UNIVERSITY OF CALIFORNIA

Santa Barbara

# Molecular Building Blocks of Life and their Photochemistry

A Thesis submitted in partial satisfaction of the

requirements for the degree Master of Science

in Chemistry

by

Si Young Lee

Committee in charge:

Professor Mattanjah S. de Vries, Chair

Professor Steven K. Buratto

Professor Luc Jaeger

September 2022

The thesis of Si Young Lee is approved.

_____

Luc Jaeger

_____

Steven K. Buratto

_____

Mattanjah S. de Vries, Committee Chair

September 2022

Molecular Building Blocks of Life and their Photochemistry

ACKNOWLEDGEMENTS

Forever grateful for my wife and her support. Thank you, Naughty, for being the best puppy ever.

ABSTRACT

Molecular Building Blocks of Life and their Photochemistry

by

Si Young Lee

The photochemistry of nucleobases upon UV exposure has been studied explicitly. Adenine, Guanine, Cytosine, and Thymine are the canonical nucleobases chosen as molecular building blocks by nature presumably under chemical selection pressure that preceded biology. Canonical nucleobases tend to have short excited-state lifetimes by removing the excited state population back to the electronic ground state via internal conversion (IC) in picoseconds or less. This fast and safe decay pathway is one of the intrinsic properties of the canonical bases and protects them against photodamage. This property is preserved since prebiotic times which makes these bases molecular fossils in modern days. The canonical bases and their derivates have been studied in the spectroscopy community, including the de Vries Lab. While canonical nucleobases are usually UV protected by having short excited-state lifetimes, a lot of derivatives and alternative compounds tend to have longer lifetimes which makes them susceptible to possibly destructive photochemical processes. Gas phase studies in the de Vries lab examine intrinsic properties of these molecules devoid of the molecular environment with isomer selectivity and supported by high level theoretical computations. Laser Desorption Jet-Cooling Mass

Spectroscopy is used. Laser desorption prevents fragmentation or thermal degradation of molecules, and the desorbed neutral molecules are entrained in a jet cooled molecular beam through supersonic expansion. This supersonic expansion lowers the internal temperature resulting in vibrational and rotational cooling. Jet cooled molecules are excited and ionized downstream, and the ions are detected with a time-of-flight mass spectrometer. Resonance enhanced multi-photon ionization (REMPI) is used. Double resonance or hole burning spectroscopy employs a burn laser and a probe laser. The burn laser, either UV or IR, is fired prior to the probe laser and is scanned while the probe laser is parked at a single resonant wavelength. Whenever the burn laser is resonant with a transition, it depletes the ground state population, and a decrease in ion signal is observed. When the burn laser is in UV range, this hole burning technique provides REMPI spectra of individual tuatomers present in the experimental set up. When the burn laser is in IR range, IR spectra of selected tautomers can be obtained and used for tautomer identification. Once the tautomers are identified, peaks from the REMPI spectrum can be selected for tautomer specific pump-probe spectroscopy that provides excited lifetime data. Nir et al. (1999) presented the first vibronic spectroscopy of guanine in the gas phase, and they were able to assign a 0-0 origin transition at 32, 878 $cm^{-1}$[1] The excited lifetime appeared to be exceeding 10 $\mu s$. Nir et al. (2001) presented an analysis based on double resonance spectroscopy, showing that the R2PI spectra included overlapping spectra of three different tautomers in Guanine.[2] IR-UV and UV-UV double resonance spectroscopy led to an initial assignment of three tautomers, as 9H-enol, and 9H- and 7H- keto tautomers respectively.[2] Brister et al. (2016) investigated promising nucleobase alternatives in solution, barbituric acid and 2,4,6-triaminopyrimidine.[3] They found both of the compounds to possess efficient electronic relaxation mechanisms for

eliminating excess amounts of absorbed energy into their aqueous molecular environment as heat within hundreds of femtoseconds.[3] On the other hand, Gengeliczki et al (2010) studied the excited state dynamics of pyrimidine derivatives such as 2,4 diaminopyrimidine and 2,6-diaminopurine, and they were able to observe the most stable tautomers of each compound respectively.[4] 2,4 diaminopyrimidine had only the diamino tautomer with its excited state lifetime bracketed between experimental limits of 10 ps and 1ns.[4] 2,6- aminopurine presented in two tautomeric forms, 9H- and 7H diamino forms with lifetimes of $6.36.3 \pm 0.4\ ns\ and\ 8.7 \pm 0.8\ ns$ respectively.[4] Lastly a pair of alternative nucleobases, isocytosine and isoguanine were studied.[5,6] They illustrate possible photochemical reasons why cytosine and guanine emerged as the building blocks of life rather than the alternatives, isocytosine and guanine.[5,6] Isocytosine presented similar dynamics to guanine while isoguanine presented similarities to cytosine.[5,6] These findings suggested that the lower photostability of the biologically relevant keto form of isoguanine can be a reason why isocytosine and isoguanine were not chosen by the nature under prebiotic conditions.[5,6]

TABLE OF CONTENTS

LIST OF FIGURES

# I. INTRODUCTION

Light has been the omnipresent source of life. Photochemistry studies the chemical behavior of electronically excited molecules through the absorption of light. In Dr. de Vries lab, we have been studying photochemistry and photo-dynamics of heterocyclic chromophores in the gas phase. Said heterocyclic chromophores include canonical nucleobases, nucleobase analogs, pigments, etc. A bottom-up approach or reductionist approach has been used where understanding was achieved on the molecular level to begin with. Complexity can be added step by step to eventually mimic the realistic environment starting from individual nucleobases followed by nucleosides where the sugar moiety is added, nucleotides where the phosphate moiety is added, base pairs, and then micro hydration where water molecules are added.

## *A. Primordial Soup on an Early Earth*

When was the origin of life on earth? In order to get a grasp about the origin of life on earth, one must go further back in time and look at how old our mother earth is. The age of earth is estimated to be 4.5 billion years old based on radiometric dating on one of the oldest meteorites.[7] The Great Oxidation Event (GOE) induced the enrichment of oxygen in the atmosphere which provided a rich oxygen source to create the ozone layer over a period of time. GOE occurred approximately 2.4 and 2.1 billion years ago.[8] Prior to the formation of the ozone layer, the surface of the earth had received an excess amount of Ultraviolet (UV) radiation. Oparin-Haldane hypothesis suggests that organic molecules were synthesized from inorganic molecules under the primordial conditions, and some of those organic molecules became building blocks of life.[9] The primordial atmospheric condition was a

reducing environment devoid of oxygen which is in accordance with the pre-GOE atmosphere on earth. This model suggests that the atmosphere consisted of methane, ammonia, carbon dioxide, water vapor, free hydrogen, and carbon.[9] Under this reducing atmospheric environment with energy input such as abundant UV radiation, the inorganic molecules could have undergone chemical reactions to form organic moleucles.[9] These organic molecules would have dissolved and accumulated in water such as the primordial ocean, and this would have created 'soup' which is now called the primordial soup.[9] More complex organic compounds would have developed over time with continued excess energy input such as lightnings and UV rays. In the early 1950s, Stanley Miller and Harold Urey performed an experiment based on this theory.[10] They set up an apparatus where methane, ammonia, water, and hydrogen gas were circulated past an electric discharge.[10] The mixture was tested for the synthesis of organic compounds such as amino acids by paper chromatography.[10]

This classic experiment demonstrated how organic compounds such as heterocyclic compounds could have become available in a prebiotic soup. What about the 'fitness' of those heterocyclic compounds?

### *B. Building Blocks of DNA*

Before the discussion of the fitness of the nucleobases, let us take a closer look at the building blocks of DNA. Adenine (A), Guanine (G), Cytosine (C), and Thymine (T) are the four letters that make up Deoxyribonucleic acid (DNA). These are the canonical nucleobases. A, G, C, T are heterocyclic nucleobases as shown in Figure 1 below. Uracil (U) is the RNA equivalent of T.

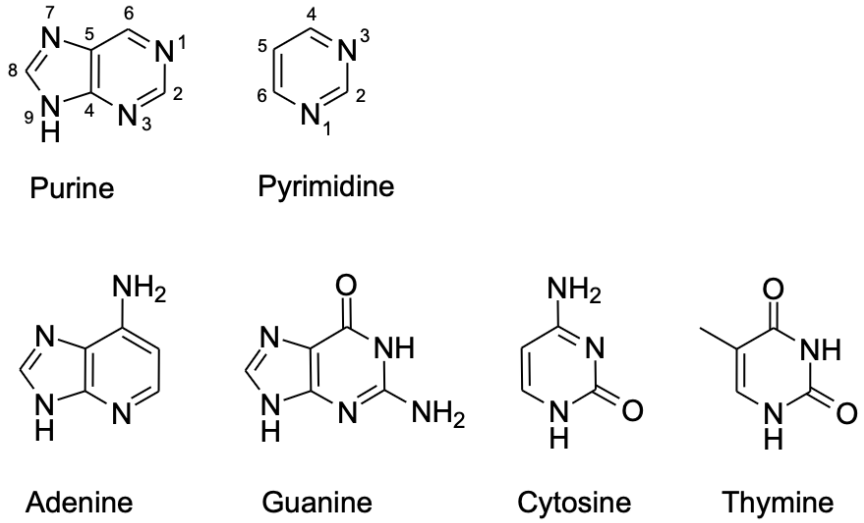


**Figure 1 The heterocyclic nucleobases. Purine and Pyrimidine are shown with conventional numbering system**

The nucleobases are the chromophore parts of DNA that absorb light in UV range. A and G are purines, and C and T are pyrimidines. Nucleosides are composed of a nucleobase and a sugar moiety, D-2-deoxyribose. Nucleic acids are biopolymers that are connected through phosphodiester linkages in between monomer units, and the monomer units are nucleotides. Most likely, various selection pressures would have molded the fittest composition of the canonical nucleobases. One of the driving factors may have been UV irradiation. The building blocks of DNA/RNA are extraordinarily photostable against UV irradiation. Photostability of the building blocks comes from rapid dissipation of excess energy absorbed through UV irradiation to the surrounding environment in the form of heat. This process is called Internal Conversion and it is indicated in Figure 3. This exquisite photostability would have given these specific nucleobases advantages over other DNA candidates especially when the amount of UV radiation was not attenuated yet due to the GOE. The rapid energy dissipation through a process called internal conversion results in short excited-state lifetimes and ultimately leads to less possibility of going through

photochemical reactions. This is an intrinsic property of the DNA building blocks that provides them protection from harmful UV exposure, which can lead to dire consequences such as apoptosis, mutagenesis, and cancer.[11] For example, 2-aminopurine and adenine (6-aminopurine both present favorable thermodynamics for base paring with uracil; however, their excited state lifetimes are drastically different from each other, 2-aminopurine having tens of nanosecond and adenine having 1 ps.[12,13] On the other hand, alternative nucleobases, derivatives, and analogs that share significant structural similarity lack this intrinsic property even when the structural difference seems miniscule. The canonical nucleobases have sub picosecond excited state lifetimes while the alternatives tend to have lifetimes that are orders of magnitude longer. The lifetimes of canonical and alternative nucleobases are visualized in Figure 2. How these heterocyclic chromophores respond upon UV exposure depends on their molecular structure. The structural dependency comes from the fact that rapid Internal Conversion occurs *via* Conical Intersections (CI). CI is defined as the locus where two potential curves cross or intersect when plotted against nuclear coordinates.[14] CIs can exists at molecular geometries that are different from the ground state equilibrium geometry which accompanies a molecular frame twist. For purines, pyramidalization at the C2 atom can govern rapid internal conversion, and for pyrimidines deformation of the C5=C6 bond can govern internal conversion. Any subtle structural differences involving these motifs can result in prolonged excited state lifetimes.

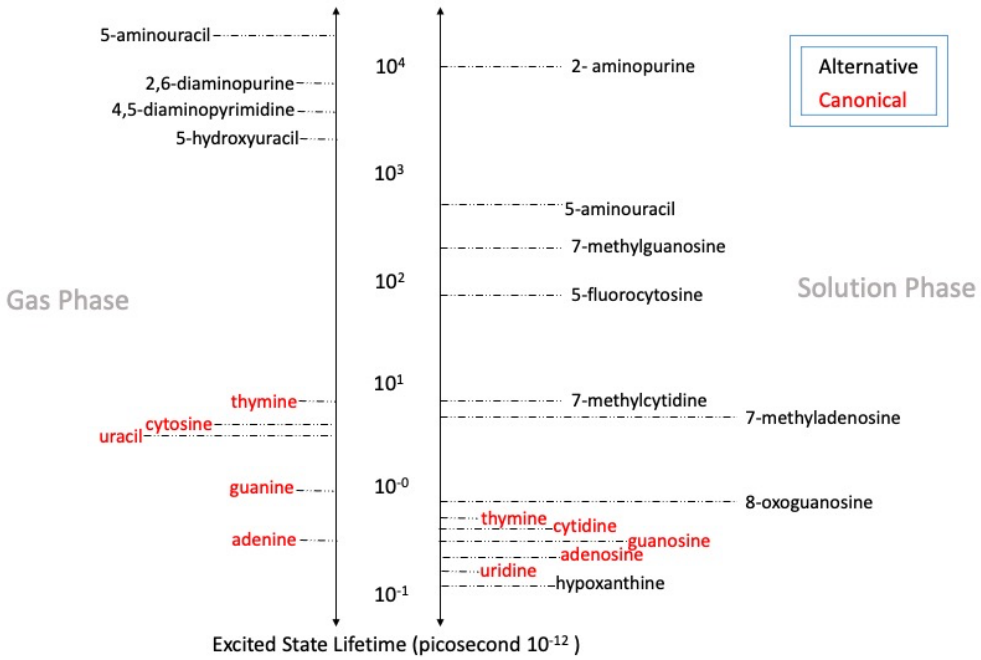


**Figure 2 Lifetimes of canonical nucleobases and alternatives/derivatives/analogs either in gas phase or solution phase.**

Prolonged excited state lifetimes can open up the pathways of possibly harmful photochemical processes. For a better understanding of photochemistry of polyatomic compounds, a simplified Jablonski diagram is shown in Figure 3.

### *C. Jablonski Diagram*

A Jablonski diagram provides a 2-dimensional visual representation of the photoexcitation and following processes of polyatomic compounds. Although the presented Jablonski diagram here is over simplified, it visualizes the possible transition pathways.

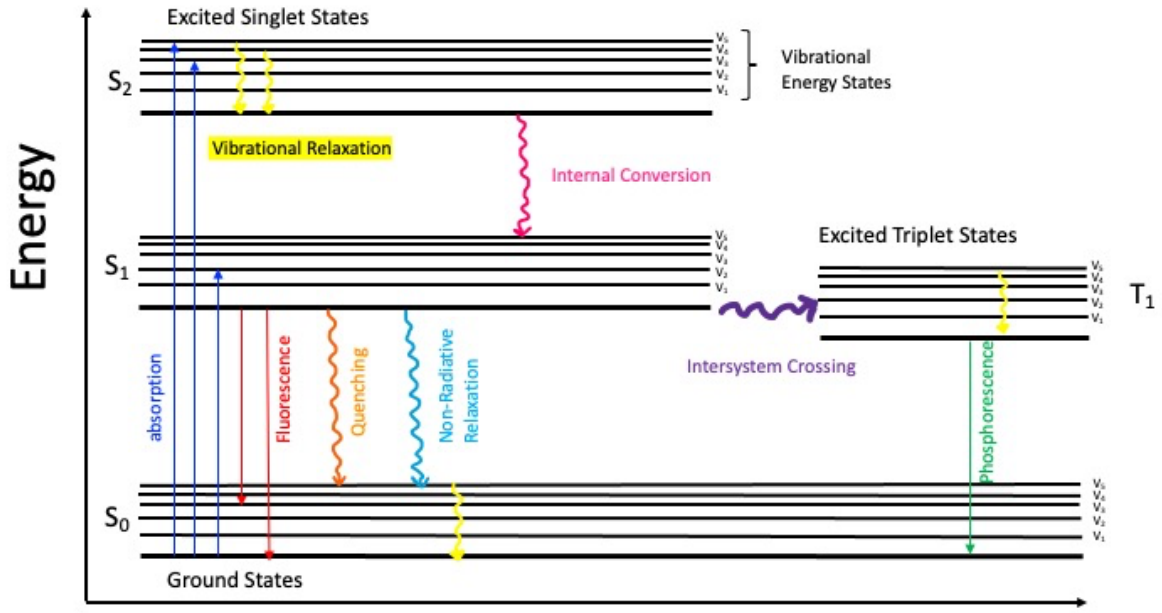


**Figure 3 Simple Jablonski Diagram showing the outcomes of excited states of polyatomic molecules. Horizontal black lines refer to the energy levels. The bold ones indicate the electronic states with the lowest vibrational level. Higher vibrational levels are indicated by thin black lines increasing from bottom to top order. Note only the first five vibrational energy states are included in this diagram. S stands for Singlet states and T stands for triplet states. $S_n$ denotes the $n^{th}$ excited singlet state and $T_n$ denotes $n^{th}$ excited triplet state. Straight arrows indicate radiative processes and wavy arrows indicate non-radiative processes.**

The ground excited state of polyatomic compounds is usually in a singlet state. A singlet state has a total electron spin angular momentum of zero where the highest occupied molecular orbital contains two opposing electron spins in a pair. An electronically excited singlet state is induced upon the absorption of light of the ground state electrons without a spin state change. An electronically excited singlet state can decay through fluorescence which is a spontaneous emission process, or it can go through a spin inversion process called intersystem crossing. Once it goes through the intersystem crossing, it will not be a singlet state and it will be the electronically excited triplet state. The electronically excited triplet state can decay through phosphorescence.

## II. METHODS

We use Laser-Desorption Jet-Cooling Mass Spectrometry (LD JC MS) to study isolated target compounds in the gas phase. Figure 4 shows a schematic diagram of the set-up.[15] The sample initially adsorbed on a surface of a translating graphite bar is desorbed by ND:YAG laser pulses. The desorbed molecules are entrained in a supersonic jet-cooled argon molecular beam. Jet-cooling lowers the internal energy of the laser desorbed molecules by multiple collisions. Down the jet stream, the jet cooled gaseous molecules are ionized by one or multiple ionization laser(s) in the collision-free region, and then the ions are directed upward into a reflectron time-of-flight mass spectrometer.

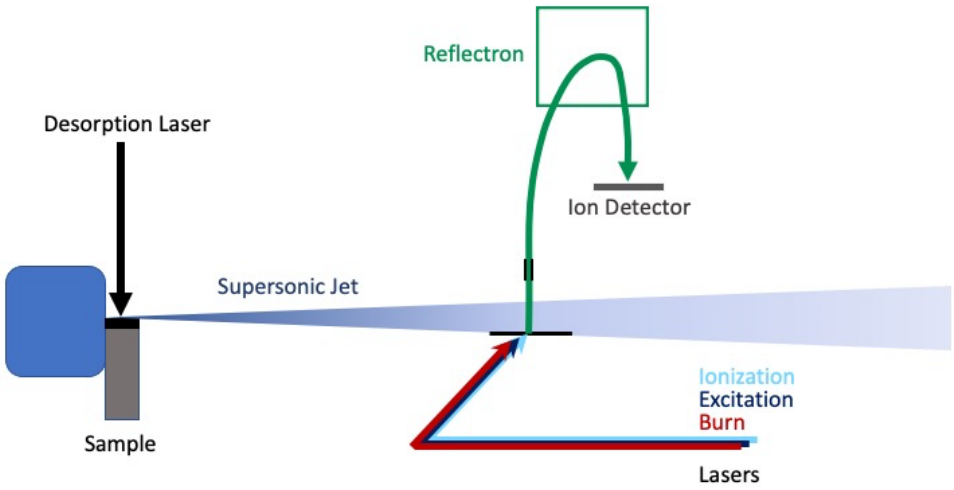


**Figure 4 The experimental Setup for LD JC MS**

### *A. Laser Desorption and Jet Cooling*

Spectroscopy is one of the most powerful tools to study the structure of molecules the and dynamics of chemical processes.[16] At room temperature, a complete absorption spectrum is composed of a massive amount of individual lines that cannot be resolved into their individual components.[16] The complexity of spectra for gas phase molecules comes

from the rotational fine structure, and the population of excited vibrational states.[16] Supersonic jet spectroscopy allows cooling of the internal degrees of freedom, specifically rotations and vibrations of the molecule.[16] There are three advantages of jet cooling; it improves spectroscopic resolution, reduces fragmentation following photoionization, and provides a pathway for cluster formation.[17] Laser desorption can induced by laser excitation of the surface itself.[18] Laser desorption minimizes fragmentation by shortening the timescale for heating.[17] At a heating rate of $10^{11}$ K/sec, this rapid laser pulse induced heating can initiate thermal desorption of surface molecules within several nanoseconds without fragmentation.[18] Desorbing the sample molecules on the lower pressure side of the pulsed valve requires less material which makes the technique of combination of laser desorption and jet-cooling sensitive.[17]

In many applications a focused Nd:YAG (1064 nm, ~1mJ pulse, at 10Hz) has been used as desorption laser. Supersonic jet cooling involves entrainment in an argon molecular beam, and the ions are detected in a reflectron time-of-flight mass spectrometer.

### *B. REMPI Ionization Technique and its Variation*

Laser desorption and jet-cooling produce neutral molecules in the gas phase. These gaseous neutral molecules undergo ionization downstream of the molecular beam. Resonance Enhanced Multi-Photon Ionization (REMPI) is an ionization technique that is used in de Vries lab. Figure 5 shows various ionization techniques used for optical spectroscopy. Figure 5(a) describes how REMPI works and produces UV excitation spectra. One photon excites the molecule to an electronic excited state from the electronic ground state when it is tuned to a resonance as shown as thicker arrow in figure 5(a). The second photon, hence the reason why it is called multi-photon ionization, then ionizes the molecule

by surpassing the ionization potential as shown as thinner arrow in figure 5(a). This second photon could be produced from the same laser or a second laser.

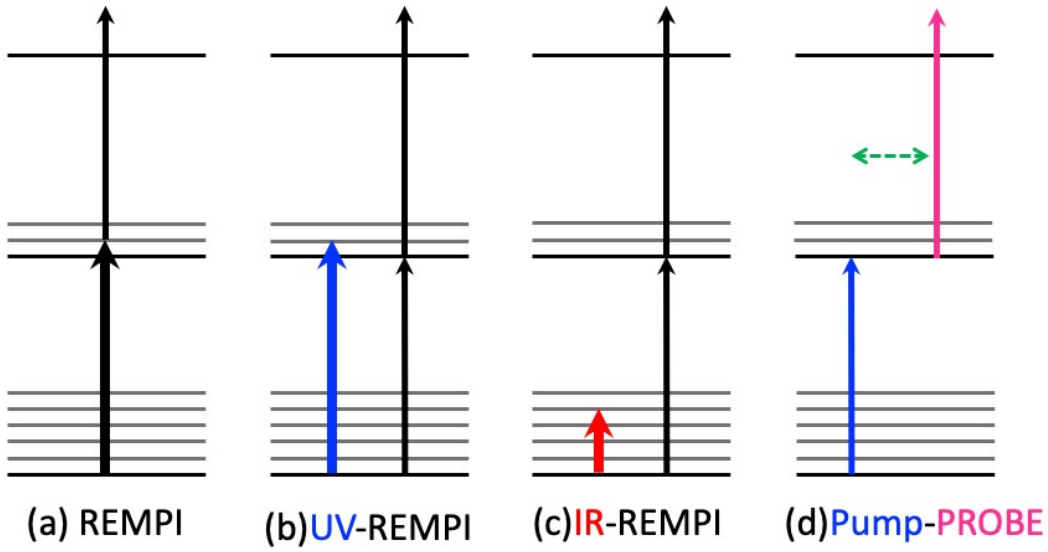


**Figure 5 Various resonant and double resonant multiphoton Ionization Techniques**

Another spectroscopy technique is called double resonant spectroscopy. Double resonance spectroscopy can provide extra information in the ionization step. In this technique, there is tunable "burn" laser pulse which can be either UV shown as blue colored arrow in figure 5(b) or IR, shown as red colored arrow in figure 5(c). When this burn laser is tuned to a resonance, it modifies/depletes the ground state population. The resonant ground state modification/depletion is then monitored by a "probe" laser that follows the burn laser, and it results in decrease in the ion signal. Due to the drastic drop in ion signals, this is also called the "hole-burning" technique. The excitation and ionization steps indicated in two black arrows in figure 5(b)&(c) serve as fixed wavelength probe. If both the excitation photon and ionization photon come from the same laser, it would be called 1-color resonant 2 photon Ionization (1C-R2PI). If they are from different laser, it would be called 2-color R2PI (2C-R2PI). With this technique, ground state IR or UV spectra of mass and optically selected species can be obtained.

For picosecond REMPI experiments, an EKSPLA PL2251 Nd:YAG laser is used, and it pumps two EKSPLA PG401 tunable optical parametric generators (OPG). Residual 1064nm and 532nm light straight from the pump laser is mixed harmonically to generate UV light 213, and this is used as the ionization pulse.[19] The fundamental has a wavelength of 1064nm, the second harmonic wavelength is 532nm, third 355, fourth 266nm, and fifth 213nm.

IR light is produced by using a Laser Vision tunable optical parametric oscillator/optical parametric amplifier (OPO/OPA). It can produce the mid-IR range of 3000-3800$cm^{-1}$ with ~1-2mJ/pulses, 3$cm^{-1}$ spectral line width. This IR laser can be used as the burn laser for IR double resonant spectroscopy. There are two variations. For the first variation, the ionization pulse is held constant at a resonant R2PI transition for a selected tautomer while the burn laser IR scans across the IR wavelength range.[19] Whenever it is resonant with vibrational states, the IR burn laser will deplete the ground state population which will result in non-resonant excitation by the excitation pulse and a decrease in ion signal. This first variation produces a tautomer specific IR spectrum. For the second variation, the ionization laser scans the R2PI spectrum while the burn IR laser is fixed at a resonant vibrational transition. [19] In this implementation, a tautomer will be labeled by their signature vibrational stretch such as OH and NH stretches.[19]

Lastly, pump-probe measurements can be performed to obtain excited state lifetimes. Pump-probe is the last variation of REMPI technique shown in figure 5(d). The pump laser pulse is indicated with a blue arrow, and a variable delay is added between this pump laser and the probe laser pulse, indicated with a pink arrow. The green double ended arrow indicates the time delay. Pump-probe measurements provide excited state lifetimes as a

function of delay between the excitation/pump pulses which is the first photon and the ionization/probe pulses which is the second photon.

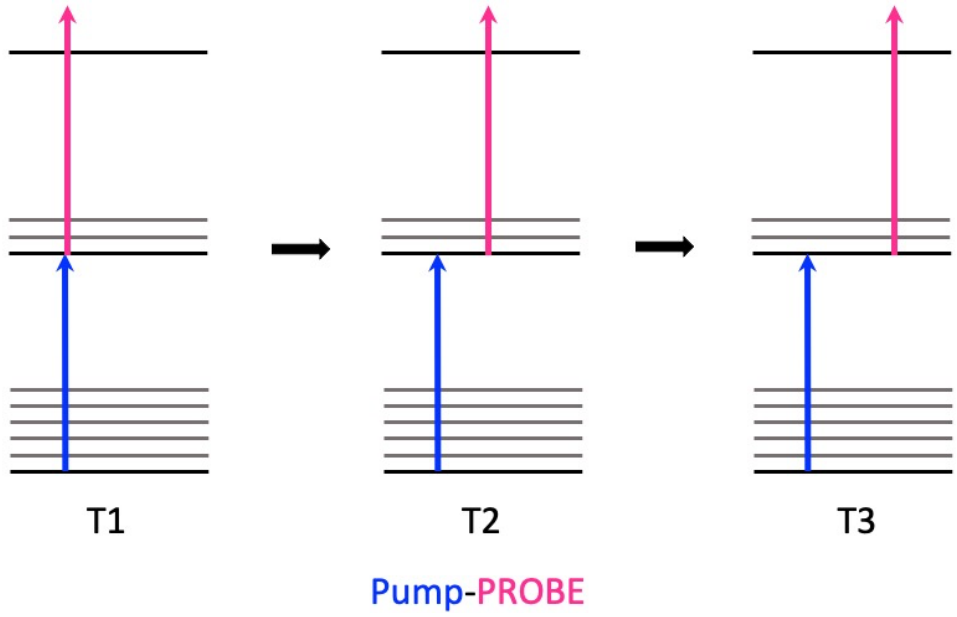


**Figure 6 Pump-Probe Measurements**

Figure 6 is provided to aid better understanding how the delay works. In the beginning pump laser pulses will be resonant and there will be no time delay between the pump laser and the probe laser. This will allow a strong ion signal. Over time, more and more time delay will be added between the excitation pulse and the ionization pulse which will cause a decrease in ion signal. This happens because there will be less and less excited state population (T1 < T2 < T3) that can be ionized by the ionization pulse, the second photon.

Figure 7 schematically shows pump-probe data. Ion signal on the Y-axis is plotted against time delay on the X-axis. As explained earlier, the highest ion signal is expected when the excitation pulse is overlapped with the ionization pulse. With more and more time delay between those two photons, the ion signal will decrease over time.

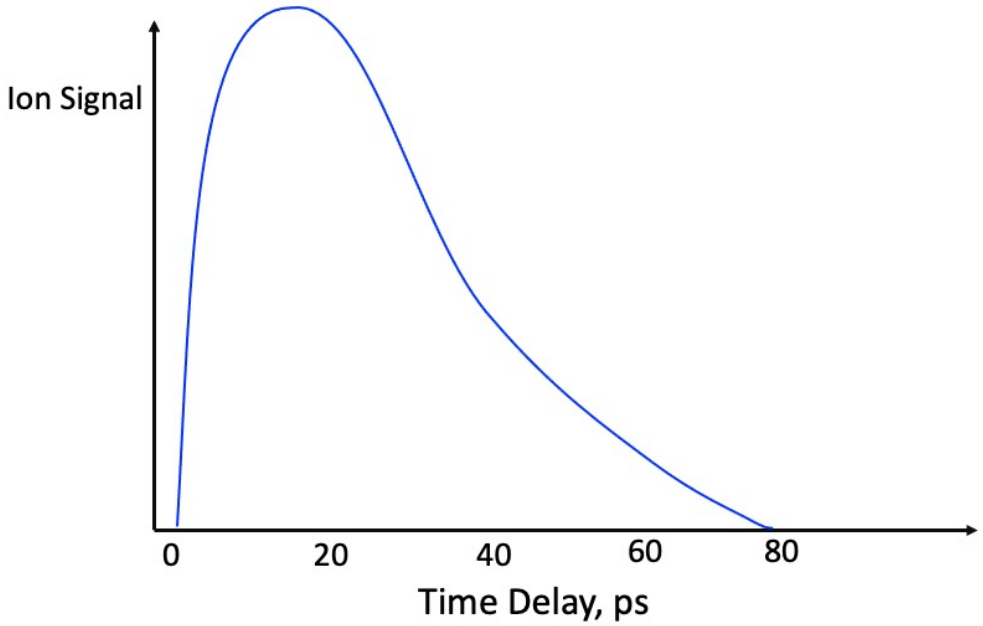


**Figure 7 Example Pump-Probe data on a Picosecond scale**

In the pump-probe experiments in picosecond regime, 30ps pump pulses are generated from the OPG laser system which is fixed to a resonant transition. The probe pulse is the 5$^{th}$ harmonic 213nm generated from the pump laser; this 5$^{th}$ harmonic can be mechanically delayed by up to ~1.5ns in time. Lifetimes are derived from the kinetic equation and ordinary differential equations, which includes convolving the instrument response function (IRF) with a mono-exponential or sometimes biexponential decay function. IRF is resented as a Gaussian function centered at around time at zero. Pump-probe experiments near the threshold for absorption provide opportunities to explore the potential energy surface/landscape which is sensitive to energy barriers and access to the most relevant conical intersections (CIs) .

# III. LITERATURE REVIEW

## *A. Successful Laser Desorption on Canonical Nucleobases*

The following literature review in this chapter demonstrates how well the mentioned experimental set up and methods are suited for studying the excited state dynamics of nucleobases and their alternatives and derivatives.

Nir et al. (1999) successfully presented the first vibronic spectroscopy of guanine in the gas phase.[1] Vibrational and electronic states of DNA bases have been studied experimentally for a long time in solution and in the gas vapor. Gas phase studies can provide high vibrational resolution, especially with entrainment or seeding of the molecules in a supersonic expansion which reduces the internal temperature by eliminating intermolecular interactions.[1] In a typical implementation by Nir et al. samples were deposited on translating graphite bar in a thin uniform layer.[1] Desorption was performed by a Nd:YAG laser at 1064nm.[1] The desorbed molecules were entrained in a supersonic expansion and with argon as the drive gas.[1] Jet-cooled gaseous sample was injected by a pulsed solenoid valve.[1] Desorption occurred on the vacuum side of the valve.[1] The authors tested different drive gases other that argon to determine the spectra were not affected by cluster formation which could induce different spectral shifts depending on the drive gases.[1] The entrained samples were two-photon ionized and the ions were detected with a reflectron time-of-flight mass spectrometer.[1] The second photon was either produced from the same laser as the first photon (1C-REMPI) or from a different laser (2C-REMPI).[1] When 2C-REMPI was utilized, the authors varied the wavelength of the first photon while monitoring selected mass peaks to obtain mass selected excitation spectra.[1] The lowest energy peak that was observed was at 32,878 $cm^{-1}$.[1] After further scanning to the red of this peak with 2C ionization without

observing any additional peaks, the lowest energy peak was determined to be 0-0 origin transition.[1] 0-0 origin transition refers to electronic transitions from and to the lowest vibrational states (v = 0). Th authors compared their REMPI spectra against IR data on polycrystalline samples and determined that they were in fair agreement.[1] Their REMPI spectra represented excited state vibronic transitions including the origin transition while the IR data probed the molecular ground state only.[1] Guanine usually stays in the keto form rather than in enol form in solution, and the keto is the more stable tautomeric form.[1] It is important to be able to study various tautomeric forms of the nucleobases, because the keto conformation is absolutely required for hydrogen bonding in Watson-Crick pair formation.[1] In the polycrystalline form, guanine exists in the keto form mostly which can be proven from the C=O stretch in the IR spectrum.[1] However, in the REMPI spectra the enol tautomer seems to exist [1] It is likely that the most stable tautomeric form of guanine would be frozen predominantly in the jet expansion. Additional measurements at shorter wavelengths and hole burning experiments are required to confirm the existence of multiple tautomeric forms in the jet cooled environment.[1]

Later, Nir et al. (2001) performed UV-UV and IR-UV hole burning spectroscopy and were able to identify three different tautomeric forms of guanine that that together constituted the R2PI spectrum from their previous work.[2] Resonant two photon ionization (R2PI) was obtained with the same experimental set up mentioned earlier.[1,2] They used two dye laser pulses that travel from the opposite direction with a delay of approximately 150ns to perform UV-UV hole burning experiment.[2] The first laser as the burn laser and the second laser as the probe laser, this produced two signals on the time-of-flight spectrum.[2] Whenever both the lasers are resonantly tuned to the same tautomer, the burn laser caused a drop in the

signal intensity.[2] The burn laser is scanned while the probe laser is fixed to a strong band of selected tautomer.[2] If a characteristic band from the R2PI is absent in the burn spectrum, that band belongs to different tautomer.[2] The next step was to probe at that frequency while the pump laser is scanning to obtain the spectrum for the next tautomer.[2] IR-UV hole burning spectroscopy used the same method but with a different frequency IR laser used as the burn laser.[2] Based on those tautomer selected IR spectra, the three tautomeric forms were initially assigned 9H-Enol, 9H- Keto, and 7H-Keto.[2] However, due to the experimental limit of wavelength range These experiments were not able to measure the C=O region, which would be most diagnostic for the keto form. In fact, later analysis revealed that the two tautomers besides the enol form were actually imino-tautomers. The keto tautomers are the lowest energy forms and are observed in the gas phase in helium drop experiments and in microwave spectroscopy in molecular beams. Both of those experiments measure ground state populations, while the REMPI experiments represent action spectroscopy of excited state populations. The most likely explanation for the absence of the keto tautomers in the multiphoton ionization experiments is therefore an extremely short excited-state lifetime which precludes ionization out of the excited state.

These authors not only assigned three tautomers in the gas phase in their R2PI spectrum, they also assigned tautomers for modified guanine, 1-methylguanine. 1-methylguanine is methylated at N1 which prevents keto-enol tautomerism. R2I and UV-UV hole burning spectra of guanine and 1-methylguanine were compared. UV-UV hole burning spectra of 1-methylguanine showed two tautomers exist in the gas phase which were assigned as 9H-keto and 7H-keto.[2] UV-UV hole burning spectra of two keto tautomers of 1-methylguanine look very similar to each other which is in accordance with the fact that these two keto

tautomers differ only at 9H and 7H position.[2] The OH and NH fingerprints on IR are very peculiar for each keto and enol forms but it was not intensive in their vibronic spectra.[2] For this specific reason the work obtained isomer specific/selective IR spectra generated by IR-UV hole burning spectroscopy.[2] Two (*a, b*) out of three (*a, b, c*) tautomer specific IR-UV spectra of guanine resembled the ones (*a', b'*) of 1-methylguanine.[2] Guanine tautomers *a* and *b* had their two significant bands at 3505/3489 and 3505/3495 $cm^{-1}$. Guanine tautomer *c* had four bands at 3587, 3577, 3516, and 3462 $cm^{-1}$. 1-methylguanine tautomers *a'* and *b'* had very similar two significant bands like the ones guanine tautomers *a* and *b* do at 3505/3489 and 3505/3495 $cm^{-1}$.

### *B. Effects of Substituents on the Modified DNA Bases: Pyrimidine Derivatives*

For a long time, the focus used to lay mainly on studying the photostability of the canonical nucleobases. However, recently there has been ample experimental and theoretical work done to answer the question, how do alternative bases which were possible 'candidates' in prebiotic era respond to UV exposure? In order to investigate the full picture of the molecular origins of life, the electronic relaxation mechanisms of purine and pyrimidine derivatives/candidates should be studied also due to their possible important roles on early earth. Brister et al. (2016) investigated barbituric acid (BA) and 2,4,6-triaminopyrimidine (TAP) in solution with a spectroscopic technique, transient absorption spectroscopy.[3] Brister et al. (2016) selected these two pyrimidine derivatives based on the findings of Cafferty and Hud (2015). [3,20] Cafferty and Hud created a list of heterocycles as the list of possible candidates for the original bases of proto-RNA.[20] Proto-RNA refers to the oldest ancestor of RNA. These heterocycles include heterocycles that have only H, $NH_2$, or O as exocyclic groups.[20] From that list they crossed out the ones that would not be

compatible with Watson-Crick base pairing, and the ones that are known not to form nucleosides with sugar moieties in prebiotic conditioned simulations.[20] Cafferty and Hud were able to identify two candidates, barbituric acid (BA) and 2,4,6-triaminopyrimidine (TAP) as promising possible nucleobases of proto-RNA.[20] The structure of BA is similar to that of the Uracil nucleobase but with an additional carbonyl group at C6 position and an additional hydrogen atom at C5 position.[3] BA is missing the C5-C6 ethylenic bond of uracil. For this reason, this study used the previously published computed data about uracil to gain insight in the possible conical intersection between the $1\pi\pi^*$ state and the $S_0$ state for both the enol and the tri-keto tautomers of BA.[3] There are two CIs that intersect between the $1\pi\pi^*$ state with the $S_0$ state in uracil which are either ethylenic $1\pi\pi^*$/ $S_0$ CI or ring opening $1\pi\pi^*$/ $S_0$ CI. [3] The ethylenic CI is induced by a twist of C5-C6 ethylenic bond and the ring opening CI is induced by the aromatic ring of the pyrimidine opening at the N3-N4 bond.[3] The authors argued that the ethylenic CI gives more satisfactory explanation for the ultrafast relaxation of the enol(C6) and tri-keto tautomers of BA.[3] The structure of TAP is similar to that of the nucleobase cytosine, but with both the carbonyl group at C2 and the hydrogen atom at C6 replaced with amino groups.[3] For this reason, previously published computed data about cytosine can be applied in which two CIs play a key role in the ultrafast relaxation of electronic energy from the $1\pi\pi^*$ state to the ground state.[3] There are a C6 puckered CI and an out-of-plane $NH_2$ (oop - $NH_2$) CI.[3] This suggests that analogous C6-puckered and oop - $NH_2$ CIs could play an important role in the ultrafast electronic relaxation in TAP.[3] A C6-puckered CI could explain the relaxation from the $1\pi\pi^*$ state to the $S_0$ state, and an oop - $NH_2$ CI could explain the relaxation from the $1n\pi^*$ states to the $S_0$ state, respectively.[3]

The aforementioned promising proto-RNA candidates BA and TAP have ultrafast sub-picosecond excited state lifetimes which could grant them photostability under UV radiation. However, other nucleobase alternatives such as 2-aminopurine and 2,6-diaminopurine present much longer excited state lifetimes which can lead to destructive photochemical reactions. Studying the effects of modifications of the heterocyclic ring such as paralyzing certain parts of the ring by substitution at certain locations that could affect the accessibility to conical intersections can elevate the understanding of the intrinsic photostability of alternative nucleobases.[4] Gengeliczki et al. (2010) performed resonant two-photon ionization (R2PI) and IR-UV hole burning spectroscopy on two of the 2-amino substituted analogs, 2,4-diaminopyrimidine (2,4-DAPy) and 2,6-diaminopurine (2,6-DAPu).[4] Amino-substitution at C2 would affect the puckering induced conical intersection, and the addition of an imidazole ring to form purine base would affect the conical intersection that is involved with ring deformation at C5 and C6.[4] Therefore, puckering at the C2 position is hindered for 2,4-DAPy and puckering at both positions are hindered for 2,6-DAPu.[4] Laser spectroscopy, including R2PI, UV-UV, and IR-UV hole burning was conducted by the de Vries lab with the previously explained experimental method. R2PI spectra of 2,4-DAPy and 2,6-DAPu were obtained with the wavenumber range between 32,000 and 36,000 $cm^{-1}$. IR-UV hole burning spectroscopy along with DFT and *ab initio* calculations showed the presence of the most stable tautomers of both compounds.[4] The $S_0$ to $S_1$ ($\pi\pi^*$) origin transition of 2,4-DAPy was determined to be at 34,459 $cm^{-1}$.[4] The excited state lifetime that was measured at this origin wavelength was longer than 10 ps but shorter than the of ~1ns time resolution limit of the instrumentation.[4] only the diamino tautomers 2,4-DAPy present in the experiment.[4] The IR-UV hole burning spectra showed N7 and N9

tautomers of 2,6-DAPu in the cold jet cooled molecular beam.[4] The $S_0$ to $S_1$ ($\pi\pi^*$) origin transition of N7 tautomer was at 32,215 $cm^{-1}$, and that of N9 tautomer was at 34,881 $cm^{-1}$.[4] The measured lifetimes for N7 and N9 tautomers of 2,6-DAPu were 8.7± 0.8 ns and 6.3± 0.4 ns respectively.[4] These longer lifetimes agree with the restriction of ring deformations in both the C2 position and the C5=C6 twist.[4] These modifications hinder the access to all conical intersections that lead 4-aminopyrimidine to have fast excited state dynamics.[4]

### *C. Photo-dynamics of alternative DNA base Isocytosine and Isoguanine*

Two promising proto-RNA candidates, 2,4,6-triaminopyrimidine and barbituric acid were studied as mentioned above, and they presented ultrashort sub-picosecond excited state lifetimes which suggest strong photostability against UV exposure.[3] On the other hand, other alternative nucleobases such as 2,4-aminopurine and 2,6-diaminopurine presented much longer excited state lifetimes which could lead to harmful photochemical reactions.[4] An experimental and theoretical study done on isocytosine(isoC) showed that isoC in the enol form follows a relatively efficient radiation-less deactivation pathway and isoC in the keto form has a longer lived $S_1$ ($n\pi$) state.[6,21,22] The alternative bases isocytosine (isoC) and isoguanine (isoG) can be compared to Watson-Crick (WC) canonical base pairs of the triple hydrogen bonded guanine/cytosine pair (G/C).[6] Isocytosine (isoC) is an isomer of C but also an analogue of guanine (G), and Isoguanine (isoG) is an isomer of G but also an analogue of Cytosine (C).[6] isoC shares the same core moiety of G and isoG shares the same core moiety of C.[6] Berenbeim et al. (2017) investigated the photochemistry of isoC by performing 2C-R2PI, IR-UV hole burning spectroscopy, pump-probe experiments.[6] This work confirms the presence of two tautomeric forms of isoC, EA1 (enol) and KA2 (keto) out of six lowest energy tautomers in the gas phase.[6] Nanosecond scale 2C-R2PI spectra showed the origin

transition at 35, 292 $cm^{-1}$, and picosecond scale 2C-R2PI spectra showed a significant band at 33,266 $cm^{-1}$.[6] IR-UV hole burning spectroscopy revealed the presence of the EA1 tautomer with the origin transition at 35, 292 $cm^{-1}$, and the KA2 tautomer with the featured band at 33,266 $cm^{-1}$ after the spectra were compared to anharmonic simulations.[6] KA2 is calculated to be the highest energy structure out of six tautomers and the fact that KA2 was present in the jet cooled molecular beam suggests the presence of the other lower energy tautomers as well.[6] Since jet-cooling is not an equilibrium process, the tautomer distribution could not be determined; however, usually the lowest energy tautomers up to approximately 50 kJ/mol are present in the experimental set up in the de Vries Lab.[6] There are three reasons why the other tautomers might not have been observed. One, some tautomers might absorb in a different UV range.[6] Second, the experimental technique is a form of action spectroscopy rather than direct absorption which means that a molecule can be excited by the first photon but not ionized by the second photon.[6] This can occur when the excited state lifetime is much shorter than the ionizing laser pulse length.[6] Third, a molecule might go through fragmentation after either excitation or ionization which could complicate the spectrum when only the parent ion mass is monitored.[6] Pump-probe experiments were performed on the origin transitions of EA1 and KA2, respectively. EA1 had lifetimes of 490 ps and 478 ns at the origin transition at 35, 292 $cm^{-1}$, and KA2 had a lifetime of 69 ps at the origin at 33,266 $cm^{-1}$.[6] The excited state dynamics of guanine might depend on the heterocyclic part of the pyrimidine moiety which is identical to isoC chemically.[6] WC isoC which is in KA2 form exhibited one excited state decay channel with a rate of approximately $10^{10}$ $s^{-1}$, and the enol form EA1 had multiple decay rates with the fastest one being 7 times longer than it was for WC isoC (KA2).[6] Berenbeim et al. (2017) found that the excited state

dynamics presented more similarities to that of guanine than to the photodynamics of cytosine.[6]

Gate et al. (2019) continued to study on alternative DNA base, and this time on isoguanine(isoG).[23] Isoguanine is a Watson-Crick (WC) partner of isocytosine, and it is also referred to as 2-hydroxyadenine.[23] There is some to explain why isoG might not have been utilized by nature as one of the canonical nucleobases. IsoG is found incorporated in DNA every now and then, and can pair with cytosine to form a parallel strand or pair with isoC to form an anti-parallel strand.[24–26] IsoC-isoG pairing is thermodynamically as stable as the canonical G-C WC pair and can actually be used as working genetic code which could result in translation of non-biological amino acid.[24–26] Gate et al. (2019) performed laser desorption jet cooling resonance enhanced multi photon ionization (REMPI), two- color resonance two-photon ionization (2C-R2PI), and IR-UV hole burning spectroscopy, and pump-probe measurements in the pico and nano second time regimes.[23] The experimental set-ups were identical to those explained earlier. 2C-R2PI was performed first to determine the transitions for possible isoG tautomers that are present in the jet cooled molecular beam.[23] gate et al. assigned the origin transitions at 30,807 $cm^{-1}$ to the keto-N3,7 tautomer and the transition at 34,340 $cm^{-1}$ as the enol-N7 tautomer after IR-UV hole burning spectroscopy was performed.[23] For IR-UV hole burning spectroscopy, the IR pulse came 200 ns earlier than the ionization pulse while the IR laser scanned across a target wavelength range with the ionization laser (UV) fixed at a constant R2PI transition wavelength (30,807 $cm^{-1}$ and 34,340 $cm^{-1}$ respectively) for each specific tautomer.[23] Whenever the IR pulse burned and depleted a resonant vibrational transition, it prevented the resonance transition that is required for ionization leading to decline in ion signal.[23] Keto-N3,7 and enol-N7

tautomers are relatively higher energy tautomeric forms of isoG compared to enol-N9, keto-N3,9 and keto N1,9.[23] There are possible reasons why the lower energy tautomers were not observable in their experimental set up. One, once it is laser desorbed, isoG might have been trapped in local minima of higher energy tautomers during the jet cooling which would block the access to other lower energy tautomers.[23] Second, if their excited state lifetime is ultra-short, even after the tautomers are excited with the first photon, they might not get ionized if the excited state lifetime is much shorter than the ionization pulse length.[23] Lastly, some of the tautomers might go through fragmentation during the photoexcitation.[23] Gate et al. (2019) also incorporated quantum chemical simulations along with their spectroscopic measurements.[23] Through 2C-R2PI they were able to pick and probe three peaks for keto-N3,7 presenting a lifetimes of 941, 972, and 932 ps.[23] This can be explained by the sloped landscape of the $\pi\pi^*/ S_0$ CI which is much higher in energy compared to its $S_1$ minimum.[23] On the other hand, the peaks for enol-N7 tautomers had poor signal to noise ratio followed by spectral broadening which could indicate their ultrafast excited state dynamics.[23] The enol isoG tautomers are affected by $\pi\pi^*$ states significantly with ultrashort lifetimes.[23] The biologically active and relevant keto forms presented long excited state lifetimes that could lead to possibly destructive photochemistry, and enol tautomers presented relatively short lifetimes which indicates photostability.[23] Their results can give a clue as to why isoG might not have been selected by nature to be part of canonical nucleobases.

# IV. SUMMARY

The chosen collection of the molecular building blocks of life by nature was most likely mediated by a chemical selection during prebiotic era. Prebiotic selection pressure might have involved photochemistry. Here in de Vries Lab, we study the role of UV radiation on photochemistry of nucleobases and their alternatives. Canonical nucleobases have the intrinsic property of "UV hardiness" that provides photostability upon the UV exposure. The canonical nucleobases tend to decay out of the excited state back to the ground state within 1 picosecond and this intrinsic property depends on the molecular structure. This structure dependency derives from rapid internal conversion *via* conical intersections (CI). Laser Desorption Jet-Cooling Mass Spectrometry in gas phase is utilized in de Vries Lab. Combined with high level theoretical computation, our experimental technique provides the landscape of photo-dynamics of the nucleobases and their alternatives. Ionization technique, Resonance Enhanced Multi-Photon Ionization (REMPI) is used. Its variation techniques include 2Color-REMPI, UV-UV and IR-UV hole burning spectroscopy or double resonance spectroscopy, and pump-probe measurement. Several optimal examples of the studies that utilized the mentioned experiment set up were included. Nir et al. (1999) successfully examined guanine with REMPI technique providing its origin transition.[1] Nir et al. (2001) successfully identified three tautomers of guanine using double resonance/ hole burning technique.[2] Brister et al. (2016) investigated promising nucleobase alternatives in solution such as barbituric acid and 2,4,6-triaminopyrimidine, and found out their rapid decay out of the excited state back to the ground state.[3] On the contrary, Gengeliczki et al (2010) found some nucleobase alternatives such as 2-aminopurine and 2,6-diaminopurine to present longer excited-state lifetimes which can lead to potentially harmful photochemistry

reactions.[4] Lastly a pair of alternative nucleobases isocytosine (isoC) and isoguanine (isoG) were studied in de Vries Lab.[5,6] The possible photochemical reasons why cytosine and guanine were selected by nature instead of their alternatives, isoC and isoG were explored.[5,6]

## REFERENCES


(1) Nir, E.; Grace, L.; Brauer, B.; de Vries, M. S. REMPI Spectroscopy of Jet-Cooled Guanine. *J. Am. Chem. Soc.* **1999**, *121* (20), 4896–4897. https://doi.org/10.1021/ja984088g.

(2) Nir, E.; Janzen, C.; Imhof, P.; Kleinermanns, K.; de Vries, M. S. Guanine Tautomerism Revealed by UV–UV and IR–UV Hole Burning Spectroscopy. *J. Chem. Phys.* **2001**, *115* (10), 4604–4611. https://doi.org/10.1063/1.1391443.

(3) Brister, M. M.; Pollum, M.; Crespo-Hernández, C. E. Photochemical Etiology of Promising Ancestors of the RNA Nucleobases. *Phys. Chem. Chem. Phys. PCCP* **2016**, *18* (30), 20097–20103. https://doi.org/10.1039/c6cp00639f.

(4) Gengeliczki, Z.; Callahan, M. P.; Svadlenak, N.; Pongor, C. I.; Sztáray, B.; Meerts, W. L.; Nachtigallová, D.; Hobza, P.; Barbatti, M.; Lischka, H.; De Vries, M. S. Effect of Substituents on the Excited-State Dynamics of the Modified DNA Bases 2,4-Diaminopyrimidine and 2,6-Diaminopurine. *Phys. Chem. Chem. Phys. PCCP* **2010**, *12* (20), 5375–5388. https://doi.org/10.1039/b917852j.

(5) Gate, G.; Szabla, R.; Haggmark, M. R.; Šponer, J. í; Sobolewski, A. L.; de Vries, M. S. Photodynamics of Alternative DNA Base Isoguanine. *Phys. Chem. Chem. Phys. PCCP* **2019**, *21* (25), 13474–13485. https://doi.org/10.1039/c9cp01622h.

(6) Berenbeim, J. A.; Boldissar, S.; Siouri, F. M.; Gate, G.; Haggmark, M. R.; Aboulache, B.; Cohen, T.; de Vries, M. S. Excited-State Dynamics of Isocytosine: A Hybrid Case of Canonical Nucleobase Photodynamics. *J. Phys. Chem. Lett.* **2017**, *8* (20), 5184–5189. https://doi.org/10.1021/acs.jpclett.7b02032.

(7) Patterson, C. Age of Meteorites and the Earth. *Geochim. Cosmochim. Acta* **1956**, *10* (4), 230–237. https://doi.org/10.1016/0016-7037(56)90036-9.

(8) Lyons, T. W.; Reinhard, C. T.; Planavsky, N. J. The Rise of Oxygen in Earth's Early Ocean and Atmosphere. *Nat. Lond.* **2014**, *506* (7488), 307–315. https://doi.org/10.1038/nature13068.

(9) Fry, I. The Origins of Research into the Origins of Life. *Endeav. New Ser.* **2006**, *30* (1), 24–28. https://doi.org/10.1016/j.endeavour.2005.12.002.

(10) MILLER, S. L. A Production of Amino Acids under Possible Primitive Earth Conditions. *Sci. Am. Assoc. Adv. Sci.* **1953**, *117* (3046), 528–529. https://doi.org/10.1126/science.117.3046.528.

(11) Schuch, A. P.; Moreno, N. C.; Schuch, N. J.; Menck, C. F. M.; Garcia, C. C. M. Sunlight Damage to Cellular DNA: Focus on Oxidatively Generated Lesions. *Free Radic. Biol. Med.* **2017**, *107*, 110–124. https://doi.org/10.1016/j.freeradbiomed.2017.01.029.

(12) Kang, H.; Lee, K. T.; Jung, B.; Ko, Y. J.; Kim, S. K. Intrinsic Lifetimes of the Excited State of DNA and RNA Bases. *J. Am. Chem. Soc.* **2002**, *124* (44), 12958–12959. https://doi.org/10.1021/ja027627x.

(13) Saenger, W. Forces Stabilizing Associations Between Bases: Hydrogen Bonding and Base Stacking. In *Principles of Nucleic Acid Structure*; Saenger, W., Ed.; Springer New York: New York, NY, 1984; pp 116–158. https://doi.org/10.1007/978-1-4612-5190-3_6.

(14) Haas, Y.; Zilberg, S. Photochemistry by Conical Intersections: A Practical Guide for Experimentalists. *Theor. Org. Photochem.* **2001**, *144* (2), 221–228. https://doi.org/10.1016/S1010-6030(01)00451-8.

(15) Meijer, G.; de Vries, M. S.; Hunziker, H. E.; Wendt, H. R. Laser Desorption Jet-Cooling of Organic Molecules. *Appl. Phys. B* **1990**, *51* (6), 395–403. https://doi.org/10.1007/BF00329101.
(16) Levy, D. H. The Spectroscopy of Very Cold Gases. *Sci. Am. Assoc. Adv. Sci.* **1981**, *214* (4518), 263–269. https://doi.org/10.1126/science.214.4518.263.
(17) *The Encyclopedia of Mass Spectrometry*, 1st ed.; Gross, M. L., Caprioli, R. M., Eds.; Elsevier: Amsterdam ; Boston, 2003.
(18) Hall, R. B. Pulsed-Laser-Induced Desorption Studies of the Kinetics of Surface Reactions. *J. Phys. Chem. 1952* **1987**, *91* (5), 1007–1015. https://doi.org/10.1021/j100289a003.
(19) Haggmark, M. R.; Gate, G.; Boldissar, S.; Berenbeim, J.; Sobolewski, A. L.; de Vries, M. S. Evidence for Competing Proton-Transfer and Hydrogen-Transfer Reactions in the S1 State of Indigo. *Chem. Phys.* **2018**, *515*, 535–542. https://doi.org/10.1016/j.chemphys.2018.09.027.
(20) Cafferty, B. J.; Hud, N. V. Was a Pyrimidine-Pyrimidine Base Pair the Ancestor of Watson-Crick Base Pairs? Insights from a Systematic Approach to the Origin of RNA. *Isr. J. Chem.* **2015**, *55* (8), 891–905. https://doi.org/10.1002/ijch.201400206.
(21) Hu, D.; Liu, Y. F.; Sobolewski, A. L.; Lan, Z. Nonadiabatic Dynamics Simulation of Keto Isocytosine: A Comparison of Dynamical Performance of Different Electronic-Structure Methods. *Phys. Chem. Chem. Phys. PCCP* **2017**, *19* (29), 19168–19177. https://doi.org/10.1039/C7CP01732D.
(22) Szabla, R.; Góra, R. W.; Šponer, J. Ultrafast Excited-State Dynamics of Isocytosine. *Phys. Chem. Chem. Phys. PCCP* **2016**, *18* (30), 20208–20218. https://doi.org/10.1039/c6cp01391k.
(23) Gate, G.; Szabla, R.; Haggmark, M. R.; Šponer, J.; Sobolewski, A. L.; de Vries, M. S. Photodynamics of Alternative DNA Base Isoguanine. *Phys. Chem. Chem. Phys.* **2019**, *21* (25), 13474–13485. https://doi.org/10.1039/C9CP01622H.
(24) Yang, X.-L.; Sugiyama, H.; Ikeda, S.; Saito, I.; Wang, A. H.-J. Structural Studies of a Stable Parallel-Stranded DNA Duplex Incorporating Isoguanine:Cytosine and Isocytosine:Guanine Basepairs by Nuclear Magnetic Resonance Spectroscopy. *Biophys. J.* **1998**, *75* (3), 1163–1171. https://doi.org/10.1016/S0006-3495(98)74035-4.
(25) Roberts, C.; Bandaru, R.; Switzer, C. Theoretical and Experimental Study of Isoguanine and Isocytosine: Base Pairing in an Expanded Genetic System. *J. Am. Chem. Soc.* **1997**, *119* (20), 4640–4649. https://doi.org/10.1021/ja970123s.
(26) Seela, F.; Wei, C.; Melenewski, A. Oligonucleotides Containing Consecutive 2′-Deoxy-Isoguanosine Residues: Synthesis, Parallel Duplex Formation and Identification of a d(T4iG4T4) Tetraplex. *Nucleosides Nucleotides* **1997**, *16* (7–9), 1523–1527. https://doi.org/10.1080/07328319708006221.