	3579.672
25	1052.084	50	3660.356
		51	3705.506

Table 29: Frequencies molecule 4

9.1.3 Explicit solvent

1			
atom	x	y	z
C	2.301806	0.392686	-1.17649
N	1.235026	0.356644	-0.32451
C	1.002512	-0.79261	0.528068
C	-0.10237	-1.60765	-0.01106
C	0.707629	-0.40962	1.981898
H	1.914573	-1.39848	0.46646
O	3.175295	-0.46609	-1.15395
H	2.112126	1.193685	-3.12815
H	0.463057	1.010383	-0.47538
N	-1.00065	-2.21281	-0.40612
H	0.585204	-1.31129	2.587783
H	-0.21086	0.179309	2.051914
H	1.543752	0.172323	2.376877
S	-1.22379	2.708665	-0.47315
H	-1.88209	1.884067	0.384838
H	-0.43571	3.206166	0.49822
O	-2.67248	0.572872	1.619676
H	-3.12523	-0.14034	1.121451
H	-3.3498	0.991973	2.156297
O	-3.70054	-1.44762	0.069062
H	-2.87612	-1.8844	-0.20981
H	-4.28212	-2.1484	0.373693
C	2.334441	1.565887	-2.12259
H	1.607212	2.338712	-1.86355
H	3.346684	1.976865	-2.14915
H	4.738258	0.070996	-1.99426
O	5.443786	0.590209	-2.41856
H	6.217004	0.022538	-2.42276
O	1.712461	2.927596	1.524923
H	2.53428	3.425445	1.504612
H	1.864079	2.151484	0.972994

Table 30: xyz coordinates molecule 1

1			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-6.888	45	843.05
2	17.671	46	923.32
3	30.093	47	991.19
4	36.393	48	1024
5	49.565	49	1048.2
6	52.136	50	1086.9
7	57.854	51	1131.4
8	66.351	52	1183.1
9	69.746	53	1230.6
10	78.709	54	1292.3
11	81.254	55	1314.4
12	97.832	56	1348.5
13	99.972	57	1389.4
14	111.26	58	1408
15	117.4	59	1453.2
16	137.37	60	1473.2
17	143.95	61	1475.1
18	156.45	62	1485.6
19	179.87	63	1562.8
20	194.41	64	1612.2
21	206.69	65	1629.9
22	213.08	66	1650.7
23	226.12	67	1666.5
24	228.5	68	1758.6
25	235.12	69	2372.9
26	241.79	70	2513.7
27	290.13	71	2692.3
28	317.26	72	3058.8
29	338.26	73	3058.8
30	371.26	74	3079.1
31	377.93	75	3135.5
32	405.96	76	3145.2
33	427.37	77	3154.1
34	435.1	78	3159.9
35	471.9	79	3384
36	553.21	80	3500.8
37	578.82	81	3641.6
38	590.59	82	3646.5
39	608.77	83	3809.3
40	645.8	84	3921.4
41	651.15	85	3928.3
42	657.19	86	3936.3
43	762.01	87	3939.8
44	793.3		

Table 31: Frequencies molecule 1

TS 4			
atom	x	y	z
C	2.305386	0.782426	-0.92763
N	1.510353	0.971611	0.172184
C	1.103586	-0.14825	0.988534
C	-0.21861	-0.69069	0.560865
C	1.089184	0.180048	2.480097
H	1.809175	-0.9628	0.778973
O	2.908084	-0.26714	-1.1167
H	2.90113	1.699079	-2.76463
H	0.794732	1.705178	0.13208
N	-0.95755	-1.54441	0.275297
H	0.784746	-0.69429	3.06174
H	0.392424	0.99804	2.678636
H	2.093413	0.483348	2.786487
S	-1.4788	1.724769	0.375728
H	-3.31745	0.969467	0.17209
H	-1.59451	2.087596	1.663247
O	-4.18219	0.452809	0.019785
H	-3.76166	-0.87913	-0.18131
H	-4.63701	0.889631	-0.70653
O	-3.33197	-1.84303	-0.27785
H	-2.1853	-1.7623	-0.01135
H	-3.7854	-2.45871	0.306873
C	2.426059	1.9876	-1.82633
H	1.444197	2.43445	-2.0112
H	3.04664	2.744029	-1.33256
H	4.032688	-0.271	-2.57887
O	4.568021	-0.07843	-3.36939
H	4.644486	-0.91287	-3.83661
O	3.877879	2.126369	1.866043
H	4.668913	1.820248	1.414113
H	3.143708	1.869152	1.294382

Table 32: xyz coordinates molecule TS 4

TS 4 Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-481.3	45	940.58
2	14.016	46	984
3	16.294	47	1023.8
4	23.049	48	1055.8
5	31.076	49	1081.3
6	36.925	50	1130.4
7	60.621	51	1183.4
8	61.702	52	1197.2
9	70.746	53	1283.3
10	88.923	54	1306
11	92.055	55	1334.2
12	102.04	56	1348.1
13	105.68	57	1394.3
14	118.87	58	1397
15	139.65	59	1449.8
16	154.91	60	1471.5
17	164.79	61	1474
18	186.84	62	1482.6
19	204.01	63	1531
20	220.6	64	1589.7
21	231.81	65	1604
22	248.94	66	1623.6
23	275.59	67	1652.3
24	298.97	68	1709.4
25	333.3	69	1762.8
26	363.86	70	2203
27	379.48	71	2276.3
28	391.4	72	2717.8
29	398.78	73	2842.8
30	421.53	74	3051.3
31	448.19	75	3059.1
32	467.1	76	3071.3
33	479.15	77	3122.1
34	549.04	78	3145
35	596.45	79	3154.4
36	606.85	80	3170.2
37	636.17	81	3356.9
38	650.35	82	3630.7
39	656.51	83	3781.7
40	675.85	84	3890.5
41	731.06	85	3898.5
42	777.45	86	3928.2
43	847.24	87	3936.7
44	931.94		

Table 33: Frequencies molecule TS 4

5			
atom	x	y	z
C	2.046617	1.291226	-0.73265
N	1.53736	0.356651	0.097557
C	1.189024	-0.98167	-0.34622
C	-0.29466	-1.01782	-0.70467
C	1.569359	-2.01056	0.709332
H	1.755615	-1.1607	-1.26664
O	2.21476	1.095345	-1.93821
H	3.508105	2.664204	-0.02961
H	1.192069	0.638906	1.021079
N	-1.14289	-1.61755	0.021002
H	1.258268	-3.01115	0.400633
H	1.093228	-1.78531	1.667434
H	2.653326	-1.99783	0.848568
S	-0.74575	-0.10681	-2.16247
H	-1.48301	1.580448	-0.5985
H	0.506292	0.375053	-2.40934
O	-1.79403	2.013295	0.215437
H	-3.17112	0.66498	0.781523
H	-2.08697	2.892176	-0.04136
O	-3.63379	-0.16046	0.98104
H	-2.11299	-1.4945	-0.27292
H	-3.18081	-0.49867	1.764771
C	2.416216	2.602575	-0.08099
H	2.069459	3.423109	-0.71323
H	2.008446	2.701225	0.927877
H	-1.23398	-1.58717	1.809556
O	-1.42769	-1.24278	2.715957
H	-1.43616	-1.99636	3.311027
O	-0.006	1.172434	2.347412
H	-0.45705	0.379856	2.686635
H	-0.65433	1.582774	1.754646

Table 34: xyz coordinates molecule 5

5			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	39.844	45	948.78
2	57.352	46	999.86
3	60.14	47	1023.9
4	72.81	48	1044.8
5	78.274	49	1052.8
6	87.082	50	1088.3
7	103.77	51	1149.2
8	109.91	52	1178.6
9	121.43	53	1300.3
10	125.2	54	1319.9
11	131.34	55	1343.6
12	152.72	56	1381
13	162.01	57	1390.5
14	172.6	58	1399.5
15	193.96	59	1454.9
16	204.98	60	1469.8
17	217.38	61	1475.9
18	221.46	62	1483
19	240.82	63	1606.7
20	252.11	64	1615.4
21	270.82	65	1623.7
22	290.27	66	1649.5
23	325.52	67	1686.3
24	343.65	68	1713.1
25	406.68	69	1720.6
26	416.46	70	2452.8
27	438.54	71	3058.5
28	456.07	72	3061.8
29	475.19	73	3084.3
30	480.17	74	3135.4
31	519.77	75	3142.9
32	554.81	76	3150.5
33	575.4	77	3156.9
34	593.21	78	3319.5
35	623.62	79	3347.1
36	635.5	80	3488
37	659.2	81	3651
38	684.72	82	3663.7
39	716.83	83	3743.4
40	747.03	84	3760.6
41	853.78	85	3830.6
42	860.5	86	3915.5
43	908.98	87	3923.3
44	927.56		

Table 35: Frequencies molecule 5

TS 5			
atom	x	y	z
C	-1.11878	1.956052	0.043294
N	-1.34649	0.662349	-0.21659
C	-1.02997	-0.39301	0.727949
C	0.423503	-0.85594	0.61008
C	-1.98374	-1.57038	0.544379
H	-1.12917	0.02515	1.732869
O	-0.71146	2.352484	1.143302
H	-2.20087	3.599152	-0.77121
H	-1.59134	0.396106	-1.16712
N	1.135076	-0.83117	1.67309
H	-1.74876	-2.35671	1.26631
H	-1.90016	-1.98883	-0.46357
H	-3.01521	-1.24253	0.701653
S	1.007714	-1.3734	-0.97306
H	3.906758	0.258247	-0.72006
H	2.224284	-0.24021	-0.87395
O	3.017136	0.626435	-0.69868
H	2.050025	-1.24837	1.496757
H	2.84557	1.092986	0.260314
O	2.579563	1.608067	1.511491
H	2.183697	2.520551	1.505661
H	1.908581	1.005756	1.887226
C	-1.40019	2.921865	-1.08287
H	-0.50481	3.526774	-1.25434
H	-1.68677	2.420116	-2.0096
H	-0.60972	-0.84883	-2.48468
O	-1.44965	-0.51707	-2.86442
H	-1.25864	-0.28821	-3.77711
H	0.455851	3.506356	1.316059
O	1.342679	3.944449	1.379438
H	1.313115	4.52189	2.145871

Table 36: xyz coordinates molecule TS 5

TS 5 Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-576	45	1003.8
2	30.804	46	1031.4
3	37.03	47	1052.9
4	52.814	48	1094.1
5	63.703	49	1107.2
6	76.612	50	1134.3
7	84.202	51	1192.2
8	92.237	52	1306.2
9	110.76	53	1317.6
10	118.24	54	1326.8
11	136.67	55	1334.4
12	155.8	56	1356.9
13	167.87	57	1393.1
14	174.24	58	1396.8
15	184.75	59	1457.6
16	219.35	60	1473.2
17	233.13	61	1476.1
18	243.72	62	1481.2
19	280.55	63	1528.5
20	295.58	64	1604.2
21	324.79	65	1616.4
22	372.85	66	1629.7
23	391.99	67	1645.3
24	401.84	68	1658.5
25	415.92	69	1702.6
26	433.01	70	1728.9
27	468.58	71	1996.3
28	492.69	72	3052
29	557.12	73	3058.7
30	579.09	74	3111.7
31	592.87	75	3134.9
32	597.64	76	3135.6
33	628.72	77	3141.3
34	638.36	78	3154.7
35	658.39	79	3205.6
36	684.86	80	3338.5
37	712.33	81	3475.7
38	730.88	82	3482.5
39	808.29	83	3535.2
40	872.8	84	3600.4
41	906.34	85	3881.7
42	928.89	86	3925.8
43	944.04	87	3926.5
44	999.72		

Table 37: Frequencies molecule TS 5

6			
atom	x	y	z
C	-1.2996	1.938227	0.465407
N	-1.49552	0.684082	0.024692
C	-1.25383	-0.48249	0.861598
C	0.246485	-0.80938	0.78687
C	-2.16902	-1.6245	0.460248
H	-1.46975	-0.17595	1.890855
O	-0.96507	2.197456	1.632757
H	-2.40238	3.611502	-0.22806
H	-1.40934	0.549385	-0.9871
N	0.986775	-0.00644	1.540187
H	-1.99028	-2.49723	1.092838
H	-1.99029	-1.92279	-0.57608
H	-3.21021	-1.30938	0.569285
S	0.85573	-1.99953	-0.20279
H	4.314937	-0.5006	-0.37555
H	2.826304	-0.96704	-0.30093
O	3.426763	-0.20021	-0.16701
H	1.994474	-0.00188	1.440672
H	2.468512	1.227759	-0.94649
O	1.803505	1.877977	-1.22923
H	1.709464	2.478186	-0.46999
H	0.536579	0.778687	1.999863
C	-1.5304	3.031358	-0.54621
H	-0.66597	3.700416	-0.54888
H	-1.6954	2.645054	-1.55339
H	0.048108	-0.33807	-2.54514
O	-0.26188	0.573584	-2.52908
H	0.477334	1.071457	-2.11321
H	0.549284	3.163629	1.498852
O	1.451372	3.362376	1.179481
H	1.652571	4.263249	1.44154

Table 38: xyz coordinates molecule 6

6			
Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	46.361	45	961.62
2	53.329	46	1000.3
3	57.178	47	1031.4
4	60.893	48	1060.9
5	67.134	49	1078.6
6	80.641	50	1155.1
7	86.452	51	1178.7
8	110.66	52	1319.5
9	120.55	53	1331
10	136.93	54	1338.9
11	160.38	55	1350.1
12	170.38	56	1393.1
13	171.66	57	1403.3
14	190.41	58	1456.3
15	201.52	59	1466.2
16	212.04	60	1474.4
17	226.69	61	1478
18	234.86	62	1481.8
19	244.42	63	1603.2
20	256.68	64	1608.5
21	261.49	65	1616.3
22	279.71	66	1658
23	332.96	67	1673.3
24	344.86	68	1695.5
25	400.78	69	1706.5
26	415.91	70	3061.1
27	426.7	71	3061.4
28	464.49	72	3082
29	498.92	73	3137.9
30	528.34	74	3143.9
31	564.08	75	3147.5
32	602.29	76	3164.9
33	634.83	77	3373.4
34	655.99	78	3437.4
35	667.27	79	3470.2
36	674.99	80	3497.8
37	697.81	81	3554.7
38	714.31	82	3626.8
39	719.07	83	3661.7
40	747.54	84	3704.2
41	765.32	85	3876.9
42	778.8	86	3920.9
43	849.08	87	3935.7
44	926.95		

Table 39: Frequencies molecule 6

9.2 Path A xyz coordinates, frequencies, and IRCs

9.2.1 Vacuum

TS 7			
atom	x	y	z
C	3.187345	-0.08953	-2.74259
C	2.470404	0.422666	-1.49503
H	2.065673	1.421068	-1.70885
N	3.307235	0.539393	-0.30395
C	1.211835	-0.41526	-1.22308
O	0.472899	-0.77963	-2.12165
S	-2.25041	-1.08859	4.394276
S	-1.92636	0.358769	2.963538
C	-1.30293	-0.18046	0.619587
O	-1.36635	0.883774	0.117523
C	-0.24796	-1.28042	0.51439
H	-0.61843	-1.8761	-0.33313
C	-0.08324	-2.15369	1.742469
N	0.985751	-0.68989	0.088934
H	3.635025	-1.07144	-2.54869
H	3.983487	0.601451	-3.03476
H	2.481376	-0.19843	-3.56867
H	3.999637	-0.20111	-0.27447
H	3.803816	1.420224	-0.28794
H	0.594805	-2.97674	1.502342
H	-1.0401	-2.56423	2.069873
H	0.331365	-1.57408	2.571623
H	1.619555	-0.25376	0.748188
H	-3.34837	-0.4511	0.111762
O	-2.78468	-0.95806	0.717456
H	-1.05237	-1.0418	5.003747
H	-2.85019	-0.50241	1.714437

Table 40: xyz coordinates molecule TS 7

Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-658.04	39	1010.678
2	19.7752	40	1080.29
3	31.7127	41	1101.022
4	41.4565	42	1155.521
5	63.5731	43	1161.812
6	79.5992	44	1205.864
7	96.6166	45	1219.249
8	134.7617	46	1242.261
9	161.2213	47	1300.579
10	177.5252	48	1316.55
11	208.7072	49	1347.63
12	221.1222	50	1384.787
13	236.8777	51	1392.568
14	240.3749	52	1406.534
15	262.4778	53	1434.489
16	267.408	54	1470.953
17	297.077	55	1472.615
18	356.486	56	1475.06
19	379.1104	57	1484.109
20	393.7826	58	1524.812
21	411.9136	59	1648.336
22	427.8972	60	1780.438
23	486.9971	61	1810.624
24	516.4843	62	1914.59
25	533.9011	63	2681.259
26	547.9886	64	3040.42
27	567.8031	65	3041.807
28	651.0494	66	3051.315
29	676.7409	67	3065.107
30	763.6095	68	3116.751
31	824.4008	69	3145.21
32	841.8004	70	3150.943
33	858.3506	71	3161.203
34	887.1232	72	3543.255
35	922.7846	73	3578.331
36	940.3136	74	3634.411
37	969.7371	75	3776.141
38	991.6829		

Table 41: Frequencies of transition state TS 7

9.2.2 Implicit solvent

atom	x	y	z
C	3.604484	0.475175	-0.29492
C	2.267462	0.793967	0.375527
H	2.119053	1.882615	0.360608
N	2.225825	0.232081	1.719085
C	1.144512	0.230024	-0.49279
O	0.822752	0.756852	-1.55232
S	-1.03051	2.094924	1.17819
S	-2.36726	0.868288	0.188285
C	-1.79141	-1.58277	-0.28501
O	-2.79263	-1.93349	-0.80062
C	-0.35911	-1.67625	-0.77728
H	-0.3935	-1.26352	-1.79011
C	0.020978	-3.16189	-0.81771
N	0.581232	-0.90623	-0.01857
H	3.774775	-0.60645	-0.31878
H	4.419963	0.943756	0.264331
H	3.619454	0.859569	-1.3174
H	3.103361	0.401575	2.197432
H	1.49376	0.676976	2.264696
H	0.048271	-3.58931	0.189664
H	1.013646	-3.25305	-1.26416
H	-0.69676	-3.72268	-1.41927
H	1.052473	-1.29723	0.792163
H	-0.998	-1.61761	1.745765
O	-1.8785	-1.63091	1.33207
H	-1.50124	1.965966	2.43142
H	-2.22427	-0.63978	1.34704

Table 42: xyz coordinates molecule TS 7

Index	Frequency (cm^{-1})	Index	Frequency (cm^{-1})
1	-443.126	39	1025.996
2	24.1571	40	1056.829
3	47.383	41	1087.501
4	66.8199	42	1113.056
5	71.5445	43	1161.095
6	86.6393	44	1203.316
7	97.9632	45	1222.468
8	128.4207	46	1268.634
9	141.9736	47	1290.368
10	182.3273	48	1340.956
11	210.305	49	1348.152
12	216.0433	50	1360.822
13	231.4661	51	1381.098
14	237.1058	52	1390.88
15	254.2674	53	1461.742
16	267.6864	54	1462.32
17	296.4529	55	1468.612
18	307.0154	56	1471.914
19	353.1413	57	1531.309
20	404.2239	58	1563.458
21	407.7271	59	1622.762
22	430.8052	60	1739.302
23	500.9168	61	1874.95
24	517.7192	62	2323.383
25	524.6274	63	2681.762
26	534.4295	64	3032.095
27	561.8479	65	3046.88
28	644.1937	66	3059.489
29	692.182	67	3107.603
30	755.1478	68	3122.515
31	839.7453	69	3142.25
32	846.9607	70	3148.427
33	873.2716	71	3162.324
34	883.2282	72	3523.227
35	890.4098	73	3527.486
36	929.8778	74	3609.524
37	971.7106	75	3743.6
38	1015.76		

Table 43: Frequencies molecule TS 7