

INSTITUTE OF GENERAL AND INORGANIC
CHEMISTRY

BULGARIAN ACADEMY OF SCIENCES



Sofia Slavova

SYNOPSIS OF DOCTORAL THESIS

For the Degree of Doctor of Philosophy

Mechanisms of prebiotic reactions based on formamide – *ab initio* modelling

Scientific supervisor: **Prof. DSc Venelin Enchev**

Sofia

2021

The PhD thesis contains 118 pages and includes 58 figures and 3 tables. The bibliography covers 218 titles. The results have been published in 3 scientific articles.

The PhD thesis was discussed and directed for defence by the Colloquium of IGIC – BAS on February 9, 2021.

PhD student is enrolled in independent training in the specialty "Theoretical Chemistry" at the Institute of General and Inorganic Chemistry - BAS. The research on the dissertation work was performed in the following laboratory: "Theoretical and Computational Chemistry" IGIC–BAS.

The defence will take place on 01.06.2021 from 14:00 in Sofia, Akad.G.Bonchev St. bld.11, fl.4, room 427, Conference Hall Boltzmann at a meeting of a scientific jury composed of:

1. Assoc. Prof. Dr. Ivelina Georgieva - IGIC, BAS (chair);
2. Prof. Dr. Natasha Trendafilova - IGIC, BAS;
3. Prof. DSc Sonya Ilieva - FCF, Sofia University (reviewer);
4. Prof. DSc Vasil Delchev - University of Plovdiv (reviewer);
5. Assoc. Prof. Dr. Diana Cheshmedzhieva - FCF, Sofia University.

The materials are available to those interested in the office of IGIC–BAS.

INTRODUCTION

When, where and how does life on Earth originate? These questions of the origin of life are among the greatest unsolved problems in the natural sciences and arouse interest among scientists. After the main publication *The Origin of Life* by Alexander Oparin, various experiments were conducted, which led to the accumulation of a significant amount of data over the last 100 years.

The presence of biomolecules plays a key role in the emergence of life on Earth. Formamide, which is one of the most abundant molecules in the universe, has been shown experimentally to be a starting molecule in the prebiotic synthesis.

The synthesis of formamide-based biomolecules catalysed or initiated by extra-terrestrial minerals is one of the most discussed in the field of prebiotic chemistry. Saladino *et al* showed that formamide-rich meteorites could deliver a wide range of biomolecules to Early Earth. The authors show that irradiation of formamide with fast protons used to simulate the solar wind in the presence of selected extra-terrestrial materials (samples from major meteorite classes) can lead to the synthesis of many major components of *primordial soup* during the early solar system. Using mass spectrometry as an analytical method, Saladino *et al* discovered nucleobases, amino acids, sugars, carboxylic acids and nucleosides. One of the most complex tasks of prebiotic synthesis is the study of a preferred reaction pathway leading to the formation of nucleobases, nucleotides, amino acids, sugars and molecules important for the origin of life.

Despite significant progress in the development of instrumental methods for the study of prebiotic reactions, for detailed understanding of reaction mechanisms, the calculation of energy barriers and the search for the most rational reaction pathways, a wider application of computer chemistry methods is needed. A powerful tool for modelling the mechanism of prebiotic chemical reactions are the *ab initio* methods, which provide reliable information about intermediates, energy barriers and transition states. This information makes it possible to choose the most preferred reaction pathway.

GOALS AND MAIN TASKS

The research provided in the PhD thesis is part of the answer to the fundamental question of the origin of life. The present work is focused on the development of unified theoretical approaches, which together with the experimental results take into account the possible influence of various factors on the origin of life. All investigations included in current work are fundamental.

The aim of the presented theoretical research is to find the reaction pathway and corresponding reaction mechanisms in which nucleobases and pterin can be formed from formamide without the presence of a catalyst.

The following **main tasks** are set within this study:

- Quantum chemical *ab initio* modelling of the mechanisms of prebiotic reactions with the participation of formamide;
- Search for plausible reaction pathways and the most rational mechanisms for reactions;
- Detailed study of intermediates, energy barriers and transition states of modelled reactions;
- Consideration of the effects of the environment on the reactions;
- Building a reaction route of Self-catalytic prebiotic reactions.

3 RESULTS

The theoretical research described in the PhD thesis was preceded by experimental studies conducted by our colleagues in the project DN09/7/2016. The experiments consisted of thermal or microwave heating of pure formamide under vacuum. Thermal heating was performed at different temperatures and different heating durations. The obtained compounds were proved by GC-MS spectral analysis.

As a result of the experiment, it was shown (Figure 3.1) that heating of pure formamide at 100 °C for two weeks, leads to the formation of urea and glycerol. Formamide heating at 120 °C for two weeks in addition to urea and glycerol produced small amounts of purine and adenine. When pure formamide was heated at 140 °C for one week, in addition to urea, which is in the largest amount, purine and adenine are obtained. Cytosine and glycine were formed in small amounts. A panel of 10 compounds was obtained after heating pure formamide at 140 °C for two weeks - urea, purine, purine nucleobases - adenine and hypoxanthine, pyrimidine nucleobases - cytosine and uracil, the amino acid glycine, (E)-urocanic acid and pterin. Pterin, also called "molecular fossil", was discovered for the first time in such experiments.

Heating pure formamide at different temperatures changes the colour of the sample. The formamide remained a colourless liquid after heating at 100 °C for 2 weeks, at 120 °C for 1 week and at 160 °C for 48 and 72 hours. When heated at 120 °C for 2 weeks and at 170 °C for 24, 48 or 72 hours, the colour of the samples is dark yellow. After heating at 140 °C for 1-2 weeks, or at 180 °C for 1-3 days, the colour of the samples is dark brown to black. The colour is probably due to thermal degradation of formamide or the formation of polymers of hydrogen cyanide, also known as HCN oligomers, which are heterogeneous, ranging in colour from yellow or orange to brown or black, depending on the degree of polymerization. The structures of these oligomers remain controversial. They are not fully characterized due to their complex composition and heterogeneous nature.

Experiments with heating pure formamide at 160 °C for 48 hours and in the presence of natural minerals or meteorites as catalysts were performed by Saladino, Di Mauro *et al.* They receive purine and pyrimidine nucleobases and amino acids.

Our results show that when pure formamide is heated at 160 °C for 48 hours, nucleobases are formed – purine, hypoxanthine, cytosine and uracil and the amino acid glycine and pterin. Purine is the compound that is produced in the largest amount, followed by urea, N-formyl glycine, cytosine and urocanic acid. After heating at 170 °C for 48 hours, adenine, glycerol and L-norvaline were obtained in addition to the previous compounds. Alanine and N-acetyl glycine were detected by heating at 180 °C for 48 hours. Purine, cytosine, adenine, urea, ethanolamine and glycerol were found when formamide was heated at 170 °C for 24 hours, while a total of nine compounds were found at 180 °C for 29 hours. Four of them, hypoxanthine, 2-amino-4,6-dihydroxy-pyrimidine, (E)-urocanic acid and pterin are in addition to the previous ones. After heating the formamide at 180 °C for 72 hours, a large panel of compounds was obtained – a total of nineteen. The amino acid alanine was found only by heating formamide at 180 °C for 48 and 72 hours. The amount of purine, urea, hypoxanthine, adenine, urocanic acid, pterin, norvaline, 2-amino-4,6-dihydroxy-pyrimidine and ethanolamine was lower than that obtained in 29 hours at the same temperature.

The main goal of this PhD thesis is to explain how purine and pyrimidine nucleobases and pterin could be formed from pure formamide alone without the presence of a catalyst, proposing reaction pathways and mechanisms of the corresponding reactions.

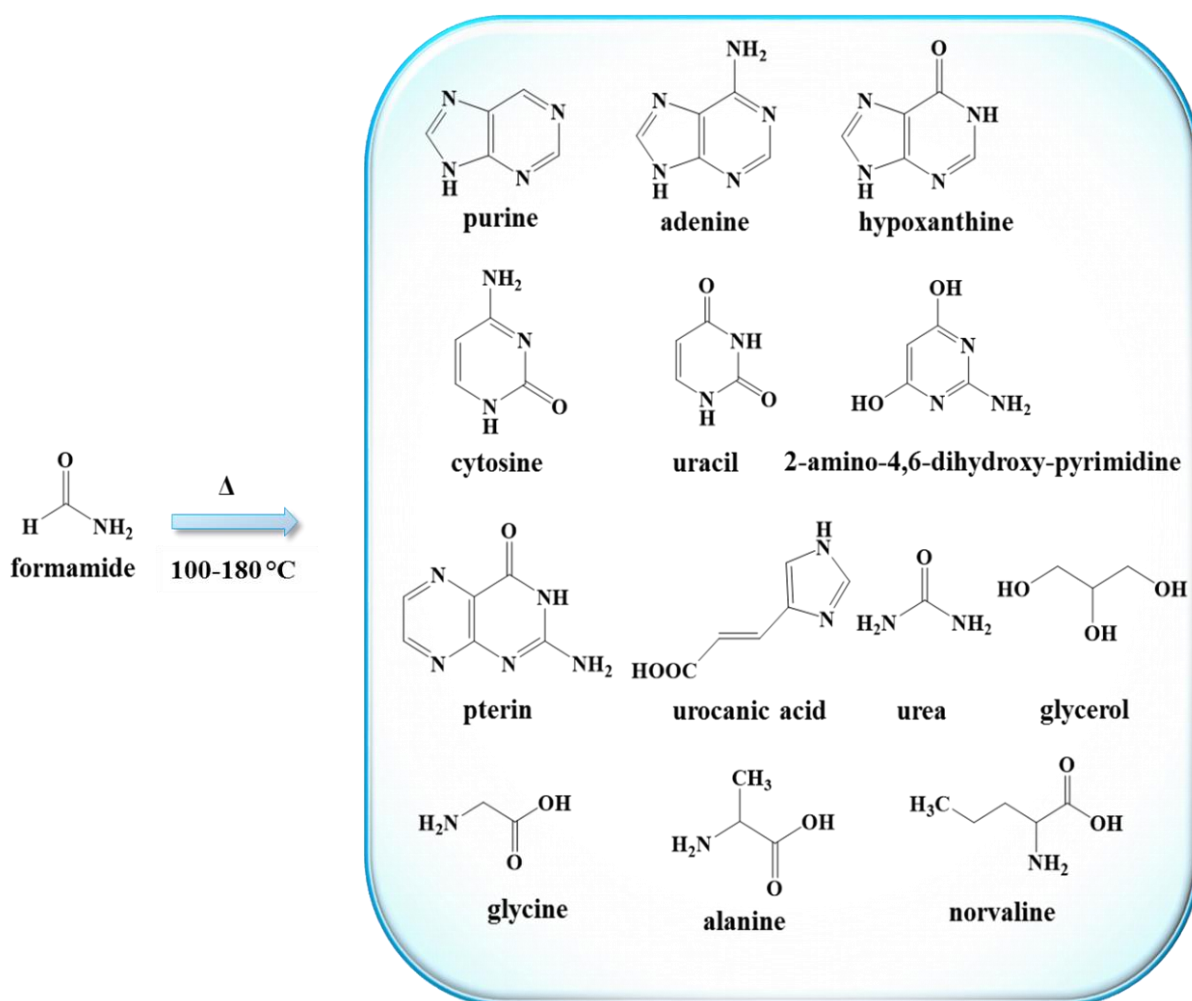


Figure 3.1 Main products after formamide heating at 100-180 °C without the presence of a catalyst.

3.1 THE FORMATION OF MAIN PREBIOTIC PRECURSORS

Ab initio quantum chemical calculations MP2 and SCS-MP2 were performed to create energetically possible reaction pathways for the synthesis of purines (adenine, hypoxanthine, guanine, isoguanine, xanthine and 2,6-diaminopurine) and pyrimidine (cytosine, uracil, isocytosine and 2,4-diaminopyrimidine) nucleobases, and pterin from formamide. All reaction pathways presented in the following sections are fully feasible in the temperature range in which formamide exists in the liquid phase.

The reaction route for the formation of main prebiotic precursors is shown in Figure 3.2. Generalized energy profile of some of these reactions is shown in Figure 3.10. The endothermic tautomeric conversion of formamide to formimidic acid is followed by strongly exothermic reaction of formation of hydrogen cyanide and water. The obtained water reacts with formamide and the exothermic reactions of dehydration, hydrolysis and decarbonylation of formamide produce hydrogen cyanide, water, ammonia, formic acid, and carbon oxide. Water, ammonia and formic acid play key role in the further prebiotic synthesis. The reaction between hydrogen cyanide and formamide leads to the formation of the key compounds, 2-iminoacetonitrile, precursor of 2-aminoacetonitrile, 2-aminomalnonitrile and isocyanic acid.

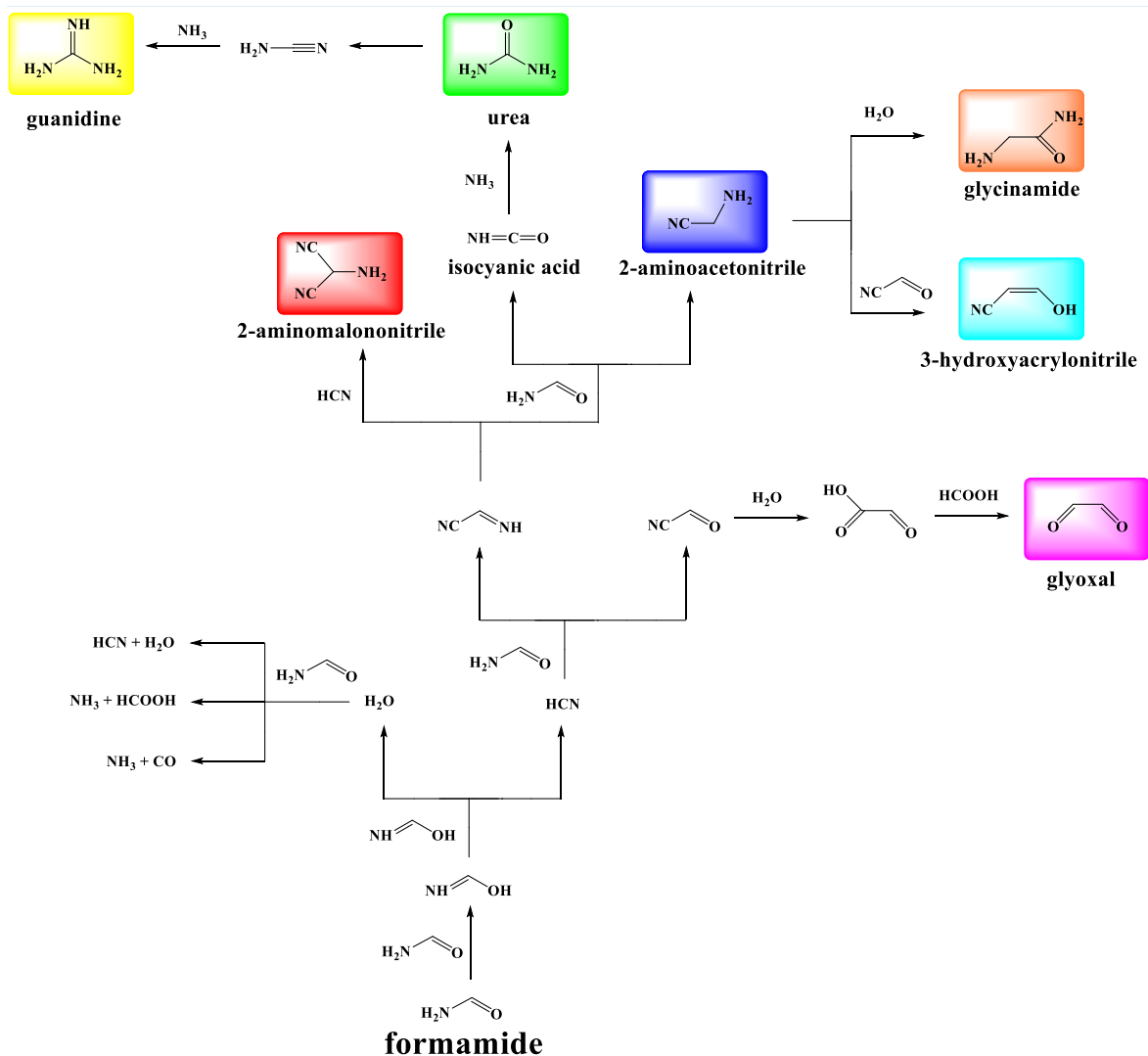


Figure 3.2 Reaction network of self-catalytic reactions for the formation of main formamide-based precursors.

The formation of 2-aminoacetonitrile and isocyanic acid from 2-iminoacetonitrile and formamide is thermodynamically preferred while 2-aminomalonalonitrile formation from 2-iminoacetonitrile and hydrogen cyanide is kinetically preferred because of lower energy barrier ($31.9 \text{ kcal mol}^{-1}$) but the reaction is endothermic.

The last precursor, urea, is formed from isocyanic acid and ammonia – the reaction is exothermic and it has very low barrier (6.2 kcal). The highest energy barrier for these reactions is calculated to be $40.3 \text{ kcal mol}^{-1}$. According to Das *et al*, chemical reactions occurring at temperatures of around 100°C have barriers in the region of $40.0 \text{ kcal mol}^{-1}$. This means that the above presented reactions in formamide should occur.

As a result of the reactions of thermal decomposition of formamide, presented on Figs. 2-4, the following compounds are formed: formimidic acid, hydrogen cyanide, water, formic acid, ammonia, carbon oxide, 2-iminoacetonitrile, isocyanic acid, urea and 2-aminoacetonitrile. A generalized energy profile of these reactions is shown on Fig. 5. The endothermic tautomeric conversion of formamide to formimidic acid is followed by strongly exothermic reaction of formation of hydrogen cyanide and water. The obtained water reacts with formamide and the exothermic reactions of dehydration, hydrolysis and decarbonylation of formamide produce hydrogen cyanide, water, ammonia, formic acid, and carbon oxide. Water, ammonia and formic acid play key role in the subsequent prebiotic synthesis. The reaction between hydrogen cyanide

and formamide leads to the formation of the key compound, 2-iminoacetonitrile, precursor of 2-aminoacetonitrile, 2-aminomalnonitrile and isocyanic acid. The formation of 2-aminoacetonitrile and the isocyanic acid from 2-iminoacetonitrile and formamide is thermodynamically preferred while 2-aminomalnonitrile formation from 2-iminoacetonitrile and hydrogen cyanide is kinetically preferred because of the lower energy barrier (31.9 kcal mol⁻¹) but the reaction is endothermic. The last precursor, urea, is formed from isocyanic acid and ammonia – the reaction is exothermic and it has very low barrier (6.2 kcal). The highest energy barrier for these reactions is calculated to be 40.3 kcal mol⁻¹. According to Das *et al* chemical reactions occurring at temperatures of around 100 °C have barriers in the region of 40.0 kcal mol⁻¹. That means that all above presented reactions in formamide should occur.

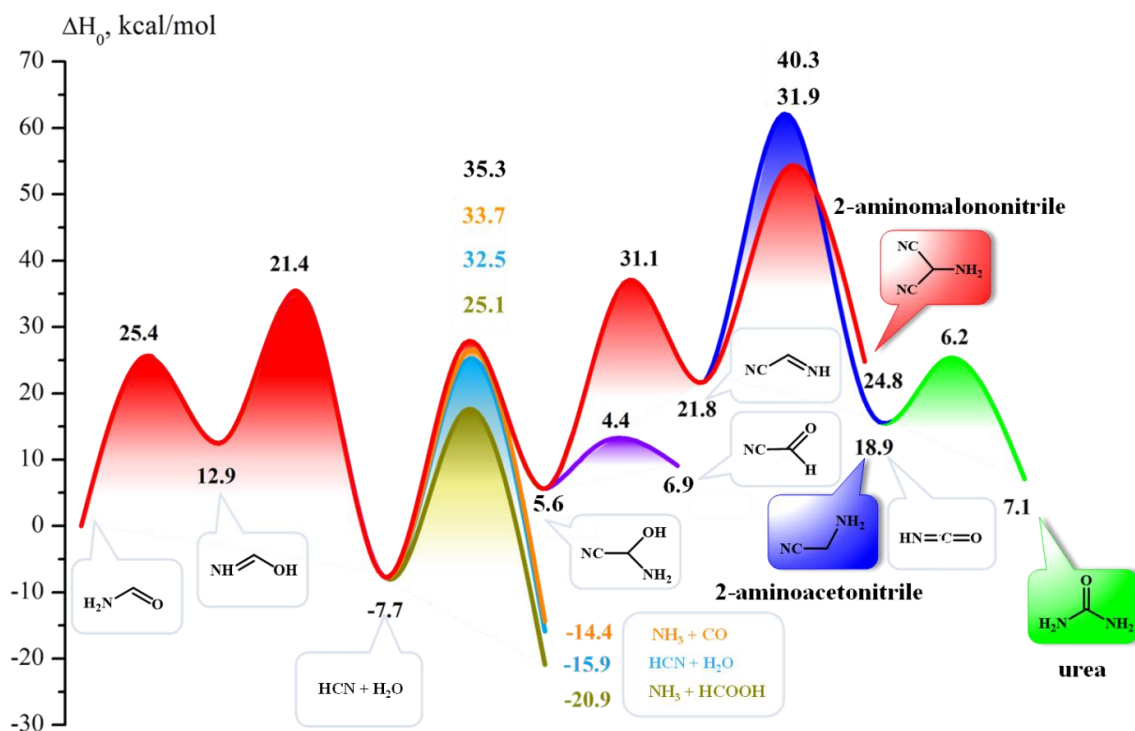


Figure 3.10 Schematic potential energy profiles along the reaction pathway of the main prebiotic precursors: 2-aminoacetonitrile (blue), isocyanic acid (blue), 2-aminomalnonitrile (red), urea (bright green) and formylcyanide (violet) from formamide. Different colours indicate reaction profiles of dehydration (light blue), hydrolysis (mustard) and decarbonylation (orange) of formamide. ΔH_0 is the zero-point energy corrected relative energy (in kcal mol⁻¹) calculated at SMD/SCS-MP2/cc-pVDZ level.

3.2 FORMATION OF PURINE NUCLEOBASES

Figure 3.11 shows a scheme for the formation of purine nucleobases (adenine, hypoxanthine, guanine, isoguanine, xanthine and 2,6-diaminopurine), starting from the main prebiotic precursors, 2-aminomalnonitrile, 2-aminoacetonitrile and urea. The formation of these precursors is shown in section 3.1.

One of the potential reaction pathways for adenine formation was proposed earlier by Leszczynski *et al*, but we suggest the existence of other alternative pathways. The proposed reaction pathway leading to purine formation (steps 1-10) is similar to the reaction pathway proposed by Leszczynski *et al*. This makes possible to compare the activation energy barriers of reactions 1-10, calculated at different theoretical levels. The reaction pathways for the formation of other nucleobases presented in Figure 3.11 were proposed for the first time.

Precursors, 2-aminoacetonitrile, 2-aminomalnonitrile, and urea, are raw material for formation of purine, adenine, hypoxanthine, 2,4-diaminopurine, xanthine, guanine and

isoguanine. They can be obtained by the formation of the intermediates *N'*-(1*H*-imidazol-5-yl)-formamidine and *N'*-(1*H*-imidazol-4-cyanide-5-yl)-formamidine.

Detailed reaction pathways for formation of *N'*-(1*H*-imidazol-5-yl)-formamidine (*IFI*) from 2-aminoacetonitrile and *N'*-(1*H*-imidazol-4-cyanide-5-yl)-formamidine (*ICNFI*) from 2-aminomalonnitrile are presented in Figure 3.12. *IFI* is precursor for purine formation (Figure 3.12), while *ICNFI*—for adenine and hypoxanthine (Figure 3.16).

Precursors, 2-aminoacetonitrile, 2-aminomalonnitrile, and urea, are raw material for formation of purine and purine bases.

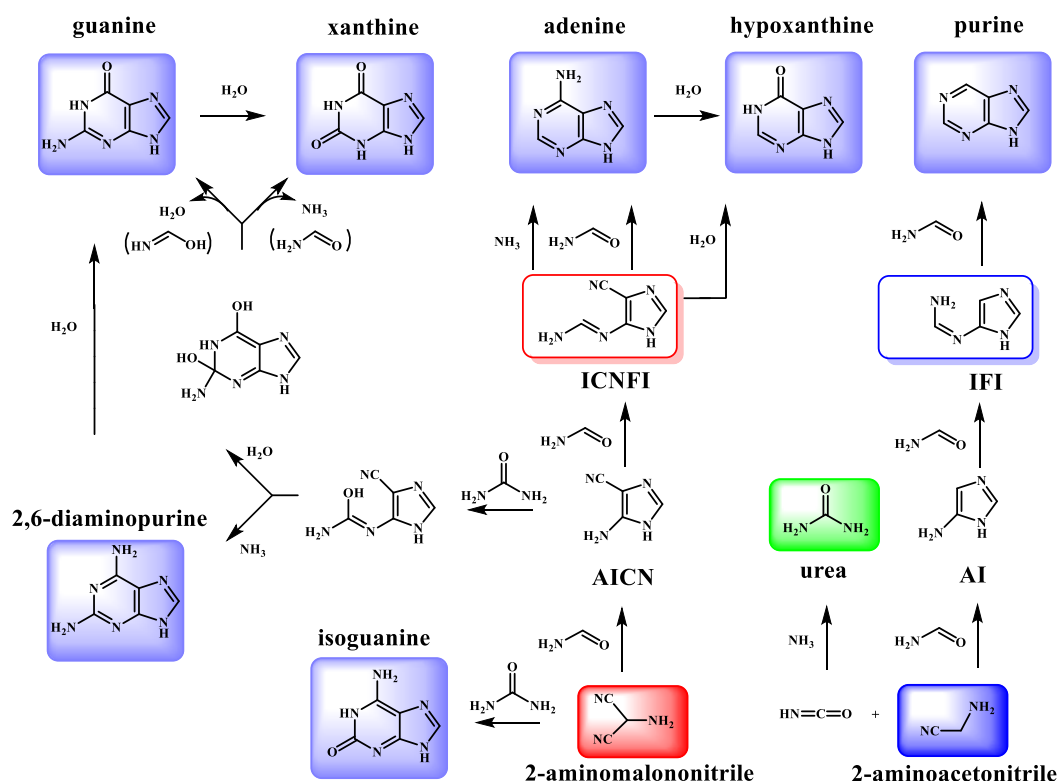


Figure 3.11 Reaction network of self-catalytic reactions for the formation of purine nucleobases.

3.2.1 Formation of 5-aminoimidazole and *N'*-(1*H*-imidazol-5-yl)formamidine

Detailed reaction profiles explaining the reaction pathways (Figure 3.12) are shown on Figure 3.13.

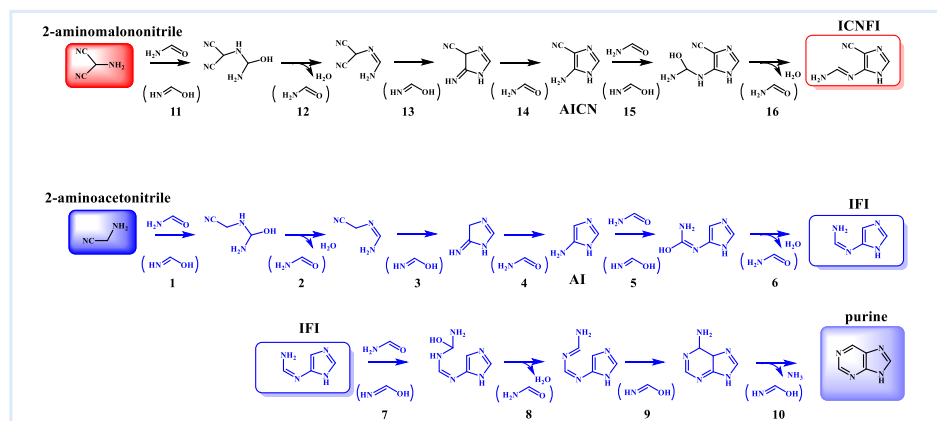


Figure 3.12 Detailed self-catalysed reaction pathways from 2-aminomalonnitrile to ICNFI (black) and from 2-aminoacetonitrile to IFI and purine (blue). For each reaction step reactant is depicted above the arrow, while

catalytic molecule – under the arrow in brackets; ΔH_0 is the zero-point energy corrected relative energy (in kcal mol⁻¹) calculated at SMD/SCS-MP2/cc-pVDZ level.

3.2.3 Route to purine

Both reaction pathways, from 2-aminoacetonitrile to **IFI** and from 2-aminomalononitrile to **ICNFI**, have the same rate-determined reaction step: five-membered ring-closure (Figure 3.15). Corresponding energy barriers are calculated to be 31.5 and 26.2 kcal mol⁻¹ (steps 3 and 13). However, in the first two steps leading to the formation of 5-iminoimidazoline, energy barriers are about 4 kcal mol⁻¹ lower than those leading to the 5-imino-4,5-dihydro-*1H*-imidazole-4-carbonitrile formation. The last two steps leading to formation of **IFI** (steps 5, 6) are also lower than those leading to **ICNFI** (steps 15, 16). The reaction pathway from 2-aminoacetonitrile to **IFI** should be kinetically preferred in comparison of that leading from 2-aminomalononitrile to **ICNFI**. This conclusion is consistent with the experimentally observed preferred formation of purine by heating of formamide because **IFI** is its precursor.

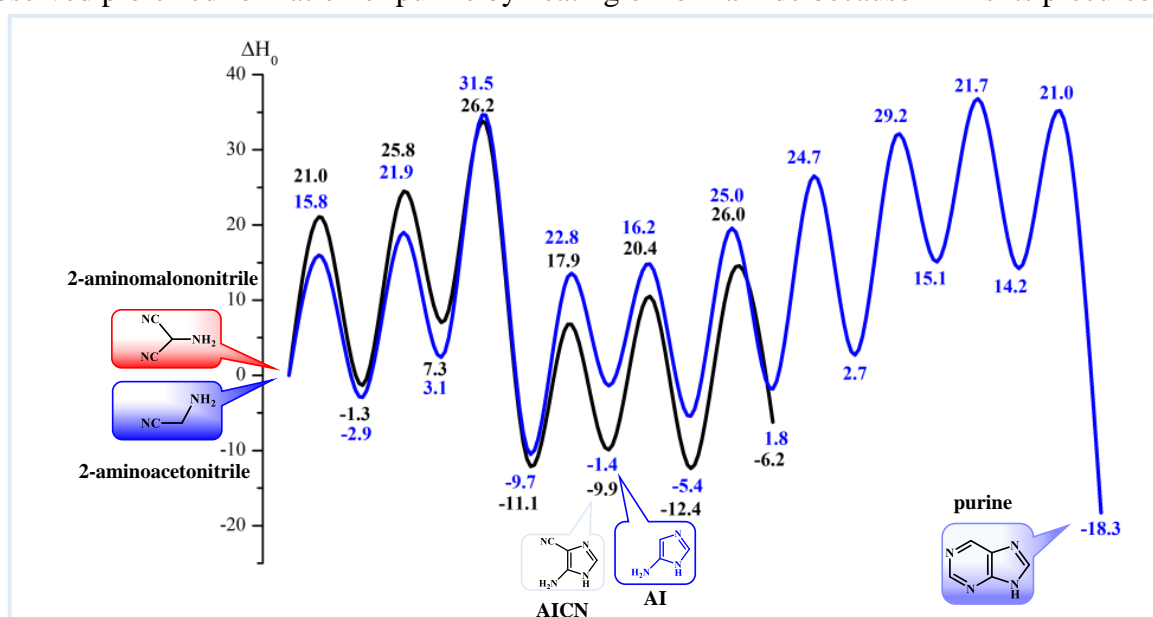


Figure 3.15 Schematic potential energy profiles along the reaction route from 2-aminomalononitrile to ICNFI (black) and from 2-aminoacetonitrile to IFI and purine (blue). For each reaction step reactant is depicted above the arrow, while catalytic molecule – under the arrow in brackets; ΔH_0 is the zero-point energy corrected relative energy (in kcal mol⁻¹) calculated at SMD/SCS-MP2/cc-pVDZ level.

3.2.4 Route to adenine

Two alternative reaction routes for adenine formation are possible (Figure 3.16). The first one begins with ammonolysis reaction and the second one is six-membered ring-closure of **ICNFI**.

The first reaction route consists of three steps: (i) addition of ammonia to the cyano group of the *N'*-(1*H*-imidazol-4-cyanide-5-yl)-formamidine; (ii) six-membered ring-closure and (iii) deamination and adenine formation. The second reaction route includes two steps: (i) ring-closure and (ii) tautomeric conversion.

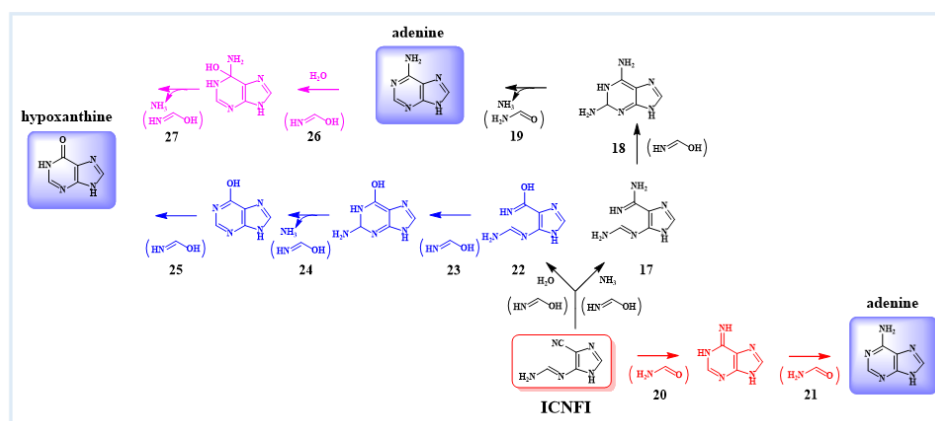


Figure 3.16 Detailed self-catalysed reaction pathways from *ICNFI* to adenine and hypoxanthine. For each reaction step reactant is depicted above the arrow, while catalytic molecule – under the arrow in brackets.

On the basis of ^{13}C - ^{15}N coupling experiments in purine and in adenine, Yamada and co-authors concluded that adenine carbons C4, C5, and C6 are derived from HCN, while formamide provides C2 and C8. Pyrimidine-based mechanism of purine formation, starting with the formation of formamide dimer, consumes two cyanide carbons: C4 and C5.

Springsteen *et al* and Leszczynski *et al* concluded that the reaction pathways for the formation of purine and adenine in which formamide, hydrogen cyanide, ammonia, and formic acid participate are interconnected. They suggest imidazole-first mechanism. Pathways to purine and adenine compete for 2-iminoacetonitrile. Formation of 2-aminoacetonitrile from 2-iminoacetonitrile is exothermic reaction but the calculated by us energy barrier is $8.4 \text{ kcal mol}^{-1}$ higher than that of 2-aminomalononitrile formation. The reaction routes from 2-aminomalononitrile to *ICNFI* and from 2-aminoacetonitrile to *IFI* and purine (Figure 3.15) also compete. Thus, an increased concentration of HCN favours the reaction pathway to adenine. The intermediates, 5-amino-imidazole-4-carboxamide and 5-formamido-imidazole-4-carboxamide in the reaction pathways to purine and adenine have been detected in mineral-doped formamide condensation reactions. In Springsteen *et al* experiments, containing low starting cyanide concentrations, the adenine fraction is diisotopically labelled – the carbon atoms C4 and C6 are derived from HCN, which is in line with the presented mechanism (Figure 3.18). In purine, formamide provides the carbon atoms C2, C5, C6 and C8.

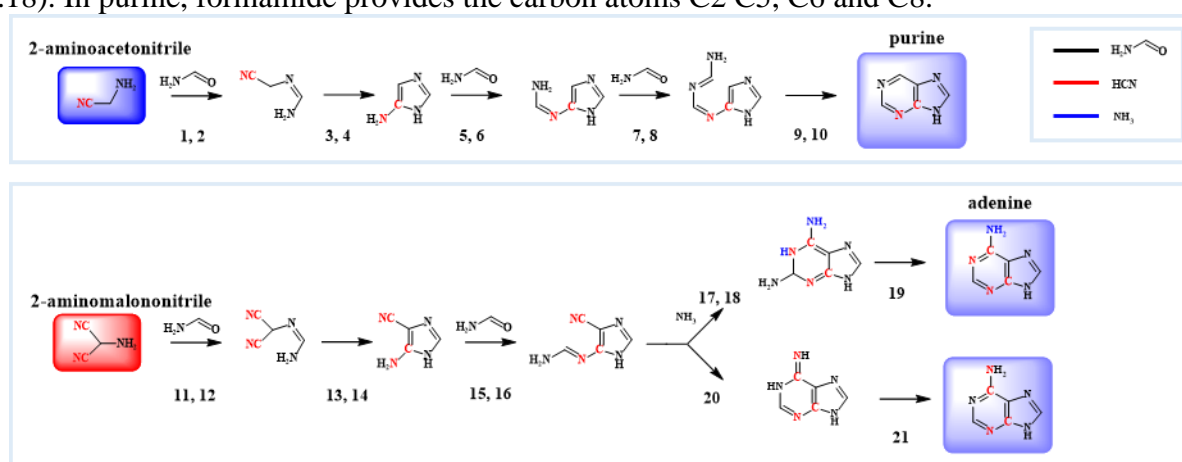


Figure 3.18 Route of purine and adenine formation from pure formamide. Formamide carbon, nitrogen and hydride atoms are black. Red labels track cyanide-derived carbon and nitrogen atoms, blue label tracks ammonia-derived nitrogen and hydride atoms, though hydrogen cyanide and ammonia are derived from formamide.

3.2.5 Route to hypoxanthine

Hypoxanthine formation (Figure 3.16) may occur in two alternative reaction routes: hydrolysis of *ICNFI* or deamination of adenine.

First reaction route, hydrolysis of *N'*-(1*H*-imidazol-4-cyanide-5-yl)-formamidine, leading to hypoxanthine formation includes four steps: (i) hydrolysis of the cyano group of the *N'*-(1*H*-imidazol-4-cyanide-5-yl)-formamidine; (ii) six-membered ring closure, (iii) deamination and (iv) tautomeric conversion. Since hypoxanthine is thermodynamically more stable than adenine,^[99] adenine can undergo easily deamination and be converted to hypoxanthine. Deamination of adenine is the second possible reaction route.

Summarizing, the reaction pathway started from *ICNFI* (Figure 3.20) presents 4 possible reaction routes: (i) ammonolysis of the cyano group (steps 17-19) or (ii) six-membered ring-closure in *ICNFI* (steps 20 and 21) leading to adenine formation, and (iii) hydrolysis of the cyano group in *ICNFI* (steps 22-25) or (iv) deamination of adenine (steps 26 and 27) leading to hypoxanthine formation. The activation energy barrier of ammonolysis reaction is calculated to be much lower (34.6 kcal mol⁻¹) than hydrolysis (41.1 kcal mol⁻¹) or ring-closure (43.8 kcal mol⁻¹). It is rate-controlling step for the adenine and hypoxanthine formation (Figure 3.20). On the other hand, describing two competitive processes – ammonolysis and hydrolysis, the reaction pathway of adenine formation is evaluated to be endothermic (5.7 kcal mol⁻¹) whereas pathway of hypoxanthine formation is strongly exothermic (-15.0 kcal mol⁻¹). In addition, deamination of adenine (step 26 and 27) is another exothermic reaction pathway for hypoxanthine generation (-6.6 kcal mol⁻¹ as is shown in Figure 3.20).

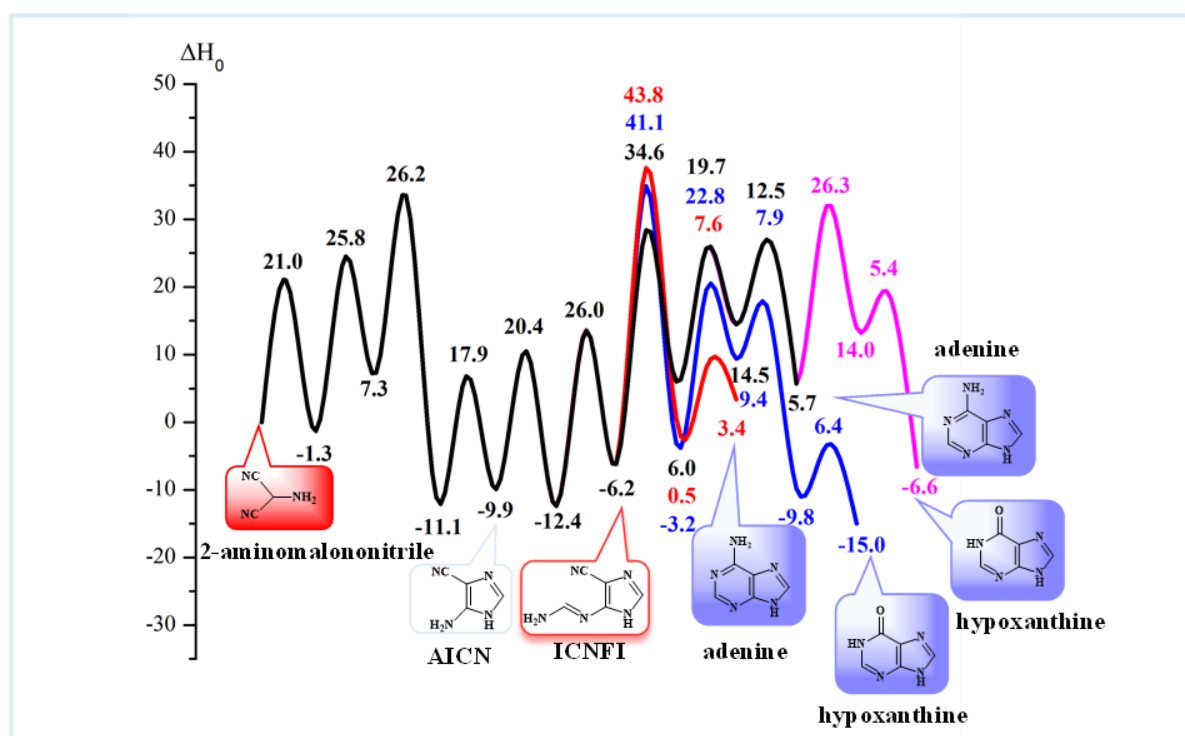


Figure 3.20 Schematic potential energy profiles along the reaction routes from 2-aminomalononitrile to adenine (black and red lines) and hypoxanthine (blue and magenta lines). ΔH_0 is the zero-point energy corrected relative energy (in kcal mol⁻¹) calculated at SMD/SCS-MP2/cc-pVDZ level.

3.2.6 Route to 2,6-diaminopurine, guanine and xanthine

As we noted above, urea can be formed as a result of the thermal decomposition of formamide (Figs. 1 and 4). This key reactant can interact with *AICN* generating *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid (Figure 3.11) which is starting compound for formation of 2,6-diaminopurine (steps 30-32), guanine (steps 33-36) and xanthine (steps 33, 34, 37-39) as

is shown on Figure 3.21. The *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid is formed from *AICN* and urea in two steps. The first one is the formation of 1-(4-cyano-1*H*-imidazol-5-yl)urea and ammonia and the second one is the formation of *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid.

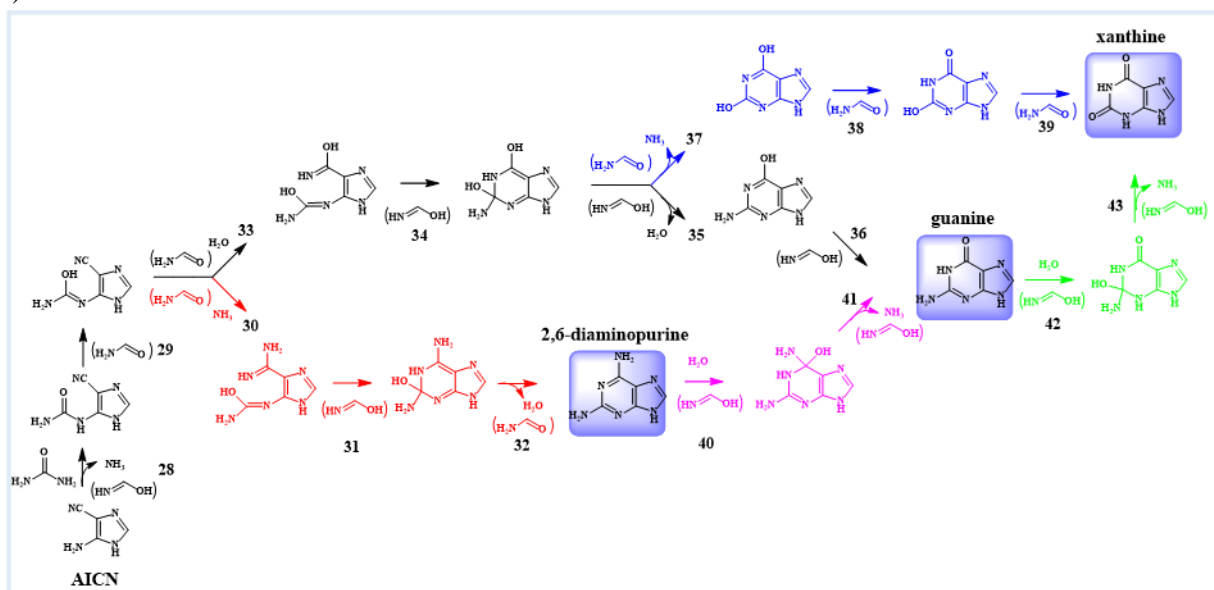


Figure 3.21 Detailed self-catalyzed reaction pathways from *AICN* to 2,6-diaminopurine, guanine and xanthine. For each reaction step reactant is depicted above the arrow while catalytic molecule – under the arrow in brackets.

The reaction route from *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid to 2,6-diaminopurine consists of three steps: (i) ammonolysis of the cyano group in 4-position in the imidazole fragment of the *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid; (ii) six-membered ring closure; (iii) dehydration and 2,6-diaminopurine formation.

The reaction route from *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid to guanine consists of four steps: (i) hydrolysis of the cyano group in 4-position in the imidazole moiety of the *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid; (ii) six-membered ring closure; (iii) dehydration and formation of guanine tautomer, and (iv) tautomeric conversion to guanine.

Guanine could also be produced by deamination reaction of 2,6-diaminopurine.

The complete energy profiles of the reaction routes from *AICN* to 2,6-diaminopurine, guanine and xanthine are illustrated on Figure 3.23. Interaction between *AICN* and urea leads to the formation of *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid. This compound is a key precursor for the further 2,6-diaminopurine, guanine and xanthine reaction pathways. Hydrolysis of its cyano group gives opportunity for guanine or xanthine generation, whereas ammonolysis of the same group – to formation of 2,6-diaminopurine. However, the energy barrier of hydrolysis, 38.8 kcal mol⁻¹, is considerably higher than that of the ammonolysis reaction, 27.1 kcal mol⁻¹. Generally, the reaction pathway for guanine (black line) is slightly exothermic (-5.2 kcal mol⁻¹) as is seen from Figure 3.23. Xanthine reaction pathway (blue line) is impeded since it is strongly endothermic (34.8 kcal mol⁻¹).

On the other hand, following the ammonolysis pathway of *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid, two elementary steps are needed to produce guanine (magenta line) from 2,6-diaminopurine. Considering this route, guanine reaction pathway is distinctly exothermic (-11.7 kcal mol⁻¹) in comparison to the one discussed above (-5.2 kcal mol⁻¹). Lastly, the deamination of guanine (green line) is energetically preferable pathway to xanthine formation (-39.0 kcal mol⁻¹ against 34.8 kcal mol⁻¹, see Figure 3.23). Overall, ammonolysis of *N'*-(4-cyano-1*H*-imidazol-5-yl)carbamimidic acid leads to energetically favored (Figure 3.23, red, magenta and green lines) reaction pathways for the further guanine and xanthine formation.

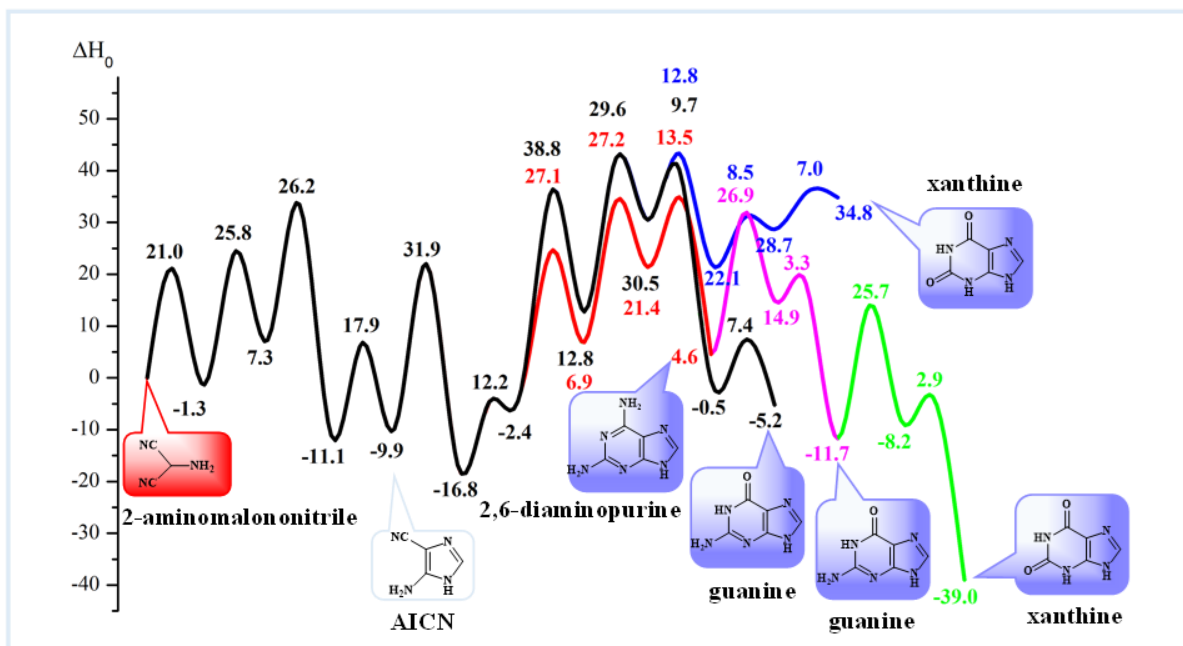


Figure 3.23 Schematic potential energy profile along the reaction route from 2-aminomalononitrile to guanine (black and magenta lines), 2,6-diaminopurine (red line), and xanthine (blue and green lines). ΔH_0 is the zero-point energy corrected relative energy (in kcal mol^{-1}) calculated at SMD/SCS-MP2/cc-pVDZ level.

3.2.7 Route to isoguanine

Two key reactants, 2-aminomalononitrile and urea, obtained by thermal decomposition of formamide may interact to generate isoguanine as it is shown on Figure 3.11. Detailed reaction route is presented on Figure 3.24. Reaction route includes 7 steps as last of them gives possibility for formation of two tautomeric forms of isoguanine. The formation of isoguanine take place in two stages: (i) formation of pyrimidine ring and (ii) additional formation of five-membered imidazole ring.

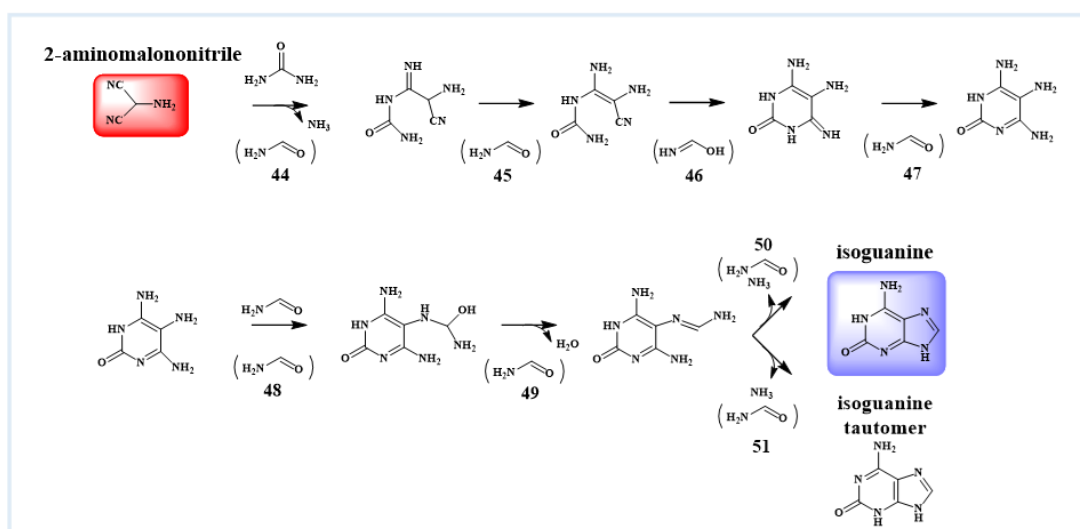


Figure 3.24 Detailed self-catalyzed reaction pathways from 2-aminomalononitrile and urea to isoguanine. For each reaction step reactant is depicted above the arrow, while catalytic molecule – under the arrow in brackets.

The first step of the suggested reaction pathway (Figure 3.26) of the reaction between urea and 2-aminomalononitrile is rate-determining step ($37.8 \text{ kcal mol}^{-1}$) while activation energies for other steps are in the range from 6 to 31 kcal mol^{-1} . Described reaction pathway of isoguanine formation can be defined as totally endothermic but as a consequence of the strongly

exothermic character of the last step ($-42.0 \text{ kcal mol}^{-1}$), the whole reaction pathway is revealed to be slightly endothermic ($4.3 \text{ kcal mol}^{-1}$).

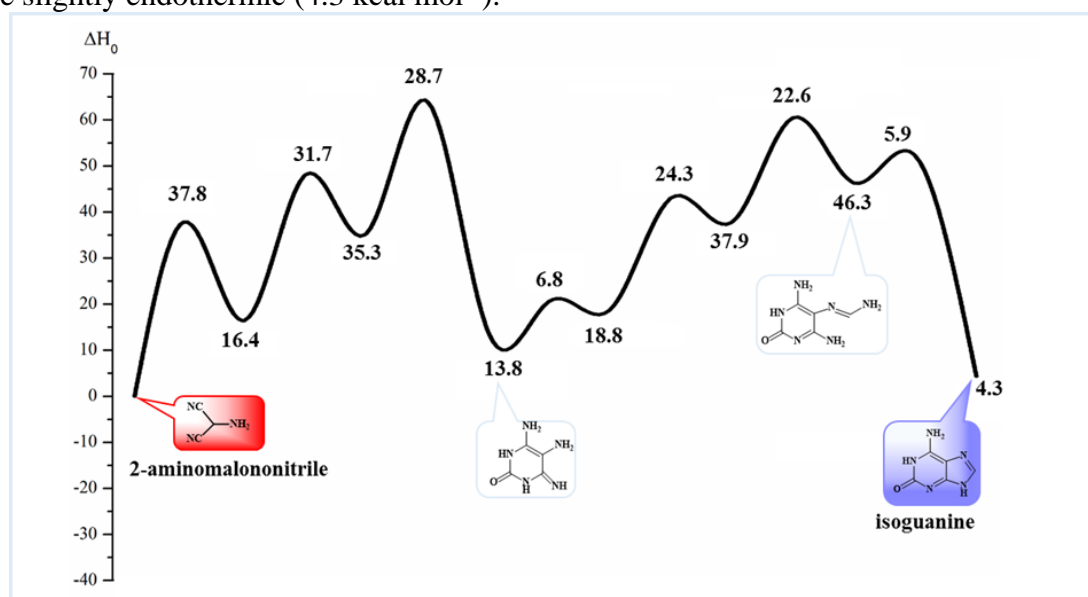


Figure 3.26 Schematic potential energy profile along the reaction route from 2-aminomalalonitrile and urea to isoguanine. ΔH_0 is the zero-point energy corrected relative energy (in kcal mol^{-1}) calculated at SMD/SCS-MP2/cc-pVDZ level.

3.2.8 Hypoxanthine formation from urea and glycinamide

A common reaction pathway for the formation of hypoxanthine from urea is shown in the work of Lagoja and Herdewijn. The preparation of hypoxanthine is carried out in 2 stages. First, the urea is dissolved in formic acid and the mixture is left for 2 weeks at room temperature. The resulting dipformylurea was then heated together with glycinamide at $175 \text{ }^\circ\text{C}$ for 2 hours to the formation of hypoxanthine.

In this section, a mechanism for the formation of hypoxanthine is proposed in accordance with the concept discussed so far, based on the prebiotic chemistry of formamide. A general scheme of the reaction pathway is presented in Figure 3.27.

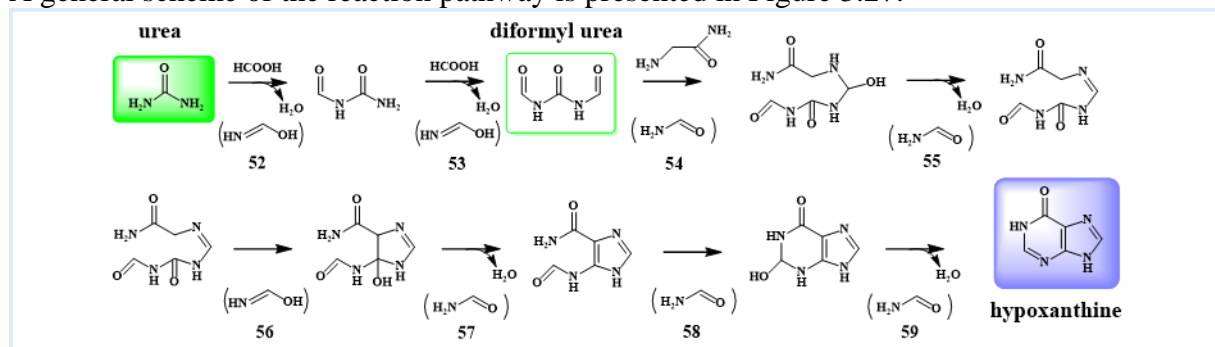


Figure 3.27 Detailed self-catalytic reaction pathways from urea and glycinamide to hypoxanthine. For each reaction step reactant is depicted above the arrow while catalytic molecule – under the arrow in brackets.

The energy profile of the reaction pathway for the formation of hypoxanthine from urea is shown in Figure 3.29. The five-membered ring closure reaction (step 56) is rate-determining ($44.0 \text{ kcal mol}^{-1}$). In general, the reaction pathway is exothermic, but due to the highly endothermic ($21.5 \text{ kcal mol}^{-1}$) reaction of the six-membered ring closure (step 58), the formation of hypoxanthine is a weakly exothermic process ($-1.6 \text{ kcal mol}^{-1}$).

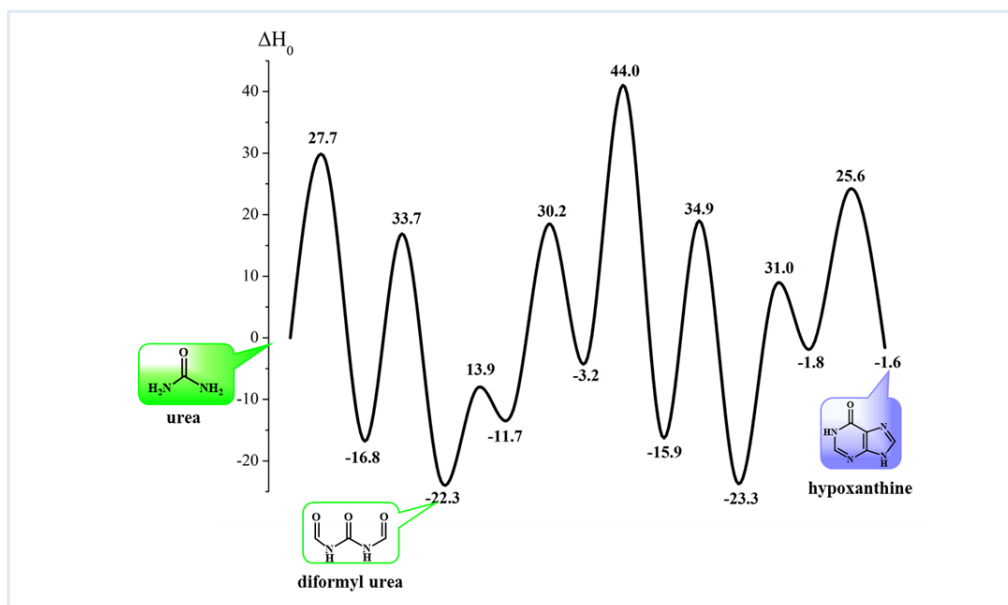


Figure 3.29 Schematic potential energy profile along the reaction route from urea to hypoxanthine. ΔH_0 is the zero-point energy corrected relative energy (in kcal mol⁻¹) calculated at SMD/SCS-MP2/cc-pVDZ level.

To identify the origin of the nitrogen atoms in hypoxanthine, Lagoja and Herdewijn synthesized hypoxanthine with ¹⁵N-labeled urea. The proton NMR experiment showed that the N3- and N9-atoms of hypoxanthine are derived from urea. As can be seen on Figure 3.30, the reaction pathway proposed here is in line with the experimental results, which shows its correctness.

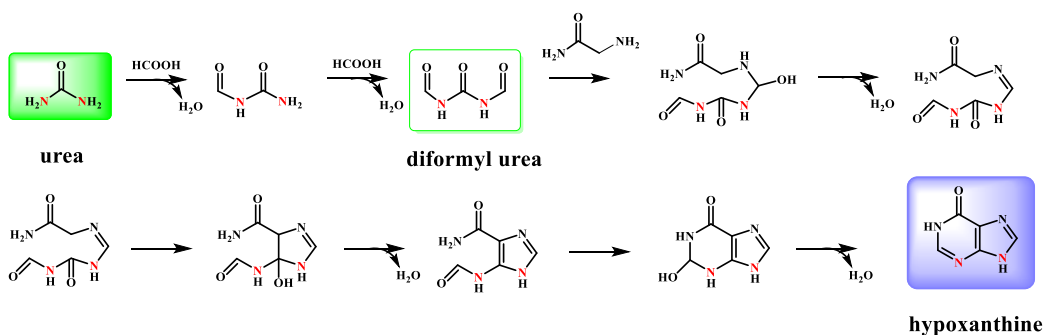


Figure 3.30 The reaction pathway from urea and glycinamide to hypoxanthine. Nitrogen atoms derived from urea are coloured in red.

3.2 THE FORMATION OF PYRIMIDINE NUCLEOBASES

The general scheme for the formation of pyrimidine nucleobases, cytosine, uracil, 2,4-diaminopyrimidine and isocytosine is shown in Figure 3.31. A reaction pathway starting with 3-hydroxyacrylonitrile, a cyanoacetaldehyde tautomer, has been proposed.

3-Hydroxyacrylonitrile and urea are starting compounds for the formation of cytosine and uracil. The other possible reaction pathway from 3-hydroxyacrylonitrile and guanidine lead to 2,4-diaminopyrimidine, isocytosine, cytosine and uracil.

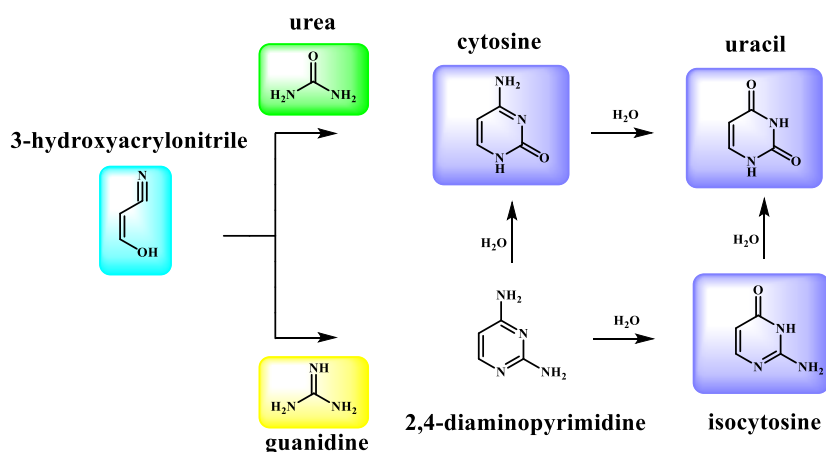


Figure 3.31 The reaction pathway of self-catalytic reactions for the formation of pyrimidine nucleobases.

3.3.1 The formation of cytosine and uracil

A detailed scheme of the reaction pathway for the formation of cytosine and uracil from 3-hydroxyacrylonitrile is shown in Figure 3.32.

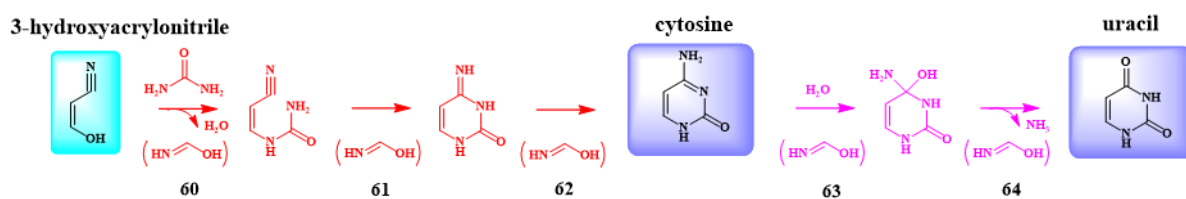


Figure 3.32 Self-catalytic reaction pathways from 3-hydroxyacrylonitrile and urea to cytosine (steps 60-62) and uracil (steps 63, 64). For each reaction step reactant is depicted above the arrow while catalytic molecule – under the arrow in brackets.

The potential energy profile of the reaction pathway to produce cytosine and uracil from 3-hydroxyacrylonitrile is shown in Figure 3.34. The first reaction (step 60), the reaction of urea and 3-hydroxyacrylonitrile, was rate-determining ($39.7 \text{ kcal mol}^{-1}$). The reaction pathway at all stages is exothermic, making the formation of cytosine ($-65.7 \text{ kcal mol}^{-1}$) and uracil ($-82.1 \text{ kcal mol}^{-1}$) thermodynamically very advantageous. This shows that the formation of cytosine and uracil is quite possible even under milder conditions.

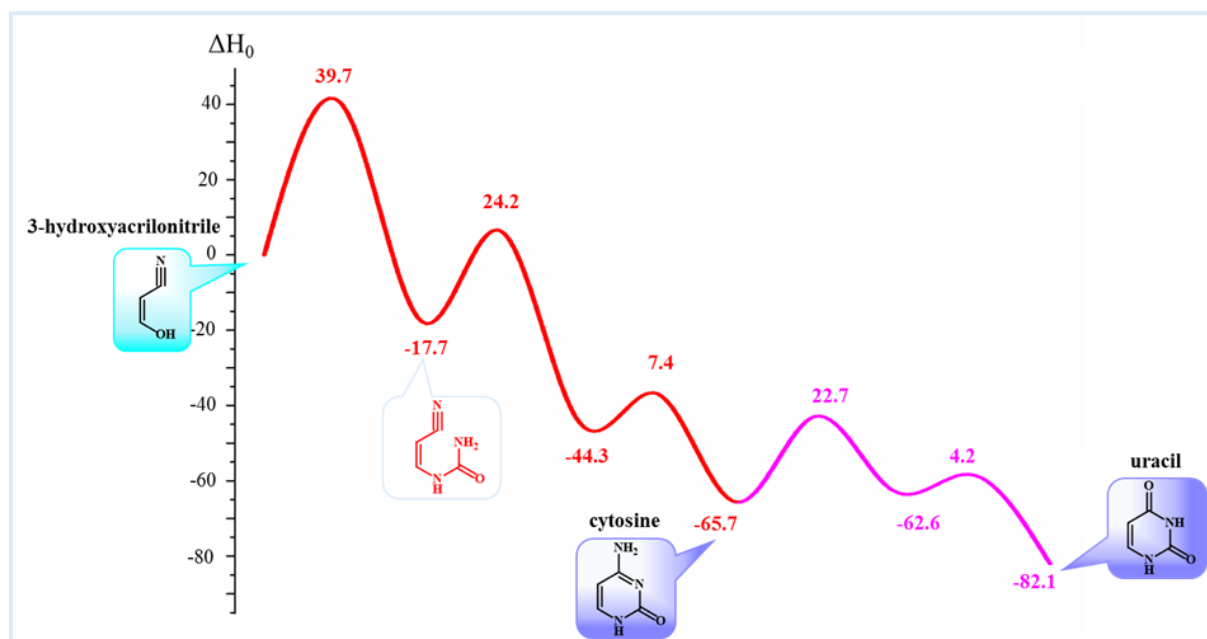


Figure 3.34 Energy profile of the reaction pathway for the formation of cytosine (red) and uracil (purple) from 3-hydroxyacrylonitrile. ΔH_0 is the zero-point energy corrected relative energy (in kcal mol⁻¹) calculated at SMD/SCS-MP2/cc-pVDZ level.

3.3.2 The formation of isocytosine, uracil and cytosine

3-Hydroxyacrylonitrile can also be a prebiotic precursor for the preparation of 2,4-diamino-pyrimidine (Figure 3.35, steps 65-68), isocytosine (Figure 3.35, steps 69, 70), uracil (Figure 3.35, steps 71, 72) and cytosine (Figure 3.35, steps 73, 74).

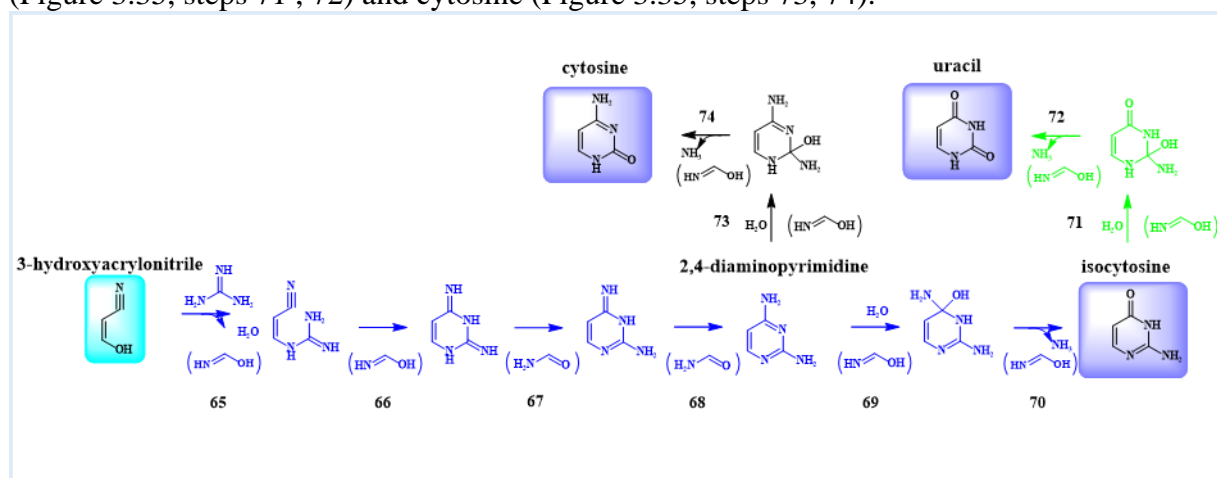


Figure 3.35 Self-catalytic reaction pathways from 3-hydroxyacrylonitrile and guanidine to isocytosine (steps 65-70), uracil (steps 71, 72) and cytosine (steps 73, 74). For each reaction step reactant is depicted above the arrow while catalytic molecule – under the arrow in brackets.

The first step (step 65) in the reaction pathway of 3-hydroxyacrylonitrile shown on Figure 3.37 is rate-determining (energy barrier is 38.5 kcal mol⁻¹). Hydrolysis of 2,4-diaminopyrimidine (step 73, black line, cytosine formation) is more preferred from a kinetic point of view than hydrolysis of step 69, which results in isocytosine with 5.1 kcal mol⁻¹. All reactions in a given reaction pathway are exothermic (-41.2 kcal mol⁻¹ for isocytosine, -43.1 kcal mol⁻¹ for cytosine and -65.4 kcal mol⁻¹ for uracil), which makes the formation of pyrimidine nucleobases very advantageous from a thermodynamic point of view.

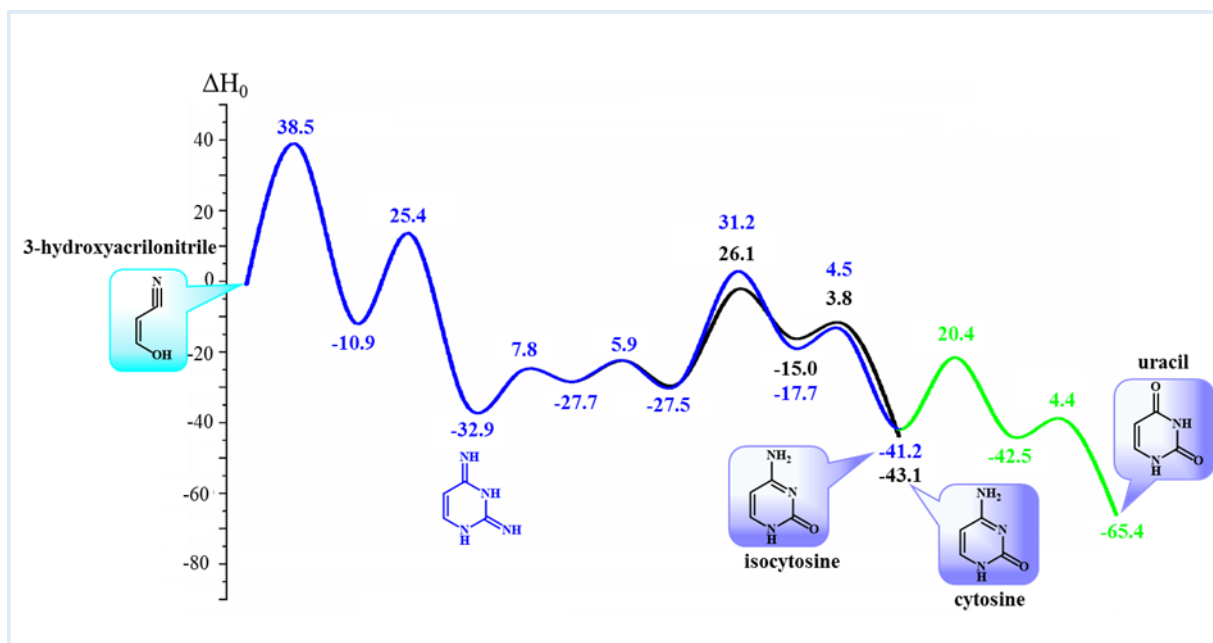


Figure 3.37 The potential energy profile of the reaction pathway from 3-hydroxyacrylonitrile to isocytosine (blue), cytosine (black) and uracil (green). ΔH_0 is the zero-point energy corrected relative energy (in kcal mol⁻¹) calculated at SMD/SCS-MP2/cc-pVDZ level.

3.4 THE FORMATION OF PTERIN AND GUANINE

A detailed scheme of the reaction pathway for the formation of pterin and guanine is presented on Figure 3.38.

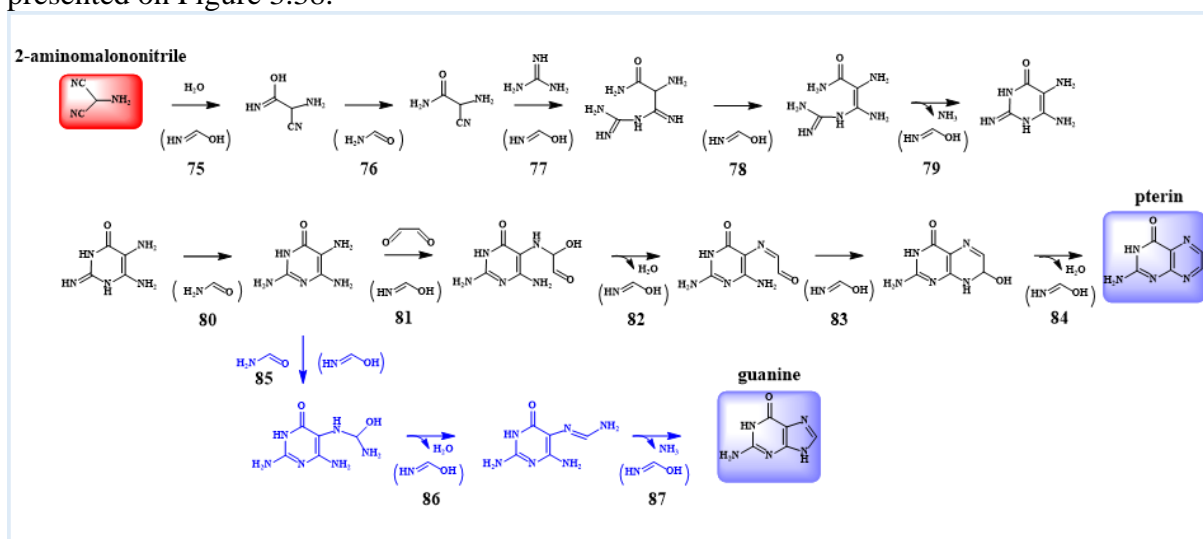


Figure 3.38 Self-catalytic reaction pathways from 2-aminomalonnitrile to pterin and guanine. For each reaction step reactant is depicted above the arrow while catalytic molecule – under the arrow in brackets.

The reaction pathways for pterin and guanine formation (steps 75-87) and the corresponding energy profiles are presented in Figure 3.40. In both cases, the formation of pterin and guanine is thermodynamically preferred, as the reaction profiles show that both reaction pathways are highly exothermic, -46.7 kcal mol⁻¹ and -43.6 kcal mol⁻¹, respectively. The formation of 2,3-diamino-3-guanidinoacrylamide (step 79) is a rate-determining reaction and has an energy barrier of 36.6 kcal mol⁻¹, which means that from a kinetic point of view these two pathways are very preferred.

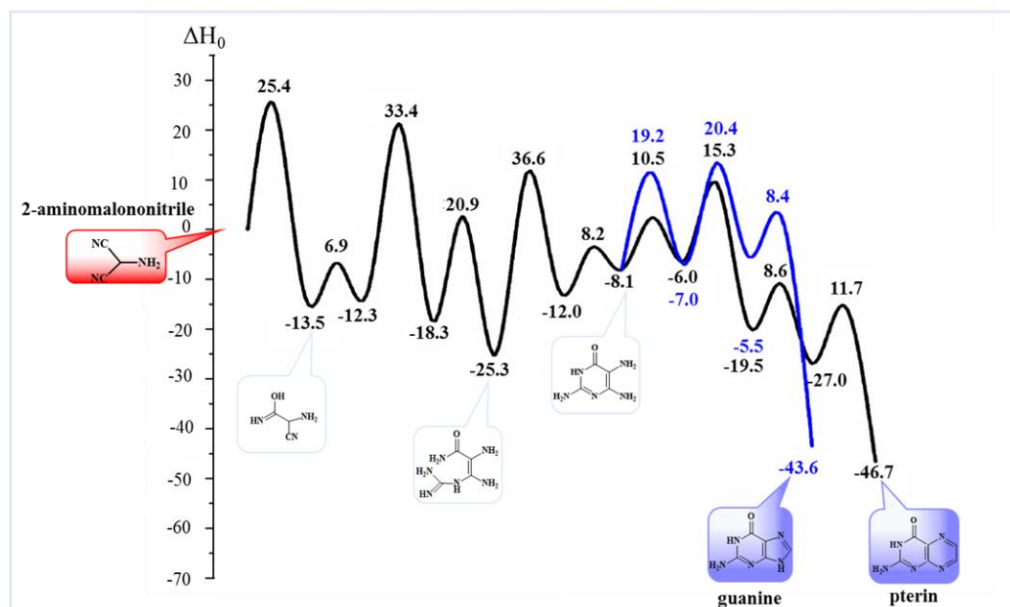


Figure 3.40 The potential energy profile along the reaction pathway from 2-aminomalononitrile to pterin (black line) and guanine (blue line). ΔH_0 is the zero-point energy corrected relative energy (in kcal mol⁻¹) calculated at SMD/SCS-MP2/cc-pVDZ level.

It is interesting to compare the reaction pathway for guanine formation with the two possible reaction pathways shown in Figure 3.23. The guanine pathway shown in Figure 3.40 is the shortest path starting from 2-aminomalononitrile, which includes 9 steps, while in the case of Figure 3.23, 10 (black line) or 11 steps are required (hydrolysis of 2,6-diaminopurine (red and purple lines)). Moreover, the formation of guanine via 2-aminomalononitrile and guanidine, Figure 3.40, is thermodynamically preferred (-43.6 kcal mol⁻¹) than the formation of guanine by the route shown in Figure 3.23 ((-5.2 kcal mol⁻¹) (black line) and -11.7 kcal mol⁻¹ (red and purple lines)). The rate-determining reactions for guanine formation (Figure 3.23) are the hydrolysis of N'-(4-cyano-1*H*-imidazol-5-yl)urea with a barrier of 38.8 kcal mol⁻¹ (black line), the formation of 1-(4-cyano-1*H*-imidazol-5-yl)urea from AICN and urea with a barrier of 31.9 kcal mol⁻¹ (red and purple lines), and the formation of 5,6-diamino-2-imino-2,3-dihydropyrimidine-4-(1*H*)-one after cyclization of (*E*)-2,3-diamino-3-guanidinoacrylamide (Figure 3.40) with a barrier of 36.6 kcal mol⁻¹. A comparison of these facts gives a clear idea that the reaction pathway for guanine formation after hydrolysis of 2-aminomalononitrile and subsequent reaction with guanidine (Figure 3.40) is kinetically preferred.

3.5 INFLUENCE OF HIGHER ORDER CORRELATION ENERGY EFFECTS, ENTROPIC CONTRIBUTIONS AND LONG-RANGE INTERACTIONS IN SOLVENT ON THE ENERGETICS OF THE REACTION

Spin-component-scaled MP2 method has been successfully used for the investigation of the reaction mechanism of tautomeric conversions or chalcone formation. Galvao *et al* reported that the value of the energy gap between *cis*-2-hydroxypyridine and 2-pyridone, calculated at SCS-MP2/aug-cc-pVTZ level is closer to the experimental value than the one calculated at CCSD(T)/aug-cc-pVTZ level. In addition, the SCS-MP2 method has the advantage of lower computational cost than QCISD(T) and CCSD(T) methods. It is shown in the literature that the calculations at SCS-MP2 level increase the energy barriers compared to the MP2 ones.

Approaches based on the refinement of the wave function, such as Møller-Plesset (MP2) and coupled-cluster (CC) theory, have become the methods of choice when high accuracy is needed, with the CCSD(T) calculations (so called the “gold standard” for accurate computations) being capable of achieving chemical accuracy. CC techniques (CCSD(T) etc.) provide the most accurate possible structures and energies, but these are methods are currently only applicable to small molecules.

Additional single-point CCSD(T)/cc-pVDZ calculations were performed using the MP2/cc-pVDZ-optimized geometries of the reactants, products, and transition states in reactions 52-59. For comparison, the results obtained at the CCSD(T) and SCS-MP2 levels of theory are presented in Table 2. Energy barriers of the reactions calculated at the MP2 level are 5% to 7% higher (reactions 54-59) or similar (reaction 53) to those calculated at the CCSD(T) level. The exception is reaction 52—the SCS-MP2-calculated barrier is 6% lower than the CCSD(T)-calculated one. However, SCS-MP2 has a substantial advantage because of lower computational cost than CCSD(T).

Table 2.

Energy differences between reactants and products at 0 K (ΔH_0) and the respective energy barriers ($\Delta H_0^\#$) in kcal mol⁻¹ for the reactions 52-59, shown in Figure 3.28, calculated at CCSD(T)/cc-pVDZ and SCS-MP2/ccpVDZ levels of theory using MP2/ccpVDZ-optimized geometry.

Reaction	CCSD(T)/cc-pVDZ		SCS-MP2/cc-pVDZ	
	ΔH_0	$\Delta H_0^\#$	ΔH_0	$\Delta H_0^\#$
52	-14.0	39.0	-16.0	36.7
53	-1.7	43.0	-2.9	43.6
54	11.0	14.5	12.6	16.1
55	6.1	27.8	6.6	29.2
56	-10.0	42.5	-10.0	45.2
57	-7.4	36.4	-9.3	38.7
58	20.1	31.0	21.9	33.0
59	-5.9	21.4	-5.2	23.0

The SCS-MP2 method was used to obtain more accurate values of the reaction energies than the standard MP2. SCS-MP2/cc-pVDZ energy calculations based on optimized MP2/cc-pVDZ structures show that the average increase in activation energy of reactions is in the range of 6 to 12%.

The comparison of the SCS-MP2/cc-pVDZ and SMD/SCS-MP2/cc-pVDZ calculations allow observing solvent effect impact on the activation energy barriers of the reactions at SCS-MP2 level. The use of the SMD model leads to a reduction of the energy barrier of the reactions in the range 0.2-13.9 kcal mol⁻¹. Exceptions are some dehydration reactions (steps 6, 8, 16, 49, 55 and 86) and hydrolysis reactions (steps 22 and 63), where the energy barrier values are higher - from 0.7 to 3.4 kcal mol⁻¹.

Comparing the influence of the solvent at the highest theoretical level, ie. comparison of SCS-MP2/cc-pVDZ and SMD/SCS-MP2/cc-pVDZ, the decrease is strongest in the reactions of: deamination 24 (52%), 27 (33%) 41 (58%), 43 (54 %), 64 (45%), 70 (41%) and 74 (55%); tautomerization of isoguanine 47 (38%); reaction between guanidine and 3-hydroxyacrylonitrile 65 (27%); reaction between 2,5,6-triaminopyrimidin-4 (3*H*)-one and glyoxal 81 (41%); formation of formyl cyanide (38%), urea (57%) and 2 - (amino(hydroxy)methyl)amino)acetonitrile 1 (31%).

The potential energy profiles of the chemical reactions (steps 1-87) reproduce zero-kelvin chemistry. However, this limits the comparisons with the experiments which are carried out at finite, and sometimes at very high, temperatures. Quite recently, a critical comparison of zero-kelvin quantum chemical calculations and *ab initio* molecular dynamics and metadynamics simulations, fully including entropic contributions, has been reported with respect to some experimental evidences. For the chemical conversion of formaldehyde and hydrogen cyanide into glycolonitrile, Ferus et al have been shown that at 300 K glycolonitrile is more stable than HCHO and HCN, but at extreme temperature (4500 K) there is the stability inversion between the reactants and the product. To taking into account the entropy contribution and the temperature in the free energy, the values of Gibbs free energies (ΔG) and the activation barriers (ΔG^\ddagger) are calculated for a temperature of 298.15 K. A comparison between thermochemistry of the reactions at 0 K and 298.15 K is presented in Table 2.

The energy difference between reactants and products at 0 K (ΔH_0) is smaller than this one at 298.15 K (ΔG_{298}) for all reaction steps of ammonolysis/hydrolysis of the cyano group in *ICNFI* and the subsequent cyclization leading to adenine/hypoxanthine formation (steps 17-27) as well as formation of 2,6-diaminopurine, guanine and xanthine from *AICN* (steps 28-42). Similar results are obtained for the reactions of: tautomerization, dehydration (Fig. 2d) and decarbonylation of formamide, 2-hydroxyl-aminoacetonitrile and formylcyanide formation from hydrogen cyanide and formamide, formation of 2-aminoacetonitrile and isocyanic acid and of 2-amino-malononitrile from 2-iminoacetonitrile. In the reactions of cytosine and uracil formation, the energy differences (ΔH_0) are smaller than (ΔG_{298}) in steps 60-62 (except the hydrolysis and deamination reactions (steps 63,64)). A similar trend is observed in most of the reactions for the formation of pterin and guanine (steps 76, 78-82, 84, 87).

The opposite trend is observed in reactions 1-3 11,12, 52, 55, 66, 68, 71, 75, 77, 83, 85, 86 and others. This trend is also valid for the formation of hydrogen cyanide and water from formimidic acid, formic acid and ammonia from formamide and water, urea from isocyanic acid and ammonia. An interesting case is the formation of 2-(amino(hydroxy) methyl) amino) malononitrile from 2-amino-malononitrile and formamide (step 11). While the reactants were less stable than the product by 1.27 kcal mol⁻¹ at 0 K, they became more stable by 0.82 kcal mol⁻¹ at 298.15 K (Table 2). A similar case is reaction 59, where an endothermic reaction at 0 K (0.18 kcal mol⁻¹) becomes exothermic (-2.40 kcal mol⁻¹) at 298.15 K.

For all cases considered, the activation barriers at 298.15 K, ($\Delta G_{298}^\#$) are higher than the barriers at 0 K ($\Delta H_0^\#$) (see Table 3), except the reaction for the formation of 3-hydroxyacrylonitrile, where the barrier is lower by 1%. The main trend is the increase of activation energy from 1% to 18%. In addition, for some reactions, the increase is in the range of 19-29% (steps 24, 25, 54, 81). The change is most pronounced for the reaction of urea formation from isocyanic acid and ammonia - 46% and step 83 (six-membered ring closure before the formation of pterin) - 32%. The rate constants at 298.15 K calculated by the Eyring equation, $k = (k_B T / h) \cdot e^{-\Delta G^\# / RT}$, where k_B , h and R are the Boltzmann, Planck and universal gas constants, respectively, and $\Delta G^\#$ is Gibbs free energy of activation are also presented in Table 3.

Table 3.

Energy differences, ΔH_0 и ΔG_{298} , and energy barriers, $\Delta H_0^\#$ and $\Delta G_{298}^\#$ (kcal/mol), for the reactions shown on Figures 3.3-3.9, 3.13, 3.14, 3.17, 3.19, 3.22, 3.25, 3.28, 3.33, 3.36 and 3.39 in the dissertation, calculated at SMD/SCS-MP2/cc-pVDZ level. The corresponding rate constants (k) are calculated in s^{-1} .

Reaction	ΔH_0	ΔG_{298}	$\Delta H_0^\#$	$\Delta G_{298}^\#$	k
fig. 3.3a	12.91	13.00	25.36	27.01	9.83×10^{-8}
fig. 3.3b	33.80	34.16	34.99	36.75	7.12×10^{-15}
fig. 3.3c	-20.64	-20.16	21.35	21.72	7.43×10^{-4}
fig. 3.3d	-8.19	-9.91	32.53	34.24	4.93×10^{-13}
fig. 3.3e	-13.16	-12.66	25.08	27.73	2.92×10^{-8}
fig. 3.3f	-6.71	-7.47	33.65	34.55	2.92×10^{-13}
fig. 3.4 (top)	13.35	14.70	35.31	36.78	6.77×10^{-15}
fig. 3.4 (left)	16.15	15.49	31.12	31.89	2.60×10^{-11}
fig. 3.4 (right)	1.27	2.02	4.35	5.16	1.02×10^9
fig. 3.5a	-2.93	-4.41	40.33	41.94	1.12×10^{-18}
fig. 3.5b	-11.77	-10.10	6.23	9.13	1.26×10^6
fig. 3.5c	2.96	4.62	31.87	34.68	2.34×10^{-13}
fig. 3.6a	-8.43	-6.00	30.98	33.34	2.25×10^{-12}
fig. 3.6b	-13.90	-14.16	3.28	4.16	5.54×10^9
fig. 3.7a	3.75	4.08	8.88	10.13	2.33×10^5
fig. 3.7b	-9.98	-11.64	29.29	29.95	6.88×10^{-10}
fig. 3.7c	-14.54	-13.22	22.78	24.66	5.19×10^{-6}
fig. 3.8a	-13.93	-11.75	25.53	27.85	2.38×10^{-8}
fig. 3.8b	0.59	0.59	6.98	7.39	2.38×10^7
fig. 3.8c	-7.30	-7.76	21.35	23.39	4.43×10^{-5}
fig. 3.8d	-2.37	-4.34	49.80	51.84	6.17×10^{-26}
fig. 3.9a	-29.03	-26.22	39.74	42.44	4.80×10^{-19}
fig. 3.9b	5.79	6.31	37.26	39.12	1.30×10^{-16}
fig. 3.9c	4.42	2.25	63.75	63.30	2.45×10^{-34}
step 1	-2.87	-0.80	15.78	18.18	2.92×10^{-1}
step 2	6.05	5.11	21.88	22.44	2.20×10^{-4}
step 3	-12.83	-12.60	31.48	33.46	1.84×10^{-12}
step 4	8.31	8.70	22.81	24.67	5.11×10^{-6}
step 5	-4.03	-1.17	16.22	18.82	9.92×10^{-2}
step 6	7.16	5.94	25.04	26.15	4.20×10^{-7}

step 7	0.86	2.37	24.70	26.70	1.66×10^{-7}
step 8	12.42	10.94	29.21	30.25	4.15×10^{-10}
step 9	-0.94	-0.24	21.73	22.93	9.63×10^{-5}
step 10	-32.49	-35.41	21.01	22.24	3.09×10^{-4}
step 11	-1.27	0.82	20.96	24.36	8.62×10^{-6}
step 12	8.57	7.12	25.82	26.67	1.75×10^{-7}
step 13	-18.38	-19.13	26.20	27.70	3.07×10^{-8}
step 14	1.20	1.36	17.95	19.57	2.80×10^{-2}
step 15	-2.53	-0.70	20.38	22.38	2.44×10^{-4}
step 16	6.17	4.24	26.00	26.31	3.21×10^{-7}
step 17	12.15	13.50	34.61	36.67	8.15×10^{-15}
step 18	8.48	9.11	19.67	20.80	3.51×10^{-3}
step 19	-8.79	-11.44	12.54	13.22	1.26×10^3
step 20	6.72	8.63	43.78	46.55	4.66×10^{-22}
step 21	2.93	3.30	7.65	8.82	2.13×10^6
step 22	2.97	4.78	41.08	44.26	2.22×10^{-20}
step 23	12.62	13.31	22.84	23.78	2.29×10^{-5}
step 24	-19.25	-21.02	7.90	9.63	5.42×10^5
step 25	-5.22	-5.41	6.36	7.58	1.72×10^7
step 26	8.30	9.70	26.27	28.71	5.58×10^{-9}
step 27	-20.62	-23.16	5.39	5.93	2.79×10^8
step 28	-6.89	-7.35	31.93	35.00	1.37×10^{-13}
step 29	14.43	14.76	12.21	12.76	2.75×10^3
step 30	9.34	12.96	27.11	31.48	5.20×10^{-11}
step 31	14.49	16.29	27.23	29.03	3.25×10^{-9}
step 32	-16.83	-18.11	13.51	13.85	4.37×10^2
step 33	15.25	15.89	38.76	40.80	7.65×10^{-18}
step 34	17.67	18.81	29.62	30.62	2.22×10^{-10}
step 35	-31.05	-31.60	9.66	10.57	1.11×10^5
step 36	-4.74	-5.48	7.36	7.90	1.00×10^7
step 37	-8.40	-10.54	12.76	14.22	2.34×10^2
step 38	6.58	7.05	8.55	9.66	5.15×10^5
step 39	6.08	6.43	7.04	7.90	1.00×10^7
step 40	10.29	11.19	26.92	28.70	5.67×10^{-9}
step 41	-26.60	-28.84	3.32	3.79	1.03×10^{10}
step 42	3.55	5.09	25.67	27.92	2.12×10^{-8}
step 43	-30.81	-33.70	2.91	3.41	1.97×10^{10}
step 44	16.43	17.19	37.83	39.81	4.07×10^{-17}
step 45	18.89	17.83	31.75	32.76	5.99×10^{-12}
step 46	-21.55	-20.44	28.74	30.57	2.42×10^{-10}
step 47	4.98	5.19	6.84	7.89	1.02×10^7
step 48	19.15	22.05	24.30	27.53	4.09×10^{-8}
step 49	8.40	7.22	22.62	23.65	2.86×10^{-5}
step 50	-41.98	-43.55	5.89	6.71	7.49×10^7
step 51	-41.40	-42.51	6.48	7.63	1.58×10^7

step 52	-16.81	-16.65	27.73	30.45	2.96×10^{-10}
step 53	-5.52	-5.62	33.71	36.43	1.22×10^{-14}
step 54	10.63	13.20	13.88	17.87	4.93×10^{-1}
step 55	8.48	7.20	30.18	31.17	8.77×10^{-11}
step 56	-12.68	-11.86	44.01	45.57	2.44×10^{-21}
step 57	-7.43	-8.74	34.93	35.77	3.72×10^{-14}
step 58	21.49	23.64	31.03	33.74	1.15×10^{-12}
step 59	0.18	-2.40	25.58	26.17	4.06×10^{-7}
step 60	-17.72	-17.23	39.75	41.91	1.17×10^{-18}
step 61	-26.57	-25.87	24.19	26.60	1.96×10^{-7}
step 62	-21.43	-21.38	7.36	8.52	3.53×10^6
step 63	3.08	3.94	22.69	24.25	1.04×10^{-5}
step 64	-19.48	-22.14	4.25	4.36	3.95×10^9
step 65	-10.89	-12.74	38.50	40.24	1.97×10^{-17}
step 66	-21.97	-21.30	25.41	27.28	6.23×10^{-8}
step 67	5.21	5.30	7.82	8.66	2.78×10^6
step 68	0.22	0.21	5.90	6.50	1.07×10^8
step 69	9.79	11.38	31.24	33.49	1.75×10^{-12}
step 70	-23.51	-25.72	4.46	4.94	1.49×10^9
step 71	-1.32	-0.31	20.36	22.35	2.56×10^{-4}
step 72	-22.87	-25.32	4.36	4.94	1.49×10^9
step 73	12.52	14.10	26.09	27.83	2.46×10^{-8}
step 74	-28.07	-30.29	3.81	4.45	3.40×10^9
step 75	-13.53	-12.51	25.39	26.90	1.18×10^{-7}
step 76	1.21	1.38	6.93	7.43	2.22×10^7
step 77	-6.05	-4.62	33.42	34.94	1.51×10^{-13}
step 78	-6.98	-7.86	20.87	21.68	7.94×10^{-4}
step 79	13.35	14.42	36.56	38.74	2.48×10^{-16}
step 80	3.89	4.21	8.21	9.02	1.52×10^6
step 81	2.08	3.79	10.48	12.83	2.44×10^3
step 82	-13.53	-14.21	15.34	16.53	4.74
step 83	-7.49	-5.03	8.56	11.26	3.46×10^4
step 84	-19.73	-21.54	11.70	12.22	6.84×10^3
step 85	1.10	2.61	19.20	20.82	3.39×10^{-3}
step 86	1.50	0.71	20.36	20.90	2.96×10^{-3}
step 87	-38.14	-39.66	8.42	9.72	4.65×10^5

3.6 SUMMARY

A unified reaction pathway for the formation of purine and pyrimidine nucleobases and pterin from pure formamide has been proposed. All reactants (HCN, ammonia, water, HCOOH, HNCO, urea, guanidine, glyoxal and 3-hydroxyacrylonitrile) required for subsequent reactions are obtained by decomposition of formamide. The reactions presented are catalyzed only by formamide or its tautomer, formimidic acid, which is formed in the first step of the reaction pathway (Figure 3.41). These “one-pot” formamide-based reactions occur due to thermal processes.

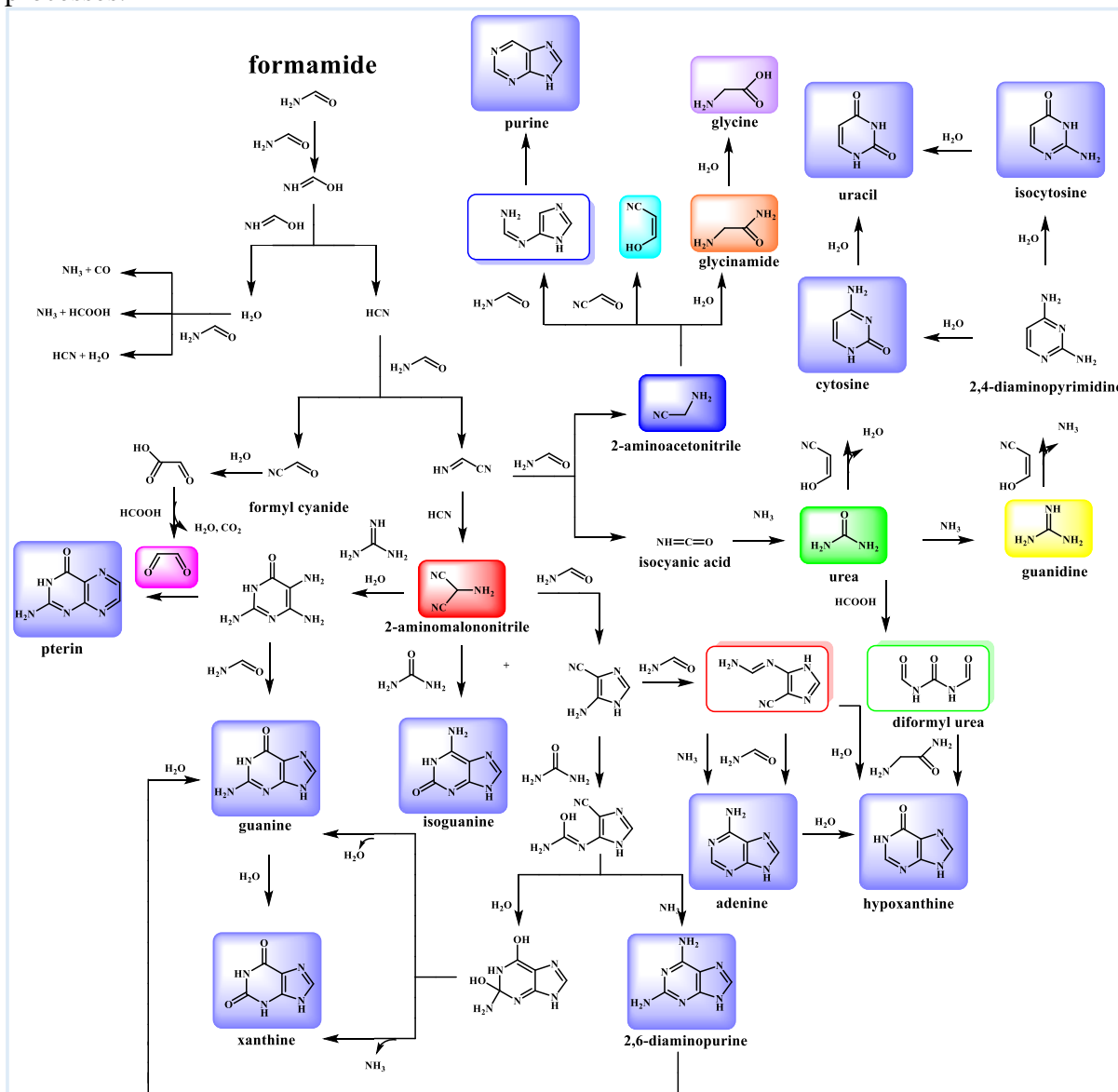


Figure 3.41 General reaction scheme showing a complete picture of the formation of prebiotic precursors, purine and pyrimidine bases and pterin from formamide.

In the proposed reaction scheme (Figure 3.41) purine, adenine, hypoxanthine, 2,6-diaminopurine, guanine and xanthine are formed by imidazole-first mechanism, while in case of isoguanine – by pyrimidine-first mechanism. Guanine can also be obtained by a mechanism in which the pyrimidine ring is formed first and after that the imidazole ring (Figure 3.40).

The formation of guanine from 2-aminomalononitrile and guanidine is thermodynamically preferred. In addition, urea plays an essential role in the formation of 2,6-diaminopurine, guanine, xanthine, isoguanine, cytosine and uracil. Guanidine, in turn, is a

starting compound for the formation of isocytosine, cytosine and uracil from 2,4-diaminopyrimidine. Glyoxal, which is a product of the hydrolysis of formyl cyanide, is a prebiotic precursor leading to the formation of pterin.

The obtained results show that there is a close connection between the prebiotic chemistry of guanine and pterin - coevolution of purines and pterin(s). Pterin (2-amino-4-oxopteridine) is the core of the major cofactors. The biochemical advantage of pterines as cofactors is their ability to form reduced (dihydro and tetrahydro) derivatives, which allows them to be the main agents in several biological redox reactions for DNA synthesis and maintenance. It is possible that in the early stages of chemical evolution, both structures were available to be incorporated into pre-RNA structures. In terms of prebiotic chemistry, purines and pterines seem destined to be together.

Following the reaction pathways shown in Figure 3.15 and Figure 3.20, purine should be formed first, as the determining step is six-membered ring closure (for *IFI* and *ICNFI*, see Figure 3.12 and Figure 3.16). Adenine is more likely to appear first (following the black line in Figure 3.20), but if the heating of formamide is for a longer period of time, adenine and hypoxanthine will be present in the mixture at the same time. From a thermodynamic point of view, it appears that guanine should also be present in the mixture (following the black line in Figure 3.23). 2,6-diamin-purine is less likely to be detected in the reaction mixture, as it should be converted to guanine and possibly from guanine to xanthine.

Probably the same problem is observed with isoguanine and xanthine. As it can be seen from Figure 3.10, urea should be present in the mixture at an early stage of heating of formamide. The amount of isoguanine depends on the amount of urea. Although all previous steps in the reaction pathway for isoguanine formation (Figure 3.26) are endothermic, the last step is highly exothermic and probably isoguanine can be detected. Liquid chromatography combined with tandem mass spectrometry (LCMS/MS) does not require sample derivatization and suitable for analysis of these compounds.

To sum up, the presented theoretical studies provide a basis for more detailed and in-depth analyzes of mixtures obtained by heating pure formamide.

CONCLUSIONS

1. Formamide heating leads to the formation of main prebiotic precursors like hydrogen cyanide, water, formic acid, ammonia, isocyanic acid, 2-aminoacetonitrile, 2-aminomalononitrile, urea, guanidine, glyoxal and formylcyanide. The energy barriers of these reactions do not exceed $40.0 \text{ kcal mol}^{-1}$ that means that all above presented reactions in formamide should occur at temperatures from $100 \text{ }^\circ\text{C}$.
2. A reaction route of self-catalytic chemical reactions for the formation of nucleobases has been proposed. Mechanisms of prebiotic self-catalytic reactions for the formation of **adenine, hypoxanthine, 2,6-diaminopurine, guanine, xanthine, isoguanine, pterin, 3-hydroxyacrylonitrile and urea from formamide have been proposed for the first time without the presence of catalysts.**
3. It has been shown that employment of higher order correlation energy effects using the SCS-MP2 method leads to an increase in the calculated activation energy of the studied reactions from 6 to 12% compared to the calculations carried out with the MP2 method.
4. Taking into account the influence of the solvent (formamide) can reduce the energy barrier of reactions up to 55%.
5. The obtained results about the reaction pathways, intermediates, energy barriers and transition states of the simulated reactions make it possible to calculate the kinetic and energy characteristics of the reactions. It has been shown that the consideration of entropy factors leads to an increase in energy barriers by an average of 9% compared to the barriers calculated at 0 K.
6. The results of the calculations performed with the SCS-MP2 method have been shown to be comparable to those performed with CCSD(T), which are so-called "gold standard" in terms of accuracy.
7. The present research shows how starting from one parent molecule (formamide) in a very logical way one can reach the building blocks of the information polymers, DNA and RNA, which are at the core of chemical evolution. The obtained results are a sound basis for further research in the field of chemical evolution based on the prebiotic reactions of formamide.

CONTRIBUTIONS

1. A unified reaction scheme has been developed for the formation of purine and pyrimidine nucleobases only from pure formamide.
2. It has been shown for the first time that pterin can be obtained by thermal heating of formamide and appropriate reaction pathways have been proposed.

Publications included in the PhD thesis

- (1) Slavova, S.; Enchev, V. **Self-catalytic Mechanism of Prebiotic Reactions: From Formamide to Purine Bases.** *Int. J. Quantum Chem.* 120 (19), 2020, e26362. <https://doi.org/10.1002/qua.26362>.
IF=2.26, Q2
- (2) Enchev, V.; Angelov, I.; Dincheva, I.; Stoyanova, N.; Slavova, S.; Rangelov, M.; Markova, N. **Chemical Evolution: From Formamide to Nucleobases and Amino Acids without the Presence of Catalyst.** *J. Biomol. Struct. Dyn.* 39, 2021, 1–16. <https://doi.org/10.1080/07391102.2020.1792986>.
IF = 3.31, Q2
- (3) Enchev, V.; Slavova, S. **Self-catalytic Mechanism of Prebiotic Reactions: II. From Urea and Glycinamide to Hypoxanthine.** *Int. J. Quantum Chem.* 121, 2021, e26508. <https://doi.org/10.1002/qua.26508>.
IF = 2.26, Q2

List of citations

- (1) Enchev, V.; Slavova, S. **Self-catalytic Mechanism of Prebiotic Reactions: II. From Urea and Glycinamide to Hypoxanthine.** *Int. J. Quantum Chem.* 121, 2021, e26508. <https://doi.org/10.1002/qua.26508>.
IF = 2.26, Q2

Cited:

Dutta, A.; Trolles-Cavalcante, S. T. Y.; Cleetus, A.; Marks, V.; Schechter, A.; Webster, R. D.; Borenstein, A. **Surface Modifications of Carbon Nanodots Reveal the Chemical Source of Their Bright Fluorescence.** *Nanoscale Adv.* 2021. <https://doi.org/10.1039/D0NA00871K>.

- (2) Enchev, V.; Angelov, I.; Dincheva, I.; Stoyanova, N.; Slavova, S.; Rangelov, M.; Markova, N. **Chemical Evolution: From Formamide to Nucleobases and Amino Acids without the Presence of Catalyst.** *J. Biomol. Struct. Dyn.* 39, 2021, 1–16. <https://doi.org/10.1080/07391102.2020.1792986>.
IF = 3.31, Q2

Cited:

Hortal, L.; Pérez-Fernández, C.; de la Fuente, J. L.; Valles, P.; Mateo-Martí, E.; Ruiz-Bermejo, M. **A Dual Perspective on the Microwave-Assisted Synthesis of HCN Polymers towards the Chemical Evolution and Design of Functional Materials.** *Sci. Rep.* 2020, 10 (1), 22350. <https://doi.org/10.1038/s41598-020-79112-5>.

Conferences:

1. 18/11/2020
Sofia,
Bulgaria

Anniversary scientific session dedicated to the 100th anniversary of the birth of Academician Georgi Bliznakov
S. Slavova, V. Enchev
Poster "Ab initio modeling of mechanisms of prebiotic reactions: from formamide to hypoxanthine"
2. 10/09/2020
Sofia,
Bulgaria

Eleventh scientific session "YOUNG SCIENTISTS IN THE WORLD OF POLYMERS"
S. Slavova, V. Enchev
Poster "Prebiotic formation of hypoxanthine from formamide"
3. 06-08/12/2019
Creative House of
BAS "Golden
Bridges" -
Vitosha

XXII Winter Seminar "INTERDISCIPLINARY PHYSICS" of young scientists and PhD students
S. Slavova, V. Enchev
Oral presentation "Formation of pterin from formamide"
4. 22-23/11/2019
Plovdiv,
Bulgaria

Fifth Scientific Conference for Students, PhD Students and Young Scientists "CHALLENGES IN CHEMISTRY"
S. Slavova, V. Enchev
Poster "Formation of pyrimidine bases cytosine and uracil from formamide"
5. 01-02/11/2019
Razgrad,
Bulgaria

Scientific conference with international participation "New industries, digital economy, society-projections of the future-II"
Enchev V., Slavova S.
Oral presentation "Ab Initio Study of Mechanism of Prebiotic Reactions: from Urea and Glycinamide to Hypoxanthine"
6. 29/08-31/08/2019
Samokov Hotel,
Borovets

Second interdisciplinary PhD students forum
S. Slavova, V. Enchev
Oral presentation "From formamide to nucleobases – puzzle arrangement"
7. 03-05/07/2019
Sofia,
Bulgaria


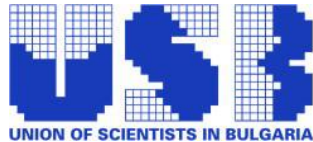
Tenth Workshop "Experimental Models and Methods in Biomedical Research"
Slavova S., Enchev V.
Oral presentation "Ab initio study of the nucleobases formation as a result of chemical evolution"
8. 15-17/05/2019
Sofia,
Bulgaria

XVIII National Conference on Chemistry for students and PhD students 2019
Slavova S., Enchev V.
Oral presentation "Chemical evolution - formation of purine bases from formamide without the presence of a catalyst"
9. 19-21/04/2019
Sofia,
Bulgaria

XII Spring Seminar of Young Scientists and Doctoral Students from BAS "Interdisciplinary Chemistry"
Slavova S., Enchev V.
Oral presentation "Quantum chemical modeling of prebiotic reactions based on formamide"

- | | | |
|-----|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10. | 04-05/04/2019
Sofia,
Bulgaria | <p><i>National Student Conference on pharmaceutical and chemical sciences</i></p> <p>Slavova S., Enchev V.
Oral presentation "Ab initio modeling of prebiotic formamide-based reactions"</p> |
| 11. | 02-03/11/2018
Razgrad,
Bulgaria | <p><i>Scientific conference with international participation "New industries, digital economy, society-projections of the future"</i></p> <p>Enchev V., Slavova S.
Poster "From formamide to glycine acid urea: an ab initio study"</p> |
| 12. | 02-03/11/2018
Razgrad,
Bulgaria | <p><i>Scientific conference with international participation "New industries, digital economy, society-projections of the future"</i></p> <p>Enchev V., Angelov I., Markova N., Stoyanova N., Slavova S., Dincheva I., Krasavin E., Kapralov M., Avramov L.
Oral presentation "From formamide to nucleic acid monomers and amino acid by using various energy sources"</p> |

Awards

- | | | | |
|----|----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| 1. | 12/05/2019
Sofia,
Bulgaria | <p>II award with the topic "The origin of life"
FameLab: science communication competition
https://nauka.offnews.bg/news/Novini_1/13-ta-laboratoriia-za-slava-FameLab-Dobrite-poslaniiia-na-mladite-uche_128505.html
https://mediabricks.bg/финалът-на-famelab-2019-състезание-за-комуникат/</p> |  |
| 2. | | <p>Special award of the Union of Scientists in Bulgaria within science communication competition FameLab
https://mediabricks.bg/финалът-на-famelab-2019-състезание-за-комуникат/</p> |  |