A theoretical investigation into complex organic molecules in the Interstellar Medium

Eren C. S. Slate

Department of Chemistry

Supervisor: Prof. Dr. Anthony J. H. M. Meijer

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

December 9, 2019

Abstract

This work aims to understand the formation routes of a variety of interstellar molecules. The formation of molecules in the interstellar medium (ISM) is a key process to understanding the chemical diversity of the universe as well as the origins of biological life on Earth. Some of the molecules (methyl formate and glycolaldehyde) under study have been identified in the ISM *via* telescopic detection. Other molecules investigated have been proposed to be present in the ISM as a result of terrestrial laboratory studies on simulated ISM ices (urea, ribose and benzoic diacids). Quantum chemical calculations were used to follow the reaction pathways of these molecules and test their feasibility under astrochemical conditions. It was found that the presence of hydrogen bonding ices around the reaction species greatly affects the energetics of the reactions and increases their likelihood to occur in the ISM. Additionally, water ice was found to have a large impact on the dissociation of methanol into radical species allowing for much more facile subsequent reactions than previously thought. Water ice was also found to affect the ease of ionisation of polycyclic aromatic hydrocarbon (PAH) species which have been proposed as a constituent of ISM dust grains.

Acknowledgements

Contents

Glossary

Arcsecond (arcsec) Angular measurement used to describe the location of an object in space. Each degree of a circle is divided into 60 arcminutes and each arcminute into 60 arcseconds.

Astronomical unit (AU) The average distance between the Earth and the Sun, 1 AU = 1.49 $\times 10^{11}$ m.

Complete Active Space Perturbation Theory nth **order (CASPTn)** A method for including electron correlation into multi-configurational calculations using perturbation theory where n stands for the order of perturbation used.

Complex Organic Molecule (COM) A molecule found in space that is composed of several atoms, primarily carbon, hydrogen, nitrogen and oxygen.

Configuration Interaction (CI) A method for solving the Schrödinger equation that makes use of multiple determinants, hence allowing for multiple electron configurations.

Cosmic Ray High-energy radiation originating, largely, from outside of the Solar System.

Density Function Theory (DFT) A computational, quantum mechanical method for finding the properties of a many-body system using the density of the electrons.

Hartree-Fock (HF) A single determinant method of approximately determining the wave-function and energy of a many-body system.

Heliopause The pressure boundary at which the Sun's solar wind is no longer strong enough to push back the stellar winds of other nearby stars.

Hot core A small, dense and hot area, occurring in regions of star formation, that is important in the formation of molecules in the ISM.

Interstellar Medium (ISM) The regions of space between stars, the boundary between a star and the ISM is usually defined to be the heliopause.

Light-year (ly) The distance travelled by light in 1 year, 1 ly $= 9.5 \times 10^{15}$ m.

Molecular Mechanics (MM) A means for finding the energy of a system using classical models.

Molecular Cloud A dust cloud in the ISM with the correct physical conditions to allow for the formation of molecules. They can also be star-forming regions.

Multi-configurational Self-consistent Field (MCSCF) Another quantum chemistry method that allows for multiple electronic configurations within a system that employs both CI-like and HF-like techniques.

Nebula An interstellar cloud largely containing dust, ionised gases, hydrogen and helium.

our Own N-layered Integrated molecular Orbital molecular Mechanics (ONIOM) A method for combining two or more computational methods, such as molecular modelling and quantum mechanics, to treat a system.

Parallax Half of the measured angle through which a star appears to move during the year.

Parsec (pc) A star with a parallax of 1 arcsec is 1 parsec away, 1 pc $= 3.08 \times 10^{16}$ m.

Protoplanet An early stage in planet formation at which a differentiated core has formed within the protoplanetary disk as a result of internal melting.

Protoplanetary Disk A rotating disk of dense gas and dust, surrounding a newly formed star, in which planets can form.

Protostar A young star, still in the first stage of star formation, that is gaining mass from a molecular cloud.

Solar Mass ($M_\odot$) The mass of the Sun, 1 $M_\odot = 2.0 \times 10^{33}$ g.

Supernova An astronomical event during the last stages of stellar evolution for a massive star wherein it is destroyed.

Time-dependent DFT (TD-DFT) A variety of DFT that allows for the investigation of the properties of a system inside a time-dependent potential such as an external magnetic field.

1. Introduction

1.1 A brief inventory of space

1.1.1 Space is empty

Much of the Universe is empty. It is also vast (our solar system alone spans about 150 000 astronomical units, AU, and the nearest star to our Sun is 300 000 AU away, where $1\text{AU} = 1.49 \times 10^{11}$ m) and cold (as low as some tens of Kelvin).[1] There are regions where defining a temperature is difficult as there are so few particles present. There is also an abundance of UV photons and cosmic rays. Cosmic rays are high-energy particles in excess of 100 MeV nucleon^{-1}.[2] Given the levels of ambient radiation, it seems unlikely that any chemical species would survive long. However, as will be discussed in the subsequent sections, this is incorrect.

The objects that can be found within the Universe, stars and comets and planets, are familiar to everyone. What may be less familiar is the nature of the gaps, as it were, between these objects. These 'gaps' are what is referred to as the interstellar medium (ISM). The ISM is defined to be everything beyond the stellar envelope, which stops at the edge of the heliopause. The heliopause is the term for the pressure boundary between the solar wind (the stream of charged particles emitted from the Sun's surface, referred to as a stellar wind for stars in general) pushing outwards and the ISM pushing in, see figure 1.1.[1]
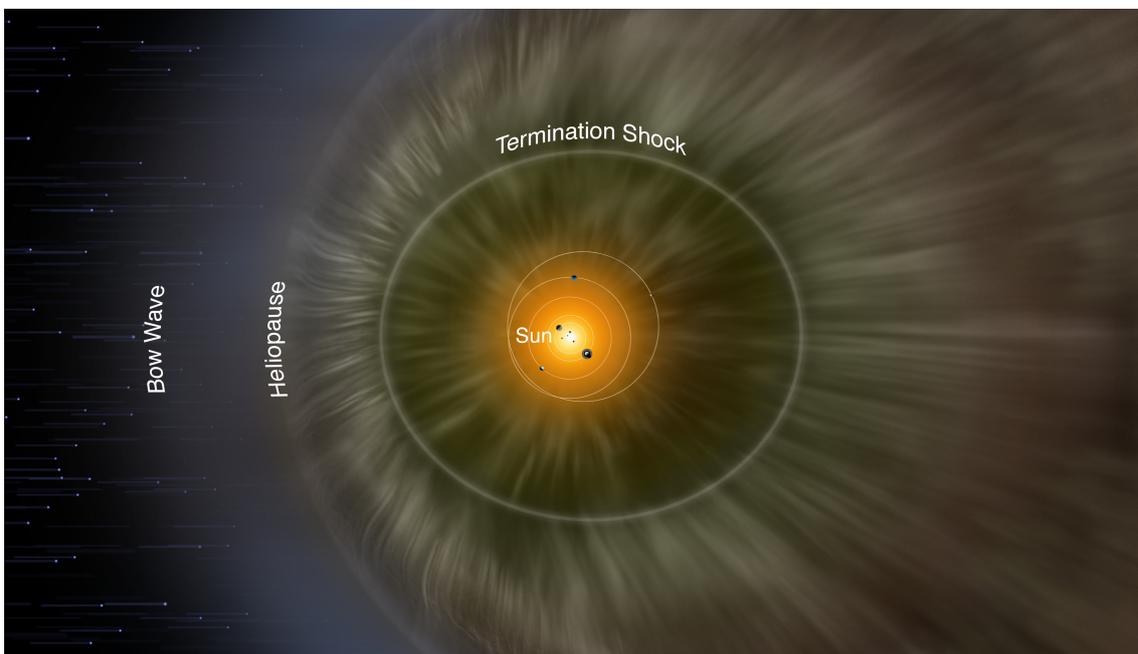
Figure 1.1: A diagram showing the heliopause of a star, image credit NASA/IBEX/Adler Planetarium.

As noted above, one might expect the ISM to be uniform and devoid of interesting chemistry. In fact, it is quite the opposite. The ISM is far from uniform, with a temperature range of 10^6 K down to as low as 10 K and densities from 10^8 particles cm^{-3} to 10^{-4} cm^{-3}.[3] To put this in context, the most dense region of the ISM is several orders of magnitude less dense than the highest vacuum that can be achieved on Earth. The vacuum system at the CERN Large Hadron Collider yields a 10^{-11} mbar pressure which is described as 'almost as rarefied as that found on the surface of the Moon'[¶] whose surface pressure is given by NASA as 2×10^5 particles cm^{-3}[**].[2] The ISM contains the ever-present hydrogen. It is 1000 times more abundant than any other more chemically interesting species, but also found in the ISM are helium, lithium, and heavier elements, see figure 1.2.[1,3] The heavier species are scattered into the ISM either by stellar winds or by supernovae.[4]
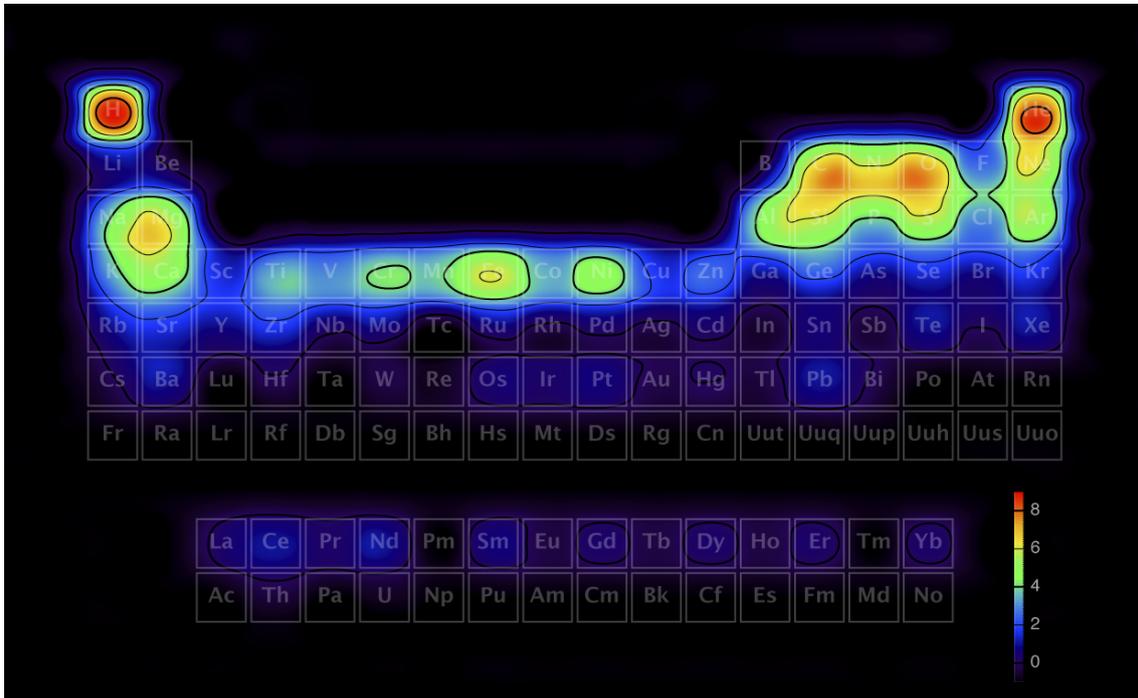
Figure 1.2: A log plot of the total abundance by weight of the elements in the Universe, reproduced with permission from Web Elements (https://www.webelements.com) copyright Prof. Mark Winter

1.1.2 Stars and the elements

Having mentioned that any element heavier than lithium on the periodic table has been scattered into the ISM by stars it seems logical to briefly describe the

---

[¶]https://home.cern/science/engineering/vacuum-empty-interstellar-space

[**]NASA Moon fact sheet https://nssdc.gsfc.nasa.gov/planetary/factsheet/moonfact.html

formation of stars. Hydrogen, helium and lithium were created during the Big Bang, the event approximately 15 billion years ago during which the Universe began. 'Big Bang nucleosynthesis' is the term used to refer to the creation of those elements through the fusion of protons and neutrons into increasingly larger elements as shown in equations 1.1 - 1.3.[1,5]

$$n + p^+ \longrightarrow {}^2H + \gamma \qquad (1.1)$$

$$^2H + n \longrightarrow {}^3H + \gamma \qquad (1.2)$$

$$^3H + p^+ \longrightarrow {}^4He + \gamma \qquad (1.3)$$

The limitations of Big Bang nucleosynthesis means that, initially, there was very little chemical diversity in the Universe. The lack of any element beyond lithium being created during the Big Bang is caused by there being no stable nucleus with five or eight nucleons. In the case of berylium-8 this is because it decays quickly back into two alpha particles.[6] This means that, in the small number of events that do produce a heavier element, it decays back into smaller nuclei.[5] Thus, to create heavier elements another nucleosynthesis process is needed. It was Hoyle *et al.* who revealed that stellar nucleosynthesis would form the way towards elements beyond lithium.[7,8]

A star is formed through gravitational collapse of a molecular or dense cloud.[9] As an aside, this process is quite similar to that of planet formation, a subject that will be discussed in more detail later. The constituent pieces in this case are molecular, or dense, clouds. This subject will be covered in more detail later on. In order for a cloud to collapse it must first exceed the Jeans mass, which is the mass above which the gravitational forces pulling the cloud inwards exceed the pressure pushing it out.[10] It is worth noting that the Jeans mass is temperature dependent, as pressure is related to temperature, and so as the Universe cooled the Jeans mass decreased, hence, larger stars would have formed earlier in the Universe's life.[11] As the cloud collapses and forms a core, the temperature and pressure increase until the pressure balances the gravitational forces and a protostar is formed.[12] Eventually enough additional mass is gained from the material in the envelope around the protostar that collapse begins again, leading to the birth of a new star.

Initially, the energy source of a young star is the fusion of protons to form helium nuclei. When hydrogen has been consumed proton fusion stops, the star contracts further, increasing the density and temperature. At roughly 10^8 K, helium begins to be consumed, which ultimately yields carbon as the product of the triple alpha process.[1,13] The triple alpha process (the fusion of three ^{4}He nuclei)

is both slow and requires high densities of helium, and so would not have generated any significant amount of carbon during the Big Bang.[14] With increasing temperature and size, stars burn carbon (giving O, Ne, Na and Mg) then neon, then oxygen. Finally, silicon is burnt, which produces elements up to iron (to contextualise the size of stars in relation to which elements they burn as fuel, our own Sun burns hydrogen and has a mass of 2×10^{33} g).[1] Any elements beyond iron in the periodic table are produced in supernovae.

1.1.3 Dust clouds

Interstellar clouds have already been mentioned in passing in relation to their part in star formation, but it is worthwhile to devote more attention to them. The colder, denser regions of the interstellar medium are referred to as interstellar clouds. There are two varieties of interstellar cloud: diffuse and dense. Diffuse clouds have low densities, a few tens of atoms per cubic centimetre, and have temperatures around 100 K.[15] Owing to the extremely low density, there is some debate as to how sensible it is to define a temperature for these regions.[1] Dense clouds have roughly 10^6 atoms cm^{-3} and temperatures in the region of 10 K.[1,2] They are up to tens of parsecs (1 pc = 3.08×10^{16} m) across and contain up to 10^5 solar masses of material.[3] It is in these dense clouds that new stars and planets are made. The star-forming regions in dense clouds are often referred to as hot cores.[4,16]

While dense clouds can remain stable for millions of years, eventually gravity will cause the cloud to collapse in on itself which leads to the formation of a new star.[3] It is not only the ability to yield new stellar objects that makes these dense clouds interesting. They are chemically diverse by astronomical standards. Some 170 different molecules have been observed via their rotational, rovibrational and electronic transitions.[4] These species range from hydrides and oxides to a selection of unsaturated radicals and ions.[2] A large number of species that have been observed are shown in Table 1.1. These molecules play their own part in star formation as molecules are required for the heat of the newly forming star to be radiated away.

It should also be noted that the ISM is not comprised of gaseous matter alone. In fact about 1% by mass of its constituents are interstellar dust grains. A schematic of a dust grain is shown in Figure 1.3.[1,18] Dust grains can be carbonaceous or silicaceous in nature. It has been suggested that silicaceous grains could be formed of silicates such as forsterite, $(Fe, Mg)_2[SiO_4]$, or a similar mineral.[19] The composition of the carbonaceous grains is somewhat more com-

---

*The Cologne Database for Molecular Spectroscopy (CDMS), https://cdms.astro.uni-koeln.de/classic/entries/

Table 1.1: A selection of identified interstellar and circumstellar species[4,17] for a complete list please see the CDMS.*

Simple hydrides, oxides, sulphides and halides				
H_2	CO	NH_3	CS	HCl
O_2	H_2O_2	PO	CO_2	$NaCl$
H_2O	SO_2	OCS	H_2S	KCl
PN	SiO	SiH_4	SiS	$AlCl$
N_2O	CH_4	$HSCN$	HF	AlF
$HONC$	$HNCO$	$AlOH$		

Nitriles and acetylene derivatives				
C_2	HCN	CH_3CN	HNC	C_2H_4
C_3	HC_3N	CH_3C_3N	$HNCO$	C_2H_2
C_5	HC_3N	CH_3C_5N	$HNCS$	C_6H_6
C_3O	HC_7N	CH_3C_2H	HNC_3	C_3H_6
C_3S	HC_9N	CH_3C_4H	CH_3NC	C_3H_7CN
C_4Si		CH_3C_6H	HC_2NC	
H_2C_4	HC_2CHO	CH_2CHCN	CH_2CCHCN	

Aldehydes, alcohols, ethers, ketones and amides				
H_2CO	CH_3OH	$HCOOH$	$HCOCN$	CH_3CH_2CN
CH_3CHO	CH_3CH_2OH	$HCOOCH_3$	CH_3NH_2	NH_2CH_2CN
CH_3CH_3CHO	CH_2CCHOH	CH_3COOH	CH_3CONH_2	NH_2CN
NH_2CHO	$(CH_2OH)_2$	$(CH_3)_2O$	H_2CCO	CH_2CHCN
CH_2OHCHO	$(CH_3)_2CO$	H_2CS		
C_2H_5OCHO		CH_3SH		

Cyclic molecules				
C_3H_2	SiC_2	$c-C_3H$	CH_2OCH_2	C_6H_6
$c-SiC_3$	H_2C_3O	C_2H_4O		

Molecular cations				
CH^+	CO^+	$HCNH^+$	OH^+	HN_2^+
CH_3^+	HCO^+	HC_3NH^+	H_2O^+	H_3^+
HS^+	HOC^+	H_2COH^+	H_3O^+	SO^+
HCS^+	$HOCO^+$	CF^+	HCl^+	H_2Cl^+

Molecular anions		
C_4H^-	C_6H^-	C_8H^-
CN^-	C_3N^-	C_5N^-

Radicals				
OH	C_2H	CN	C_2O	C_2S
CH	C_3H	C_3N	NO	NS
CH_2	C_4H	$HCCN$	SO	SiC
NH	C_5H	CH_2CN	HCO	SiN
SH	C_7H	$NaCN$	KCN	$MgCN$
C_3H_2	C_8H	$MgNC$	$FeCN$	
C_4H_2	HNO	H_2CN	HNC_3	HO_2
C_6H_2	$AlNC$	$SiNC$	C_4Si	$SiCN$
HCP	CCP	AlO		

Fullerenes		
C_{60}	C_{70}	C_{60}^+

plex. The dust grains may be made from amorphous carbon, carbynes, diamond, fullerene, graphite or even polycyclic aromatic hydrocarbons (PAHs).[20–22]

Ice mantle

Mantle of organic material

Mantle of refractory organic material

Silicate core

Carbonaceous particles
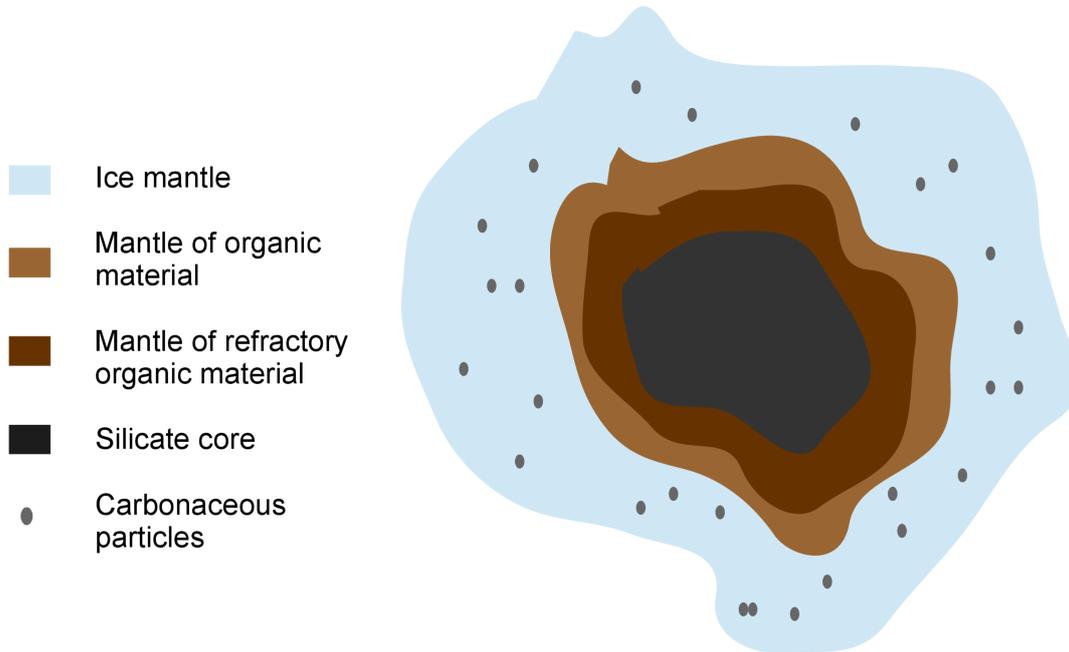
Figure 1.3: The structure of a dust grain in a dense cloud in the ISM[1]

These dust particles are capable of scattering light, which makes them the main source of opacity in the ISM. Gas phase species can also accrete on grain surfaces, which can lead to the formation of ice mantles around the dust grains, much like those found around comets.[2] Water ice is a common ice mantle component but CO, CO_2, CH_4 and NH_3 have also been detected, along with some larger species like methanol and formic acid.[4,23] In recent years the potential for these dust grains to facilitate surface reactions between species found in the ISM has been shown to be increasingly important and is a subject that will be discussed in more detail later.[24,25]

1.1.4   Planets, asteroids and comets

As well as forming part of interstellar clouds, dust grains form the basis of asteroids, meteoroids, and comets. Asteroids and meteoroids, as well as their cousins residing in the outer solar system, comets, are largely composed of minerals and some metals. The most common of the minerals are from the olivine or pyroxene families.[26] They can be dry, in the sense that they lack an outer layer of ice, or ice-bearing. These ices can be similar in composition to those ices found on dust grains.[26]

Planets are largely formed from the conglomeration of other entities found within the interstellar medium which, as they condense energy, begin to heat

internally.[1,27] This heating causes differentiation of the now molten materials, allowing the heavier elements to sink and form a core surrounded by the less dense elements.[1,28] This led to the question of how heavy elements such as gold or platinum have been found in the Earth's surface. It was found that precious metal ions are capable of diffusing from the core to the surface along the boundaries between grains on mantle rock.[29] The chemistry that occurs on the constituent parts of the forming planetesimal, as well as the reactions that happen during the formation, greatly dictate the resulting chemical state of the planet itself. The chemical diversity of the newly-made planet can be enhanced by collisions with meteors and comets. It is thought that such a bombardment occurred on early Earth.[1,30] As Earth amply shows, planetary chemistry can grow to be incredibly complex. There are many questions as yet unanswered as to how the initially simple chemical processes available to newly-created Earth led to the myriad complex chemical and biological reactions taking place now.

1.1.5 Molecules in space

The presence of molecules in the ISM, and some examples of what is present, has already been touched upon during the discussion of interstellar clouds. However, this raises the question of how we know that these species are indeed present. The simple answer is *via* astronomy. Astronomy is closely related to high-resolution spectroscopy. It makes use of rotational, vibrational, or even electronic transitions to identify molecules.[1] Since the first detection of Ca^+ along the line of sight towards the binary star δ-Orionis in 1904, the number of molecules identified has increased greatly.[1,31] The capabilities of the instruments used have also improved markedly. A variety of telescopes are used nowadays to study space. One of the newest is the Atacama Large Millimeter/submillimeter Array (ALMA) which offers the high sensitivity and spectral resolution needed to study young stars.[32] Figure 1.4 shows an example of the imaging possible using telescopes. It shows the Helix nebula imaged in infrared and visible light as well as in the molecular emission of S-band H_2, which is the release of energy from rotationally excited dihydrogen molecules within the nebula.[33] If astronomers are to image space they must select a region of it to study using these telescopes. One frequently used source for studying molecules in the ISM is Sagittarius (Sgr) B2, a molecular cloud complex and one of the most massive star-forming regions in the Galaxy.[16] Many organic molecules have been detected for the first time towards this source.[34]

There is much interest in the organic molecules found in the ISM, particularly those that can be classified as complex organic molecules (COMs). In this con-
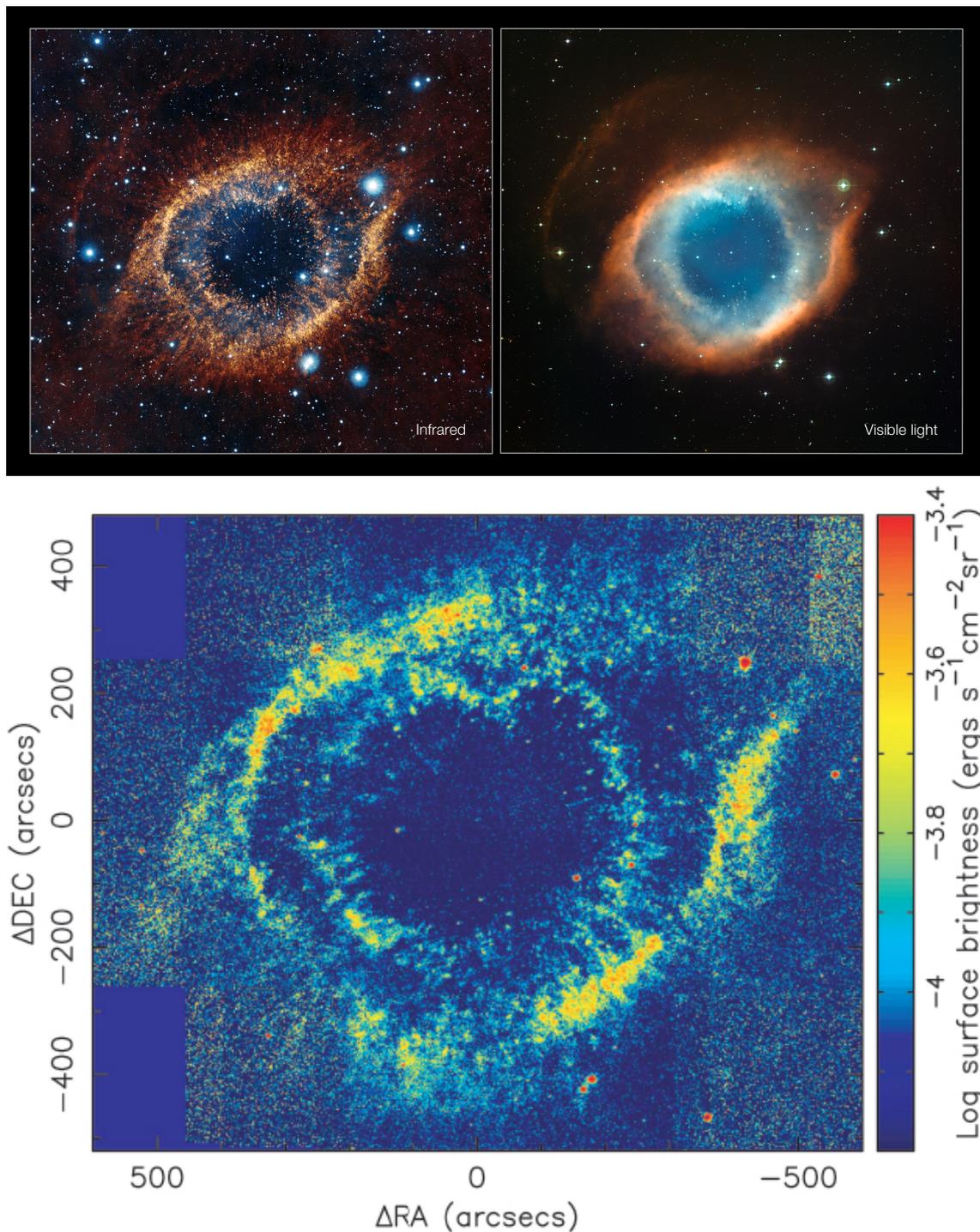
Figure 1.4: Upper: the Helix nebula imaged in IR (using the VISTA telescope) and visible light (using the MPG/ESO 2.2-metre telescope) produced by the European Southern Observatory (ESO) in 2012 (https://www.eso.org/public/images/eso1205d/) and lower: the Helix nebula imaged in the molecular S-branch emission of H_2 using ISOPHOT on the Infrared Space Observatory, reproduced from Speck 2002.[33]

text complex simply means that the molecule contains several atoms and, as one would expect, organic means that it is made from a combination of carbon, hydrogen, nitrogen and oxygen.[35] Some of these COMs are well known on Earth, others less so, such as C_4H^-.[36] The focus of much research in the field is the understanding of the formation of these species, so that we can better understand the universe and the mechanisms and processes that led to the beginning of life, some theories of which are outlined below.

1.1.6 The origins of life

The origins of life discussion as it exists today, namely as a scientific topic rather than a philosophical and religious one, came with the publication of Oparin's theories of a 'primordial broth'.[37] This broth referred to the oceans on early Earth that became enriched in organic molecules as a result of lightning acting on a reducing atmosphere, a reducing atmosphere simply being an atmosphere that acts as a reductant. The work of Urey and Miller in the 1950's, which is discussed in more detail later, gave further weight to these arguments when their experiments were found to produce amino acids. However, geological record data does not support the existence of a reducing atmosphere, and Urey-Miller type experiments using alternative, less severe, atmospheres fail to replicate the diversity of biologically relevant molecules.[38]

An alternative theory to the necessary components for life being synthesised on early Earth is that they were brought to Earth *via* cometary impacts. The relative abundance of organic matter outside of the asteroid belt compared to the amount inside it suggests this may be true.[39] This is coupled with the evidence from the surface of the moon of a period of 'heavy bombardment' of Earth by comets.[40] The discovery of organic molecules on both meteorites (Murchison, Orgueil and Taggish Lake) and comets supports this.[41–43] These discoveries show the need for understanding how these organic species are formed in an ISM setting.

The next issue to address is how biological life started from the building blocks formed in space. Again, there are a variety of theories for this, two of which will be detailed here. There have been suggestions that life started with the formation of an abiotic metabolism. The citric acid cycle has been proposed as one potential starting point, as it can produce amino acids, purines and pyrimidines. Another option is based on fixing carbon from CO_2 by oxidising FeS_2, see figure 1.5 for a simplified version of the cycle.[44,45] There have also been cycles put forward that are based on mineral clays. These clays could have adsorbed the organic molecules found in early Earth's oceans and acted as a heterogeneous catalyst
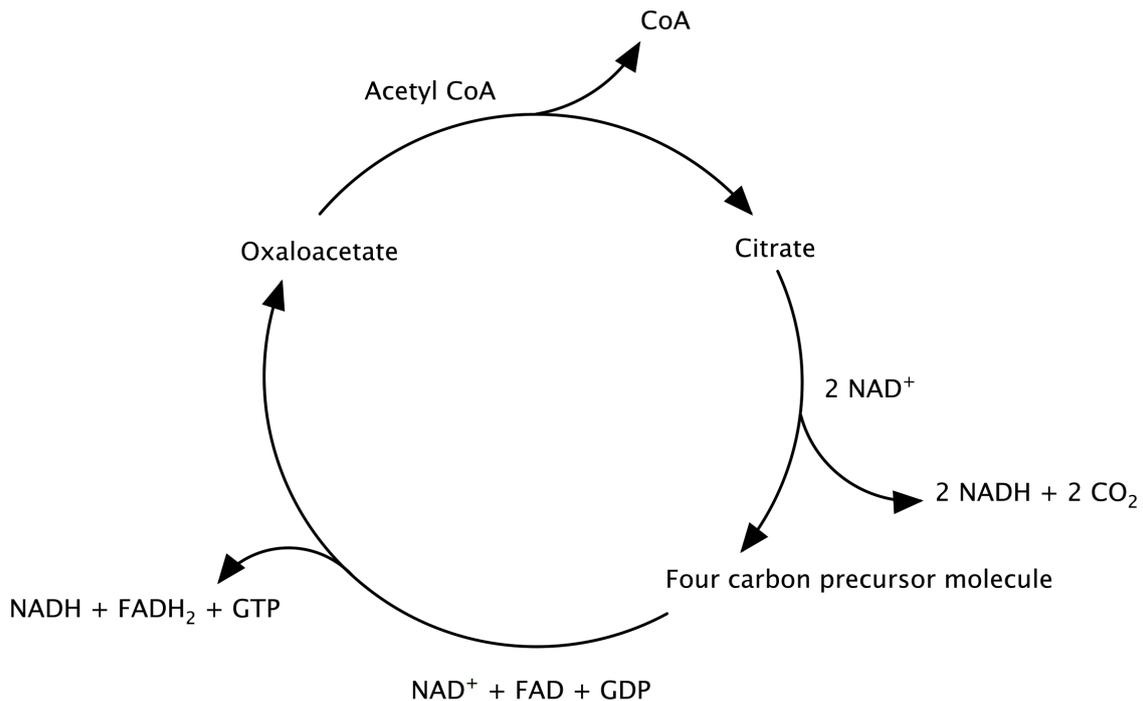
Figure 1.5: A simplified version of the citric acid cycle where citrate is made from the reaction between oxaloacetate and coenzyme A (CoA). The product citrate is then converted back into oxaloacetate through oxidising reactions with NAD or FAD (electron carrier molecules).

for the formation of larger species.[46,47] This would require the organic molecules to be retained on the clay surface and so it has been proposed that the clay surfaces may have formed barriers analogous to modern biology's lipid bilayers.[46] Additionally, the ordered stacking of mica layers suggest how mineral systems could store and pass on genetic information in a manner similar to DNA base pair stacking.[46]

Alternatively to this, in 1962 Rich first introduced the concept of an RNA world, which posits that life began with replication of genetic information coming first.[48] Later, the discovery of the catalytic capabilities of RNA in the form of ribozymes, which are RNA molecules that catalyse the hydrolysis and esterification of RNA substrates, furthered the case for this theory.[49–51] The RNA world hypothesis does rely on the constituent parts of RNA being synthesised either in space or on early Earth. Both urea and ribose are precursors to RNA and the feasibility of their formation under astrochemical conditions is explored in later sections. First, however, we will discuss how the formation of complex molecules in the ISM is modelled by astrochemists and review some examples of molecules that have been detected.

1.2   The formation of complex organic molecules

1.2.1   Astrochemical modelling

Modelling gas-phase reactions

Much of the gas-phase chemistry in the ISM takes place in the star-forming regions of dense molecular clouds. Local heating by the protostar causes species that had previously accreted onto the surfaces of dust grains to sublime back into the gas phase.[4] Then, in these warm regions of gas, simple saturated molecules can react.[36] However, despite the local warming, the temperatures reached are still low (some tens of Kelvin) which means that endothermic reactions and reactions that have energy barriers are still unlikely to proceed.[52,53] Some types of reaction that do not possess energy barriers, and thus can occur under any interstellar conditions, are shown along with examples in Table 1.2.

Table 1.2: The general types, with examples, of reactions which possess no energy barrier and thus are viable in an interstellar context.[52]

Reaction Type	Example
Ion-Molecule	$H_2^+ + H_2 \longrightarrow H_3^+ + H$
Radiative Association	$C^+ + H_2 \longrightarrow CH_2^+ + \gamma$
Dissociative Recombination	$H_3^+ + e^- \longrightarrow H + H_2$
Radical-Radical	$CN + C_2H_2 \longrightarrow H_3CN + H$

Of these available reactions, it is those between ions and neutral molecules that dominate.[54,55] Positive ions in molecular clouds are largely formed by cosmic rays or *via* reactions with species that have been ionised by cosmic rays.[36] However, ionisation can also be caused by secondary electrons (electrons generated as the products of an ionisation process).[56] Cosmic rays can also lead to the production of UV photons within the molecular clouds. This occurs via secondary electrons electronically exciting H_2 which then re-radiates.[57] The ability of cloud species to generate UV radiation internally means that photodissociation and photoionisation are possible processes despite the fact that dense molecular clouds are opaque in the UV.

Now that we have established what kinds of chemical reactions are occurring in the ISM, the question of how to model such reactions arises. Interstellar chemistry is not thermodynamically driven, but governed by kinetics instead, providing the reaction is exothermic. Thus, chemistry in the ISM can be modelled using a network of all possible process, with their associated rate constants, referred to as a chemical reaction network. These networks help to mimic the growing complexity of the molecules as the successive reactions occur. The complexity

arises from the simple compounds made in early steps combining to form larger and larger species. There are several chemical networks which work on this principle.[58–60] In these networks, rate constants are generally of the form shown in equation 1.4.

$$k(T) = \alpha(\frac{T}{300})^{\beta} e^{(-\frac{\gamma}{T})}. \tag{1.4}$$

This is a modified Arrhenius equation for rate constant k as a function of temperature, T, where α is the value of the reaction rate, β characterises the temperature dependence of the rate and γ is the activation barrier in Kelvin.[52,60]

Modelling astrochemical systems in this manner, by using coupled rate equations, has been shown to be both efficient and accurate in many cases.[61] These equations propagate the average concentration of species through non-linear differential equations, which contain both source and sink terms that correspond to available chemical reactions.[61] In these models, the rates of ionisation and dissociation, caused by cosmic rays and UV photons, must be included in addition to the relevant rates of reaction. This is done by equations such as 1.5 and 1.6 respectively.[60]

$$k_{CR} = A_i \zeta_{CR}, \tag{1.5}$$

$$k_{UV} = A e^{-(CA_v)} \chi, \tag{1.6}$$

where ζ_{CR} is the cosmic ray ionisation rate, A_v is the visual extinction and χ is the UV flux. A_i and A are pre-exponential parameters which represent the rate coefficients and $e^{-(CA_v)}$ takes into account the continuum attenuation from the dust grains present in interstellar environments.[60]

Gas-phase chemistry, however, does not always successfully reproduce the observed abundances of detected molecules. For example, the measured abundances of alcohols like CH_3OH and C_2H_5OH in hot cores are higher than those predicted by models in which these species are formed entirely *via* gas-phase chemistry.[62] Equally, many models assume that both $(CH_3)_2O$ and $HCOOCH_3$ are formed in the gas phase after the ejection of methanol from grain surfaces. However, $HCOOCH_3$ formation from this was found to be less efficient than previously assumed, even though $(CH_3)_2O$ can be efficiently produced this way.[62]

The use of gas-phase chemistry alone also fails to reproduce the observed abundance of methyl formate, which can be seen alongside two of its structural isomers in Figure 1.6.[63] These isomers are of interest to astrochemists as they may have relevance in searching for the origins of life. If gas-phase chemistry alone cannot accurately predict the abundances of the molecules observed in the ISM then logically some other process must be occurring to cause the discrepancy.
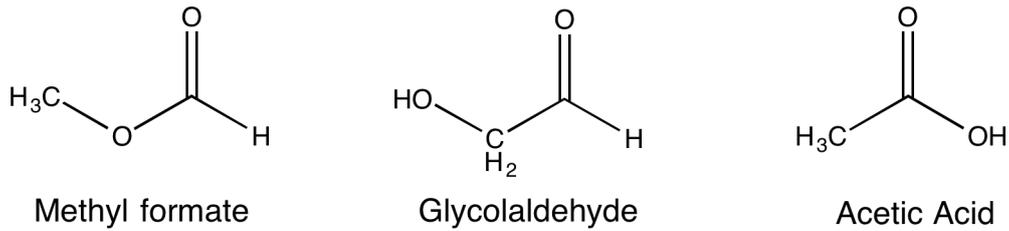
Figure 1.6: The $C_2H_4O_2$ isomers methyl formate, glycolaldehyde and acetic acid.

Modelling grain surface reactions

In recent years it has become increasingly apparent in the field that a large array of complex organic molecules found in the interstellar medium originate on the surfaces of the dust grains. In particular, these dust grains act as catalysts for astrochemical reactions.[64] In addition, the dust grain can act as a heat sink, thus preventing or moderating the rapid dissociation of the product.[65] Therefore, it seems worthwhile to add such chemistries into existing astrochemical models to ascertain whether this leads to predicted abundances in better agreement with those observed. One way to do this is to alter existing astrochemical networks to include terms that take grain-surface chemistry into account. This requires the inclusion of equations that detail the rates of adsorption and desorption from grain surfaces as well as for the rate of reactions taking place on the grain surface.

The rate of adsorption of a molecule onto a dust grain can be modelled using equation 1.7, and the first order rate of desorption via 1.8,[60,66]

$$k_{ads}(i) = s_A \nu_A n_{grain} \pi r_{grain}^2 n_g(A),$$
(1.7)

$$k_{des}(i) = k_{evap,A} n_s(A).$$
(1.8)

Here, $n_g(A)$ is the number density of species A in the gas-phase, r_{grain} is the average radius if a dust grain (on the order of 0.1 μm) with number density n_{grain}, and ν_A is the average gas-phase thermal velocity. For the desorption equation $k_{evap,A}$ is the thermal desorption rate which depends on the binding energy of the species and $n_s(A)$ is the number density of the species A on the surface.[60,66]

Thus, only the rate of reaction on the grain surface is left to consider. In studying this several assumptions are commonly made. One is which surface-reaction mechanism is taking place. This could be the Langmuir-Hinshelwood mechanism[67] which requires both reacting species to be adsorbed on the grain surface, after which they diffuse across the surface until they meet and react (see figure 1.7 for a schematic representation). This is as opposed to the Eley-Rideal or Hot Atom mechanisms, in which only one species is adsorbed on the surface while the other remains in the gas-phase. Commonly, one assumes that

the grains are spherical with a radius of 0.1μm and that the species are attached to the surface via physisorption thus allowing for significant mobility. Using these assumptions the surface rate of reaction, $R(ij)$, between species i and j due to diffusion can be determined using equation 1.9.[60]

$$R(ij) = \kappa(ij)\left\{R_{diff}(i) + R_{diff}(j)\right\}N_iN_jn_d,\qquad(1.9)$$

where N_i and N_j are the mean number of molecules of i and j on the surface, R_{diff} is the rate of diffusion for each species and $\kappa(ij)$ is the probability for a given reaction between i and j to occur. If there is no barrier to the reaction then $\kappa(ij) = 1$, but if there is an energy barrier then equation 1.10 must be used.[60]

$$\kappa(ij) = \alpha\exp\left(-\frac{E_a}{T_g}\right).\qquad(1.10)$$
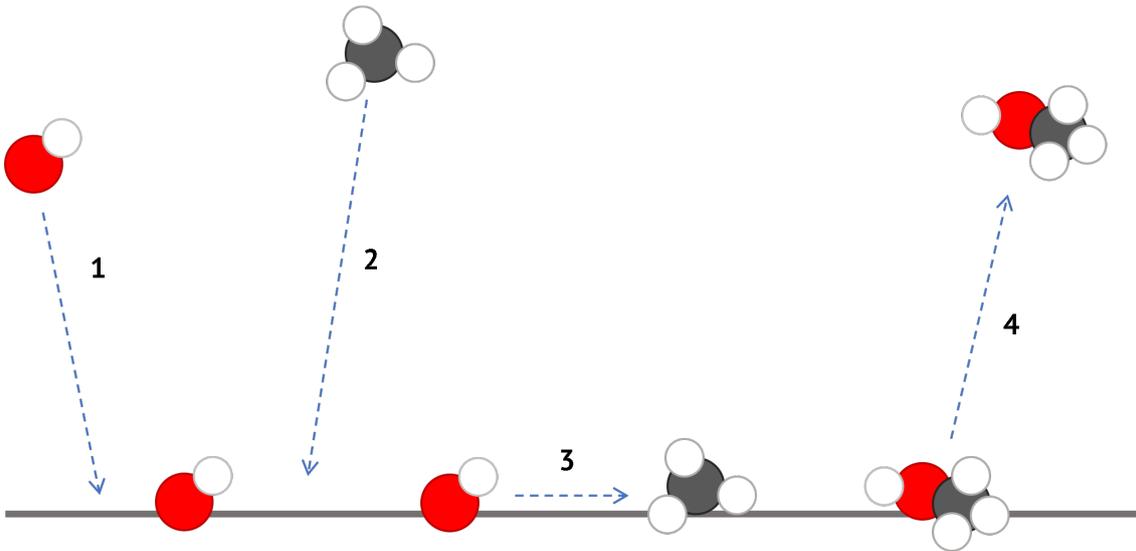
Figure 1.7: A diagram of the Langmuir-Hinshelwood surface reaction mechanism where both species are adsorbed onto the surface before the reaction occurs.

It has been shown that, by adding in grain surface chemistry to an astrochemical network, surface reactions can be an efficient means of making complex organic molecules.[68] It is also apparent, unsurprisingly, that the majority of surface chemistry takes place at higher grain temperatures, when diffusion across the surface becomes more feasible for precursor species.[69] In terms of the abundances of the isomers of methyl formate, acetic acid, and glycolaldehyde, the addition of grain surface chemistry does not make the predicted abundance match those

observed but it does indicate that methyl formate almost certainly has a grain-surface origin.[68,69]

Models for dust grains

If dust grain surfaces are as important to the formation of some complex organic molecules as they appear to be then we need a way to model them efficiently. Polycyclic aromatic hydrocarbons (PAHs) have been suggested as nucleation sites for carbonaceous dust grains and thus are often used as substitutes for dust grains when modelling grain surface processes.[64,65] They have been shown to be good catalysts in some astrochemical reactions, as outlined below.

It has been suggested that PAHs have a role in the formation of H_2 in the ISM.[70] There have also been suggestions that H_2O_2 could be formed on grain surfaces which would ultimately lead to the formation of H_2O via the reactions 1.11 and 1.12.[65]

$$O_2 + H + H \longrightarrow H_2O_2 \tag{1.11}$$

$$H_2O_2 + H \longrightarrow H_2O + OH \tag{1.12}$$

The ionisation energies of polycyclic aromatic hydrocarbons are low and so, due to the large amount of radiation in the ISM, it is possible that they exist as molecular cations in space.[71] Both hydrogen and molecular oxygen can adsorb onto PAH surfaces and so reactions between these species are possible.[65,72] The reaction between O_2 and hydrogen on the surface is barrierless and yields a species that can either dissociate into benzene$^{\bullet+}$ (often used as a small PAH analogue) and $HO_2^{\bullet}$, which is a precursor to hydrogen peroxide, or undergo further reaction with hydrogen to give H_2O_2.

Alternative processes leading to water involve the hydrogenation of O, O_2 and O_3 on surfaces, where hydrogen peroxide is not an intermediate.[73] However, the addition of hydrogen to di- and triatomic oxygen is not barrierless and so the sequential, and barrierless, addition of H to oxygen atoms adsorbed onto the grain surface is more likely.[19] Additionally, some doubt has been cast on the likelihood of solid water remaining on grain surfaces.[74] It has been posited that the exothermic nature of the formation process of water would cause immediate dissociation from the surface and that, in diffuse clouds, the grain surface would remain too warm for efficient water-ice build up.[74] However, in colder dense clouds simulations and experimental studies have found that water ice can and does form on carbonaceous surfaces.[75–77]

Another example of PAHs acting as catalysts for astrochemical reactions is the formation of sulphur-containing species. Sulphur species are often used as an evolutionary tracer in star-forming regions.[78] However, the form of the sulphur-

containing species is not well understood. The large abundance of hydrogen makes it sensible to suggest that hydrogenated sulphur, such as H_2S, should be present but it is not observed.[64] What has been detected is OCS.[79] Some potential OCS formation reactions are given in Table 1.3.[64]

Table 1.3: Potential formation reactions for OCS in star-forming regions.[64]

Reaction Pathway
$CO + S \longrightarrow OCS$
$CO + HS \longrightarrow OCS + H$
$CS + O \longrightarrow OCS$
$CS + OH \longrightarrow OCS + H$

The energy barriers involved in these reactions are noticeably reduced when they are modelled as taking place on the grain surface rather than in the gas phase. In the case of the last two reactions listed they become barrierless on the surface and thus are eminently feasible in the ISM.[64] Adriaens *et al.* suggest that the catalytic effect of coronene, the PAH they used to mimic the grain surface, lies in the role of this molecule in the polarisation and weakening of the CO bond via induced-dipole-induced-dipole interactions.[64]

In light of the literature reviewed above, it seems that polycyclic aromatic hydrocarbons may be suitable models for carbonaceous dust grains. The use of PAHs as grain models could help to elucidate the nature of reactions taking place on dust grain surfaces as well as showing how the grains themselves catalyse these reactions.

Diffusion and desorption processes

The way in which grain-surface chemistry can be included in chemical network models of astrochemical reactions was described earlier. However, for chemical species on grain surfaces to be able to react with one another they need to be able to diffuse across the grain surface. Thus, it seems logical to attempt to understand this process as well as possible. A key assumption was that grain-surface reactions take place *via* the Langmuir-Hinshelwood mechanism, which has been previously discussed during the coverage of grain surface models. A key part of this mechanism is that it requires the reacting species to be able to diffuse across the surface. Higher grain-surface temperatures aid the diffusion of 'heavy' surface radicals. It has been suggested that the stochastic heating caused by cosmic rays could aid and, indeed, speed up diffusion processes on the grain surface.[60] It has also been suggested that it is possible that quantum tunnelling could have an effect of the rate of diffusion.[60]

The inclusion of these tunnelling processes into chemical networks does have an effect on the predicted abundances. By allowing small, light atoms and molecules such as H, H_2 and O to tunnel through energy barriers, the formation of H- and O-bearing species is enhanced.[60] The enhancement stems from the fact that with quantum tunnelling species can move across the surface at much lower grain temperatures, which then leads to greater abundances of related product species.[60] That oxygen atom tunnelling is sufficient to give rise to a noticeable increase in O-bearing species is perhaps surprising for oxygen given its mass relative to hydrogen. There is clearly more chemical complexity to be found in cold regions than one would initially expect.

Cosmic-ray-induced heating also increases the abundance of complex species on the grain surface by increasing the mobility of precursor molecules. However, this effect is only efficient when the visual extinction (the absorption and scattering of radiation by dust particles) is low.[60] Moreover, even though this causes an increased abundance of the more complex species, this is accompanied by a decrease in the abundances of simpler molecules. This is a result of the photodissociation of surface molecules, which produces more radicals, which can react to generate the more complex species.[60]

It has been suggested that the barriers to surface migration are too large for anything other than hydrogen atoms and so diffusion-limited processes would be unlikely to generate many COMs.[80] However, the ice mantles of interstellar dust grains are subject to irradiation from UV photons which can have an effect on the ice. In particular, the energy of these photons is high enough to ionise the molecules that comprise the ice mantle.[81] When these molecules are ionised they break down into more reactive species, which subsequently react with neighbouring molecules or radicals.

This mechanism is evidenced by the fact that, when a binary mixture of H_3COH and CO ices was irradiated with $5\,\text{keV}$ electrons, both methyl formate and glycolaldehyde were observed as products.[81] The predicted formation routes of both isomers are summarised in figure 1.8.

Given that methanol and carbon monoxide have been observed as grain ice constituents, this could be a promising formation route to some of the $C_2H_4O_2$ isomers.[82] The abundances agree qualitatively with those found via observation, in that methyl formate is more abundant than glycolaldehyde, but acetic acid was observed only when a methane and carbon dioxide ice mixture was used.[83]

It is also important to understand the desorption process of any COMs formed. Many astrochemical models treat desorption as an instantaneous process, even though this unlikely to be the case.[84] A better understanding of how and when species desorb from the grain surface may help shed light on the relative abun-
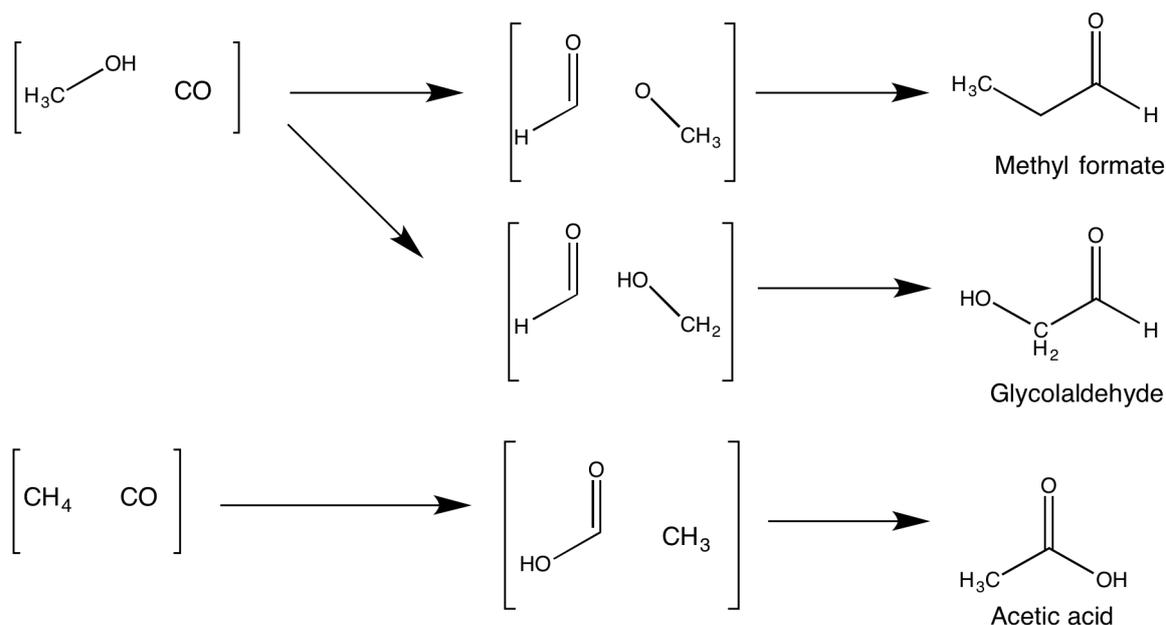
Figure 1.8: The formation routes for the $C_2H_4O_2$ isomers when a binary ice mix of methanol and carbon monoxide was subjected to irradiation.[81]

dances of the COMs that have been observed. Water-ice is capable of trapping CO and retaining it within the ice such that it desorbs at higher temperatures than expected given the strength of the interaction between it and the ice.[84]

Collings *et al.* posited that the desorption behaviour of species within a water ice mantle is largely controlled by the behaviour of the surrounding water.[84] Species that are similar to water, i.e species that can hydrogen bond with the water, co-desorb with water whereas CO-like species display what is known as a 'molecular volcano' as well as co-desorption. A molecular volcano is a desorption process during which trapped molecules are released much like a volcano erupting.[84] There are also intermediate species, those whose behaviour is between water-like and CO-like, which can be modelled using suitable scaled kinetics for the release of CO-like species.[84]

Further studies have looked at the desorption behaviour of more complex astrochemically relevant molecules, in particular the $C_2O_2H_4$ isomers in order to ascertain whether a difference in desorption process could account for the differences in observations.[85] As with the simpler species the nature of the interactions with the water ice determine the desorption behaviour. Burke *et al.* found that the COMs could be put into categories very similar to those used by Collings *et al.* giving a further two: complex intermediate, and complex water-like.[86] The categories and the molecules that fall into them are summarised in Table 1.4. The difference in desorption behaviour between the $C_2O_2H_4$ isomers may play a part

Table 1.4: The desorption behaviour of various astrochemically relevant species as categorised in Collings 2004 and Burke 2015.[84,86]

Desorption category	Species
Water-like	NH_3
	CH_3OH
	$HCOOH$
CO-like	N_2
	O_2
	CH_4
	NO
Intermediate	C_2D_2
	C_2H_2
	CO_2
	OCS
	CS_2
	SO_2
	CH_3CN
	H_2S
Complex Intermediate	$HCOOCH_3$
	H_3COCH_3
	$OC(CH_3)_2$
Complex Water-like	H_3CCOOH
	H_3CCH_2OH
	$H_3CCH(OH)CH_3$
	$HCOCH_2OH$

in why methyl formate has been observed in a wider variety of objects than acetic acid and glycolaldehyde which could have remained trapped in the ice.[85]

1.2.2 A brief survey of some examples

Now that we have covered some of the processes behind identifying chemical species in the ISM as well as understanding their formation it seems prudent to cover some examples. These examples show how detections, computational modelling and laboratory studies work in tandem to elucidate how the chemical diversity of the ISM has arisen.

Formic acid ($HCOOH$) represents a possible precursor to both acetic acid and glycine and has been detected towards a wide variety of sources in the ISM.[62,87–90] This variety of detection sources includes in cometary comae (the envelopes of sublimed material formed around comets when they pass close to the Sun) and as a solid on interstellar grains.[82,91] This latter detection strongly suggests that there is a grain surface formation route to formic acid. Both CO and H_2O ice are common in the ISM and the photolysis of such ices has proved to

yield formic acid amongst other products.[92] Additionally, irradiation of these ices with high energy electrons leads to the formation of HCO and OH, which can the react to give HCOOH.[93] An alternative route using OH and CO has been proposed that goes via the intermediate species HOCO, which can then react with a hydrogen atom to give formic acid.[94]

Model ice studies disagree as to the importance of this alternative route. Bennett *et al.* found it to only be a minor route, with the HCO and OH reaction pair playing a more key part in formic acid generation.[95] Ioppolo *et al.*, however, suggest that HOCO could have a more key role than previously outlined and could explain the presence of formic acid in dense molecular clouds and at the beginning of the warm up phase of protostars in star forming regions.[96] Some gas-phase routes have been proposed for the formation of formic acid from either HCO^+ and H_2O or CH_4 and O_2^+.[97,98] Both of these routes first lead to $HCOOH^+$ which then dissociatively recombines to give formic acid. However, this dissociative recombination was found to be less efficient than thought, making the grain surface routes the much more likely option.[99]

Both benzene and larger PAHs have been detected in the ISM as well as some benzene derivatives (benzonitrile and phenol).[100–104] Despite this, the formation route for benzene is not well understood, to the extent that including the currently proposed formation routes for benzene and benzonitrile in the kinetic database KIDA leads to column densities that do not match those observed. This implies that there must be formation routes as yet unaccounted for.[103] It has been suggested that benzene could be formed from the irradiation of acetylene ices by electrons.[105] Computational investigations found a plausible reaction route in which three acetylene units join together sequentially and then cyclise to form the product benzene ring. However, these experiments were designed to probe benzene formation on Saturn's moon Titan and so may not be applicable in an ISM setting. Nonetheless acetylene has been detected in the ISM and so this may still represent a viable formation route.[106] Propargyl alcohol (HC_2CH_2OH) has also been put forward as a possible benzene precursor. This is due to its ability to dissociate into OH and C_3H_3, the latter of which can dimerise into benzene.[107] Indeed, irradiating a propargyl alcohol ice does lead to benzene as a major product.[108] There is, as yet, no confirmed detection of propargyl alcohol, but its isomer propenal has been detected and work has been done to suggest where propargyl alcohol may exist in the ISM.[109,110]

Of particular interest in an astrobiological sense are ways in which amino acids could be formed. Research on this topic has been conducted for decades. One, now famous, early example is the Urey-Miller experiment from the 1950's.[111] In this experiment, a proposed reducing atmosphere of H_2O, CH_4, NH_3 and H_2,

modelled on the atmosphere of Jupiter, was sparked with electrodes over water to mimic lightning in an early Earth atmosphere (shown in figure 1.9). The
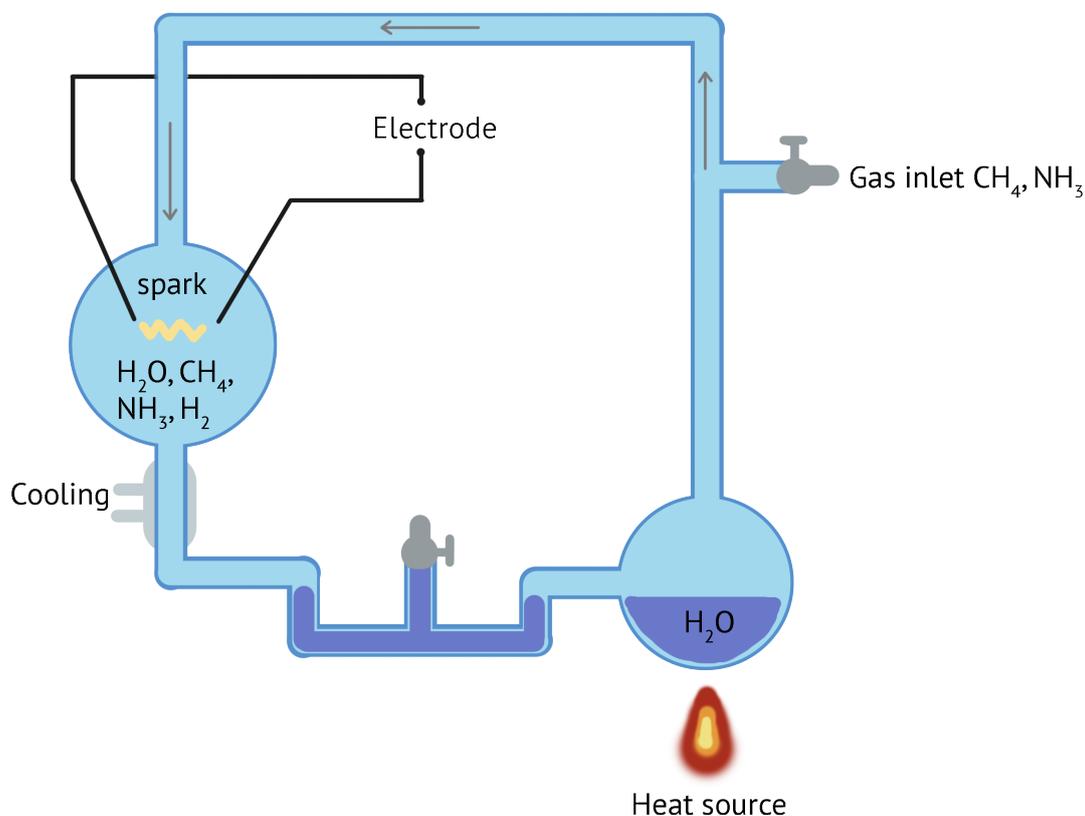
Figure 1.9: A diagram of the experimental setup used by Miller and Urey to investigate the formation of biologically relevant organic molecules on early Earth.[111]

condensed vapours were then analysed and found to contain a variety of biologically relevant molecules including amino acids. The Urey-Miller experiment is not uncontroversial. Firstly, geological records do not support the presence of a strongly reducing atmosphere on early Earth. Secondly, the results were racemic whereas biological systems are well known to contain only single entantiomers.[112] Nonetheless, the biological importance of amino acids means that a plausible abiotic synthesis needs to be found.

Two precursors to amino acids that could be formed, abiotically, on dust grain surfaces are acetamide and acetonitrile. Acetamide formation also begins with the complexation of O_2 to the PAH cation surface. Subsequent reactions with CH_3CN and hydrogen lead to $CH_3C(OH)NH$, which can undergo exothermic hydrogen transfer to yield acetamide.[65] Acetonitrile may be produced by a reaction between CH_2CN^+ and ammonia but, as yet, the formation mechanism is poorly understood.[65] It has been found that the suggested formation route is both bar-

rierless and exothermic. However, it is so exothermic that, were the molecule to dissipate the reaction enthalpy via vibrational relaxation, it would dissociate again immediately. Thus, if acetonitrile is produced this way, the process must use the grain surface as a heat sink.[65]

Another astrobiologically important COM that could be formed on the surface of dust grains is urea, H_2NCONH_2, which has been detected in the Murchison meteorite.[113] In this instance it is the ice mantle of the grain, as opposed to the grain surface itself, that is providing the means to form the product.[114] Irradiation of an ammonia/carbon monoxide ice mix yielded OCN^-, H_2CO, formamide, NH_4^+, urea and formylhydrazine. It was thought that urea would form via reactions 1.13 and 1.14,

$$NH_4^+ + OCN^- \longrightarrow NH_3 + HNCO \tag{1.13}$$

$$NH_3 + HNCO \longrightarrow CO(NH_2)_2 \tag{1.14}$$
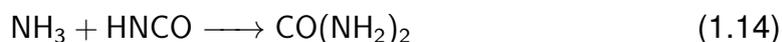
However, Förstel *et al.* conducted an experimental study which indicates that this formation route is unlikely.[114] They suggest instead that urea could be formed via the insertion of NH into the $C-H$ bond of formamide as shown in Figure 1.10 though further investigation is required to ascertain the feasibility of this reaction (see Chapter 4).
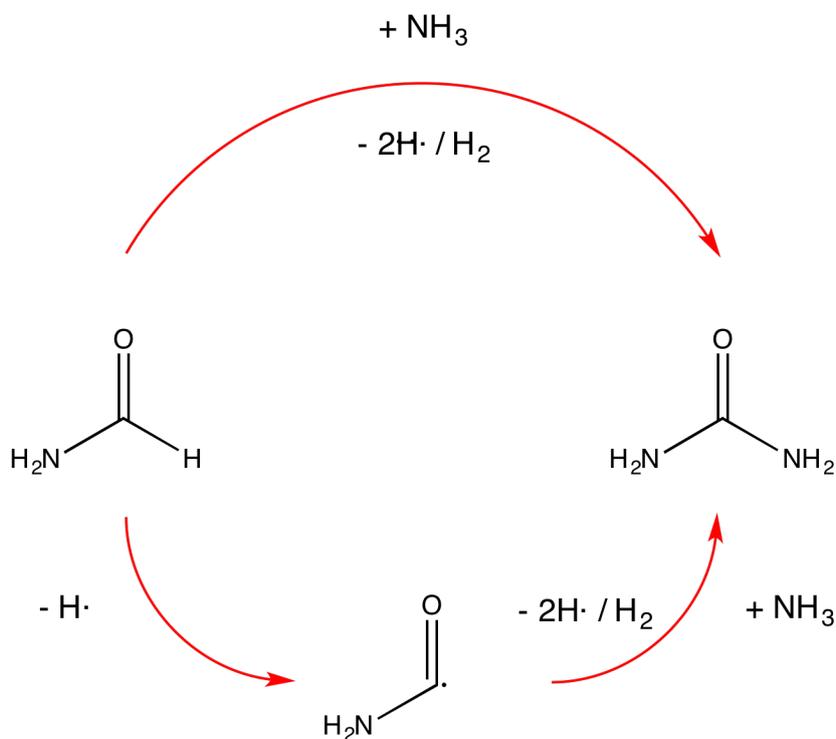
Figure 1.10: The synthetic routes to urea from formamide via irradiation by cosmic rays. The lower route is that suggested by Förstel et al. in which NH inserts into the $C-H$ bond of formamide.[114]

Urea is not the only COM that has been found on the surfaces of comets or meteorites; amino acids and nucleobases have also been found.[113,115] Recently the simplest amino acid glycine has been found on the comet Wild 2 and in the coma of comet 67P/Churyumov-Gerasimenko.[116] Until this detection, glycine had remained elusive despite the fact that it had been shown in laboratory studies that irradiation of ices containing water, methanol, carbon mono- and dioxide and ammonia or hydrogen cyanide can produce glycine amongst other amino acids.[117,118]

Various formation routes have been suggested, both on grain surfaces and in the gas-phase. These involve the addition of either the NH_2 radical or NH_3OH^+ to acetic acid, or the reaction between the radicals NH_2CH_2 and $HOCO$.[119–121] Adding in the surface reactions for the formation of glycine leads to a significant abundance thereof. The actual peak abundance, as well as the temperature dependence of the main formation routes, varies depending on the warming timescale of the model.[122] The gas-phase reactions that can lead to glycine never achieve significance in comparison with the grain-surface reactions and thus seem unlikely to contribute greatly to interstellar glycine.[122]

1.2.3 Aims

Advances in astrochemical modelling, both theoretical and experimental, are bringing us closer to understanding the observations. The importance of grain surface reactions has become more apparent and they are beginning to offer explanations as to why seemingly very similar molecules have different abundances.[68,69,123] We are building a better picture of the reactions, and of the diffusion and desorption processes.[60,84–86]

However, there is still much to be done. As yet, various molecules important for understanding the origins of life do not have clear astrochemical synthesis routes.[114,122] For example, it has been suggested that biologically relevant sugars could be produced via the formose reaction, which is the autocatalytic self-condensation of formaldehyde.[124,125] However, the vast range of products produced by the formose reaction means the yield of the most astrobiologically desirable sugars is low, and there is still the question of synthesising the amino acids and nucleoside bases. There is more work to be done in the field, even as our understanding of it grows and deepens.

Herein, a variety of ISM chemistry is considered. Firstly, the ionisation behaviour of PAHs of increasing size in water ice is investigated (Chapter 3) and the UV-vis spectrum of benzoic acid and benzene diacid mixtures are predicted (Chapter 4). In Chapter 5 the production of methyl formate and its isomer glycolaldehyde from the radical species HCO and $HOCH_2$ or OCH_3 is probed using

multiconfigurational methods. Various formation routes to urea based on those proposed by Förstel *et al.* are explored using Density Functional Theory (DFT) and multiconfigurational methods, see Chapter 6. Finally, the ability of the formose reaction to generate ribose under astrochemical conditions is tested by calculating transition states and barrier heights using DFT in Chapter 7.

2.   Theory

2.1   Quantum mechanical background

Quantum mechanics attempts to explain phenomena that classical physics cannot, and quantum chemistry does the same for chemical systems specifically.[126] These phenomena include the structures and properties of molecules, nuclear motion, and chemical reactivities. This allows chemists to apply quantum chemical theories and methods to explore chemical systems and their properties. Quantum chemistry can inform us on stable geometries and conformers; transition states; spectral properties and reaction mechanisms. It is of particular use in studying systems or properties that are not easily dealt with in a laboratory setting. As much of astrochemistry occurs at such low temperatures and densities not easily achievable in a laboratory, computational methods can shed a great deal of light on the reactions taking place.

At the heart of quantum mechanics lies the time-independent Schrödinger equation (equation 2.1).[127–130]

$$\hat{H}\Psi = E\Psi. \tag{2.1}$$

Here, Ψ is the wavefunction, E is the energy, and $\hat{H}$ is the Hamiltonian operator. The latter is composed of the operators for kinetic and potential energy, $\hat{H} = \hat{T} + \hat{V}$. The kinetic energy operator takes the form;

$$\hat{T} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\nabla^2, \tag{2.2}$$

where, $\hbar$ is the reduced Planck's constant and $\nabla^2 = (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})$ and is the Laplacian. The potential energy operator, $\hat{V}$, describes the Coulombic potential between the charged sub-atomic particles. These interactions can be separated into the nuclear-nuclear, nuclear-electron and electron-electron terms. Additionally, the Hamiltonian can be simplified by making use of atomic units where $e = 1, m_e = 1, \hbar = 1, 4\pi\epsilon_o = 1$ leading to equation 2.3.

$$\hat{H} = -\frac{1}{2}\sum_A \frac{1}{M_A}\hat{\nabla}_R^{\,2} - \frac{1}{2}\sum_a \hat{\nabla}_r^{\,2} + [\sum_{A>B}\frac{Z_A Z_B}{R_{AB}} - \sum_{a,a}\frac{Z_A}{r_{Aa}} - \sum_{a>b}\frac{1}{r_{ab}}]. \tag{2.3}$$

In this, M_A, Z_A and R_{AB} are the mass and charge of nucleus A and then the distance between nuclei A and B. The remaining terms, r_{Aa} and r_{ab}, are the distances between nucleus A and electron a, and electron a and electron b respectively. The first two terms are the nuclear and electronic kinetic energies and the latter three are the nuclear-nuclear, nuclear-electronic and electronic-electronic potential energy terms.

Frequently, we are concerned with the probability amplitude of the wavefunction, $|\Psi^2|\, d\tau$, which indicates the likelihood of finding a quantum-particle in a given volume element which includes all space spin coordinates. Also, in general we deal with the time-independent, non-relativistic Schrödinger equation.[131]

However, for any system larger than a hydrogenic system, it becomes essentially impossible to solve the Schrödinger equation exactly, and we have to make some approximations. One of the approximations used to simplify solving the Schrödinger equation is the orbital approximation which states that each electron in a many-electron molecule occupies its own one-electron function or orbital. Another is the Born-Oppenheimer approximation, also known as the adiabatic approximation. The Born-Oppenheimer approximation states that, as nuclei are vastly heavier than electrons, the electrons move significantly faster than the nuclei.[132] As such, electrons can instantaneously adjust to changes in nuclear position, and so the nuclei appear stationary to the electrons. This allows one to simplify the Schrödinger equation by separating nuclear and electronic motion. This means we can now write our wavefunction as in equation 2.4.

$$\psi(\mathbf{r}, \mathbf{R}) = \Phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}). \tag{2.4}$$

Here, Φ is the electronic wavefunction, χ is the nuclear wavefunction, $\mathbf{R}$ is a set of the nuclear coordinates, and $\mathbf{r}$ is a set of electronic coordinates. If we then apply the Born-Oppenheimer approximation to $\hat{H}\Psi = E\Psi$ given that $\psi(\mathbf{r}, \mathbf{R}) = \Phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R})$ we end up with equation 2.5.

$$\hat{H}[\phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R})] = E[\phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R})]. \tag{2.5}$$

Expanding this out and then, as $M_A >> m_a$, recognising that terms inversely proportional to the mass of the nuclei should be very small removing these non-adiabatic contributions to the kinetic energy the following equation results:

$$\hat{H}[\phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R})] = -\frac{\hbar^2}{2}\sum_a \phi_a(\mathbf{r})\sum_A \frac{1}{M_A}\hat{\nabla}_R^2 \chi(\mathbf{R})$$
$$-\frac{\hbar^2}{2m_e}\sum_a \hat{\nabla}_r^2 [\phi(\mathbf{r})] + \hat{V}[\phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R})] \tag{2.6}$$
$$= E^{BO}[\phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R})],$$

in which E^{BO} is the total energy within the Born-Oppenheimer approximation and differs from the exact, non-relativistic energy by a non-adiabatic correction.

If we then apply the clamped-nucleus approximation, in which we fix the nuclear framework such that the nuclei are motionless, we can calculate the elec-

tronic energies for fixed nuclear positions, via equation 2.7.[131]

$$\hat{H}_{el}\Phi(\mathbf{r}) = E_{el}(\mathbf{R})\Phi(\mathbf{r};\mathbf{R}). \tag{2.7}$$

In this, $\hat{H}_{el}$ is the electronic Hamiltonian which is written in full (in atomic units) in equation 2.8, where Z_A is the charge on nucleus A, r_{Aa} is the distance between nucleus A and electron a, r_{ab} is the distance between electron a and electron b, and R_{AB} is the distance between nuclei A and B.

$$\hat{H}_{el} = -\frac{1}{2}\sum_a \nabla_a^2 - \sum_{A,a}\frac{Z_A}{r_{Aa}} + \sum_{a>b}\frac{1}{r_{ab}} + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}}. \tag{2.8}$$

The result of the clamped nucleus approximation is that the last sum in the electronic Hamiltonian becomes a constant given that the inter-nuclei distance does not change. This then leads to the idea of a potential energy surface where the electronic Schrödinger equation for various sets of nuclear coordinates is solved and the geometry with the lowest potential energy can be found.

The above approximations allow us to simplify the process of solving the Schrödinger equation somewhat. As previously stated, it is not practically possible to achieve exact solutions for systems with multiple electrons, but we can find approximate solutions by building upon what we know about one-electron systems. It is then possible to create the many-electron wavefunction by combining together several one-electron wavefunctions. We do so by forming the anti-symmetrised product (a so-called Slater determinant, see section 2.2) of several one-electron wavefunctions, $|\tilde{\Phi}\rangle$, given here in Dirac's bra-ket notation. Bra-ket notation is a convenient shorthand in which the ket, $|a\rangle$, represents the wavefunction and the bra, $\langle a'|$, is its complex conjugate. If the bra and ket are written as follows, $\langle a'||a\rangle$ then integration over $d\tau$ is implied (often the two central lines are merged into a single line).

Having found an approximate solution, we can use the *Variation Principle*, equation 2.9, which states that the energy obtained is an upper bound for the true energy of the system. By minimising this expression, it is possible to find the most accurate wavefunction within our assumptions,

$$\tilde{E} = \frac{\langle\tilde{\Phi}|\hat{H}|\tilde{\Phi}\rangle}{\langle\tilde{\Phi}|\tilde{\Phi}\rangle} \geq E_0, \tag{2.9}$$

hence, allowing the properties of the system to be accurately predicted.

2.2 The Slater-Condon Rules

The Hamiltonian under the Born-Oppenheimer approximation is composed of kinetic and potential energy terms split into nuclear and electronic contributions.

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{nm} - \hat{V}_{ne} + \hat{V}_{ee}, \tag{2.10}$$

where $\hat{T}_n$ is the purely nuclear kinetic operator, $\hat{T}_e$ is the purely electronic kinetic operator, $\hat{V}_{nm}$ is the nuclear-nuclear potential operator, $\hat{V}_{ne}$ is the nuclear-electronic potential operator and finally, $\hat{V}_{ee}$ is the electronic-electronic potential operator. If the electrons are treated as independent particles, it is possible to write an *N*-particle wavefunction as the product of *N* independent, orthonormal, single electron functions as in equation 2.11.

$$\phi(r_1, r_2, ..., r_N) = \varphi_1(r_1)\varphi_2(r_2)...\varphi_N(r_N) = \prod_i^N \varphi_i(r_i). \tag{2.11}$$

This is the Hartree product form of the electronic wavefunction. However, electrons possess spin as well as space coordinates. Thus, this must be included as well. This is done by rewriting the Hartree product in terms of the space-spin wavefunction for *N* particles $\Theta(\tau) = \Phi(r)\Xi(\sigma)$ and the single particle space-spin wavefunction $\theta_i(\tau_i) = \phi_i(r_i)\zeta_i(\sigma_i)$ to get

$$\Theta_i(\tau_i, \tau_2, ..., \tau_N) = \theta_1(\tau_1)\theta_2(\tau_2)...\theta_n(\tau_N) = \prod_i^N \theta_i(\tau_i). \tag{2.12}$$

However, single electrons are indistinguishable fermions which means that they must obey the Pauli antisymmetry principle. This means that there must be an accompanying change of sign upon the swapping of two electrons.[126,133] This principle can be obeyed by writing the electron wavefunction as a normalised determinant, for example this two electron case:

$$\Theta(\tau_1, \tau_2) = \frac{1}{\sqrt{2!}} \begin{bmatrix} \theta_1\tau_1 & \theta_2\tau_1 \\ \theta_1\tau_2 & \theta_2\tau_2 \end{bmatrix} = \frac{1}{\sqrt{2!}}[\theta_1\tau_1\theta_2\tau_2 - \theta_1\tau_2\theta_2\tau_1]. \tag{2.13}$$

The above is referred to as a Slater determinant and is inherently antisymmetric. This, simple, two-electron case can be extended to an N-electron one with the general normalisation constant $\frac{1}{\sqrt{N!}}$. The same result can be achieved by using

the antisymmetriser operator, $\hat{A}$, which takes the form in equation 2.14.[134]

$$\hat{A} = \frac{1}{N!} \sum_{\hat{\pi} \in S_n} \epsilon_\pi \hat{\pi}.$$ (2.14)

The antisymmetriser includes the permutation operator $\hat{\pi}$ which swaps the coordinates of a pair of particles, and ϵ_π which is the sign change associated with the permutation. The wavefunction equation that results from the use of the antisymmetriser is equivalent to the Slater determinant wavefunction.

Using an antisymmetric wavefunction within the Born-Oppenheimer approximation has useful consequences as can be seen when applied to the Hamiltonian. If the terms are collected based on the number of electrons they involve the following equation is obtained.

$$\hat{H} = -\frac{1}{2} \sum_a \hat{\nabla}^2 - \sum_{A,a} \frac{Z_a}{r_{Aa}} + \sum_{a>b} \frac{1}{r_{ab}} + \hat{T}_n + \hat{V}_{nn}$$
$$= \sum_a \hat{h}_a + \sum_{a>b} \hat{v}_{ab} + \hat{T}_n + \hat{V}_{nn}.$$ (2.15)

The first term, $\hat{h}_a$, is the one electron operator, while the second term, $\hat{v}_{ab}$, is the two electron operator. This means that the electronic energy can be expressed as in equation 2.16.

$$\tilde{E} = \sum_a \langle \tilde{\Theta} | \hat{h}_a | \tilde{\Theta} \rangle + \sum_{ab} \langle \tilde{\Theta} | \hat{v_{ab}} | \tilde{\Theta} \rangle.$$ (2.16)

If Θ is written out in terms of the spin-orbitals, then the one-electron term becomes $\langle \theta_1 ... \theta_n \hat{A} | \hat{h}_a | \hat{A} \theta_1 ... \theta_n \rangle$. Expanding out Θ by using the antisymmetriser operator and rearranging due to the idempotence of $\hat{A}$, means that this can be simplified into equation 2.17.

$$\langle \theta_1 ... \psi_n | \hat{h}_a \hat{A} | \theta_1 ... \theta_n \rangle = \sum_{\hat{\pi}} \langle \theta_1 ... \theta_n | \hat{h}_a \hat{\pi} | \theta_1 ... \theta_n \rangle.$$ (2.17)

Assuming the one-electron orbitals are orthonormal, the only non-zero term in this is for the identity $\hat{\pi}$ operator as, if two of the electrons are swapped (a and b for example), $\hat{h}_a$ only acts on one electron yielding equation 2.18.

$$\langle \theta_1 | \theta_1 \rangle ... \langle \theta_a | \hat{h}_a | \theta_b \rangle \langle \theta_b | \theta_a \rangle ... \langle \theta_n | \theta_n \rangle,$$ (2.18)

where $\langle \theta_a | \theta_b \rangle = 0$ as a does not equal b. This is further simplified as, under the identity operator, normalisation means that all $\langle \theta_c | \theta_c \rangle = 1$ and thus, the matrix element simply becomes $\langle \theta_1 | \hat{b}_i | \theta_a \rangle$, which mean that the one-electron parts reduce

down to equation 2.19.

$$h_a = \langle \theta_a | - \frac{1}{2} \nabla^2 - \sum_A \frac{Z_A}{r_A} | \theta_a \rangle. \tag{2.19}$$

An analogous derivation can be performed for the two-electron parts, however, in this case both the identity operator and the π-operator that swaps a and b (as $\hat{v}_{ab}$ operates on both of them) yield non-zero terms leading to equation 2.20.

$$\langle \theta_1 | \theta_1 \rangle ... \langle \theta_a \theta_b | \hat{v}_{ab} | \theta_a \theta_b \rangle ... \langle \theta_N | \theta_N \rangle - $$
$$\langle \theta_1 | \theta_1 \rangle ... \langle \theta_a \theta_b | \hat{v}_{ab} | \theta_b \theta_a \rangle ... \langle \theta_N | \theta_N \rangle \tag{2.20}$$

Two of these terms, $\langle \theta_a \theta_b | \hat{v}_{ab} | \theta_a \theta_b \rangle$ and $\langle \theta_a \theta_b | \hat{v}_{ab} | \theta_b \theta_a \rangle$, are the Coulomb (identity) and Exchange (swap) operators and can also be expressed as shown in equations 2.21 and 2.22, respectively

$$J_{ab} = \langle \theta_a \theta_b | \frac{1}{r_{12}} | \theta_a \theta_b \rangle, \tag{2.21}$$

$$K_{ab} = \langle \theta_a \theta_b | \frac{1}{r_{12}} | \theta_b \theta_a \rangle, \tag{2.22}$$

where θ_a and θ_b are orbitals and r_{12} is the inter-electron separation. Equation 2.21 describes the repulsion an electron in orbital θ_m experiences from an electron in orbital θ_n. It is simply an electrostatic, and thus purely classical, interaction. J is often approximated as a one-electron integral meaning that each electron only experiences an average of the interactions with all the other electrons in the system. The Exchange term arises from the antisymmetry of the wavefunction in that electrons of the same spin tend to avoid one another. Unlike the Coulomb term it is entirely quantum mechanical in nature. If we take these along with the one-electron Hamiltonian, then the energy can be expressed as in equation 2.23.

$$\tilde{E} = \sum_{ia=1}^{n} h_a + \frac{1}{2} \sum_{a=1}^{n} \sum_{b=1}^{n} (J_{ab} - K_{ab}). \tag{2.23}$$

Splitting up the one and two electron terms in the manner that we have are referred to as the Slater-Condon rules.[135–137]

Given that these terms deal with the repulsion and exchange interactions between electrons in different orbitals we need to be aware of the origin of these orbitals. It is commonplace for orbitals to be constructed from a linear combina-

tion of atomic orbitals, LCAO, as outlined in equation 2.24.

$$\theta_i = \sum_p^m c_{pi}\chi_p.$$ (2.24)

In which c_{pi} is a variational parameter and χ_p is a basis function. Once the atomic orbitals have been combined to form the molecular orbitals these molecular orbitals then make up the Slater determinant, $\Theta = |\theta_i(1)...\theta_n(n)\rangle$, which then leads to the energy.

2.3 Hartree-Fock self-consistent field theory

As a result of the LCAO approximation the trial energy depends on the values of the variational coefficients, c_{pi}, which form an *N* by *N* matrix. To obtain the best energy the c_{pi} parameters are varied, and through them the orbitals, until a minimum is reached. In contrast to the Molecular Orbitals, the atomic orbitals are not necessarily orthogonal. This means we need to consider the overlap between them. We can do this by using the overlap matrix, also *N* by *N*, which contains the overlap integrals of the basis functions as shown by equation 2.25.

$$S_{pq} = \int \chi_p\chi_q d\tau.$$ (2.25)

Thus, the variation principle gives rise to a set of secular equations, see equation 2.26,

$$\sum_{q=1}^m F_{pq}c_{qi} = \epsilon_i \sum_{q=1}^m S_{pq}c_{qi},$$ (2.26)

which equates to equation 2.27

$$|\mathbf{F} - \epsilon\mathbf{S}| = 0,$$ (2.27)

but can also be written in matrix form as $\mathbf{Fc} = \epsilon\mathbf{Sc}$, known as the Roothaan-Hall secular equation.[138] Here $\mathbf{S}$ is the overlap matrix, as defined in equation 2.25, and $\mathbf{c}$ is the matrix of variational parameters; ϵ is a diagonal matrix of the orbital energies and $\mathbf{F}$ is the Fock matrix. The Fock matrix is given by equation 2.28

$$F_{pq} = h_{pq} + \sum \langle\chi_p\chi_r||\chi_q\chi_s\rangle\rho_{sr},$$ (2.28)

where $h_{pq} = \langle \chi_p | -\frac{1}{2}\nabla^2 - \frac{Z_A}{r_A} | \chi_q \rangle$ and ρ_{sr} is a matrix element of the one-particle reduced density matrix ρ which is given by equation 2.29.

$$\rho_{sr} = \sum_1 c_{is}^* c_{ir}. \tag{2.29}$$

If the basis were complete, continuous, and spanned all of space these matrix elements would be $\rho(r, r')$, where r and r' are positions. The diagonal terms would then simply be the electron density, and so, represent the probability of finding an electron at position r. However, the basis is not continuous or complete, but discrete, thus s and r stand for orbitals (which span a discrete region of space), and so the diagonal terms, ρ_{ss}, represent the chances of finding an electron in orbital s.

The self-consistent field (SCF) method uses a single Slater determinant. During an SCF calculation, a set of basis functions and trial coefficients are chosen and used to form the wavefunction. From this, the overlap and Fock matrices can be evaluated. From here, the secular equations can then be solved to yield a set of orbital energies and a new set of coefficients. If the energy and density changes from one step to the next are smaller than the convergence criteria, the solution is self-consistent. The energy is then at a minimum which could be global or local. If this is not true, then the new coefficients are used in a new cycle, and the procedure is repeated until convergence occurs and the best energy is found. This process means that the HF energy of the system slowly approaches the true energy of the system and the result is optimised.

Hartree-Fock theory however, is only exact for a system with a single electron in it. Any system more complex than this will experience something known as correlation. Correlation is a term used to refer to the interactions between electrons in a system. HF does not include these interactions between particles, instead, particles only experience average interactions from all the other particles in the system due to approximating J, the Coulomb operator, as a one-electron integral. Thus the Hartree-Fock energy differs from the exact energy by something known as the correlation energy. The physical basis of correlation energy is made up of two parts, the dynamic and static correlation.[139–142] Dynamic correlation covers the electrons being located unphysically close to one another due to only experiencing an average interaction. Static correlation is more complicated and can be most easily thought of as the correlation stemming from only using one electronic configuration. Methods that go beyond Hartree-Fock attempt to find ways to reintroduce this correlation and thus come closer to the exact energy of the system.

2.4 Density functional theory

The basis of density functional theory (DFT) for non-degenerate states are the Hohenberg-Kohn (HK) theorems, the first of which states that any property of the ground state of a system is a functional of the ground state electron density, ρ.[143] The integral of ρ defines the number of electrons, the cusps define the positions of nuclei and their heights define the nuclear charge. Therefore ρ contains all the information required to determine the properties of a system. The second Hohenberg-Kohn theorem says that, for a trial density, the resulting energy is always an upper bound on the energy for the true density. This means that the trial ρ is variational and so the *Variation Principle* can be used to find the optimum density.

Hohenberg and Kohn though, did not suggest a form for the functional which would give the exact energy from the density. The most common of approaches to this is the Kohn-Sham (KS) approach, which uses mean-field and self-consistent methods.[144] Thus, whilst also making use of the Born-Oppenheimer approximation, the following equation for the electronic energy can be written.

$$
\begin{aligned}
E[\rho_0] &= V_{ne}[\rho_0] + F_{HK}[\rho_0] \\
&= \langle \rho_0 | \hat{V}_{ne} | \rho_0 \rangle + \langle \rho_0 | \hat{T}_{ee} | \rho_0 \rangle + \langle \rho_0 | \hat{V}_{ee} | \rho_0 \rangle \\
&= V_{ne}[\rho_0] + T_e[\rho_0] + V_{ee}[\rho_0],
\end{aligned}
\tag{2.30}
$$

where F_{HK} is the Hohenberg-Kohn functional which can be split into kinetic and potential elements. The potential operator, $\hat{V}_{ee}$, includes some Coulombic interactions and so can be written as $\hat{V}_{ee} = \frac{1}{2}\hat{V}_{12} + \hat{V}_{nc}$, with the first term describing the Coulombic part and the second the remaining non-Coulombic interactions. The exact nature of both $\hat{V}_{nc}$ and the kinetic operator $\hat{T}_e$ are unknown but based on Hartree-Fock theory, $\hat{V}_{nc}$ should include some kind of exchange.

KS DFT has some similarities with Hartree-Fock theory in that it swaps the true electronic density, ρ, for the electronic density of a system of non-interacting particles ϱ . Doing so allows for the density to be written as the sum of a series of single particle densities.

$$
\varrho(\tau) = \sum_i |\varphi_i(\tau_i)|^2.
\tag{2.31}
$$

Here φ_i is the single particle function of particle i. From here we can construct an equation for the energy using this substitute density as in equation 2.32,

$$E[\varrho] = \langle \varrho | \hat{T}_e | \varrho \rangle - \langle \varrho | \hat{V}_{ne} | \varrho \rangle + \langle \varrho | \frac{1}{2} V_{12} | \varrho \rangle + \langle \varrho | \hat{f}_{xc} | \varrho \rangle$$

$$= -\frac{1}{2} \sum_i \langle \varphi_i | \hat{\nabla}_r^2 | \varphi_i \rangle - \sum_i \langle \varphi_i | \frac{Z_A}{r_{Ai}} | \varphi_i \rangle + \frac{1}{2} \langle \varphi_i | \hat{V}_{ij} | \varphi_j \rangle + \langle \varrho | \hat{f}_{xc} | \varrho \rangle \qquad (2.32)$$

$$= T_e[\varrho] + V_{ne}[\varrho] + J[\varrho] + E_{xc}[\varrho],$$

where $T_e[\varrho]$ is the non-interacting kinetic energy; $V_{ne}[\varrho]$ is the Coulombic interactions between the nuclei and the non-interacting electronic density; $J[\varrho]$ is the Coulombic interactions between the particles in ϱ and $E_{xc}[\varrho]$ is the exchange-correlation energy. The form of the functional operator for the exchange-correlation, $\hat{f}_{xc}$, is at present not known.

An analogy between the Fock operator from Hartee-Fock theory and the Kohn-Sham exhange-correlation operator can be made by considering how $\hat{f}_{xc}$ acts on the single particle function φ_i. Combine this with expanding the single particle functions out in terms of basis functions and an equivalent of the Roothaan equations can be formed.[145]

$$\varphi_i = \sum_\mu c_{\mu i} \vartheta_\mu. \qquad (2.33)$$

Here, in equation 2.33, $c_{\mu i}$ is a coefficient and ϑ_μ is a basis function. From here it is then possible to write a matrix equation, again in analogy to Hartree-Fock theory, which can be solved self-consistently.

As previously mentioned, the exact form of the exchange-correlation operator $\hat{f}_{xc}$ is not known meaning that the term for the exchange-correlation energy, E_{xc}, is not known either. Consequently it must be approximated. The way in which this is done varies depending on which generation of DFT functional is being used. The earliest is the local density approximation (LDA) in which E_{xc} takes the form shown in equation 2.34,

$$E_{xc}[\rho] = -c_x \int \rho^{\frac{4}{3}}(\mathbf{r}) d\mathbf{r}, \qquad (2.34)$$

which is derived from the uniform electron gas, making it effective for use in solid state physics but less so for chemistry.[144] The generalised gradient approximation (GGA) marks the second generation of DFT functionals. GGAs depend both on ρ and its derivative as can be seen in equation 2.35

$$E_{xc}^{B88} = E_{xc}^{LDA} - \beta \rho^{\frac{1}{3}} \frac{x^2}{1 + 6\beta x \sinh^{-1} x}, \qquad (2.35)$$

where $x = |\nabla\rho|/\rho^{\frac{4}{3}}$. β is simply a parameter that has been fitted to a set of test data.[146] GGAs are more accurate for thermochemistry than LDAs, but both underestimate barrier heights.[147]

The next set of functionals are the hybrid functionals, so called, because they contain Hartree-Fock-like exchange terms in conjunction with LDA exchange terms. They are often more accurate than pure functionals. One, well-known, example of a hybrid functional is that of Becke, Lee, Yang and Parr with three parameters, B3LYP. The E_{xc} term for B3LYP, see equation 2.36, contains, as the name suggests, three parameters which were fitted to a training set of experimental atomisation energies, ionisation potentials and proton affinities.[148,149]

$$E_{xc}^{B3LYP} = (1-a)E_x^{LDA} + aE_x^{HF} + b\Delta E_x^B + (1-c)E_c^{LDA} + cE_c^{LYP}. \qquad (2.36)$$

Such functionals have been found to perform well in calculating properties such as vibrational frequencies and electronic energies.[150–153]

2.5 Time-dependent density functional theory

In all previous sections the time-independent Schrödinger equation has been used. However, if the matter of interest is the response of a system to some stimulus over time, then the time-independent form of the equation is no longer appropriate. A means to include time-dependence into the HK theorems was developed by Runge and Gross. These are, unsurprisingly, known as the Runge-Gross (RG) theorems.[154] The first RG theorem is a time-dependent parallel to the first Hohenberg-Kohn theorem. If an external potential is turned on at t_0 then the time-dependent electronic density, $\rho(\tau, t)$, determines the wavefunction plus some phase factor as described in equation 2.37,

$$\Theta(\tau, t) = e^{i\psi(t)}\Theta[\rho, \Theta_0](\tau, t), \qquad (2.37)$$

where Θ_0 is the initial wavefunction $\Theta(\tau_0, t_0)$. If the system under study is in the ground state then the first RG theorem can be combined with first HK theorem (see equation 2.30) so that the time-dependent wavefunction depends on the initial electronic density of the system up to some phase factor. The second HK theorem, the one that states the energy functional is variational in nature, can also be extended to a time-dependent case. To do this one must use the time-dependent action which, in its functional form, can be seen in equation 2.38,

$$A[\rho(\tau, t)] = \int_{t_0}^{t_1} \langle\Theta[\rho(\tau, t)]|i\frac{\partial}{\partial t} - \hat{H}(\tau, t)|\Theta[\rho(\tau, t)]\rangle\, dt, \qquad (2.38)$$

which has a stationary point at the correct time-dependent electronic density.[155,156]

It is convenient to rewrite the action function in terms of the time-dependent analogue of the HK functional and an external potential operator, see equation 2.39.

$$A[\rho(\tau, t)] = F^{HK}[\rho(\tau, t)] - \int_{t_0}^{t_1} \int \hat{u}(\tau, t)\rho(\tau, t)d\tau dt. \qquad (2.39)$$

As in time-independent DFT, if the electronic density ρ is replaced with that of a system of non-interacting particles, ϱ, then it is possible to form a time-dependent equivalent of the Kohn-Sham single-particle functional equation, equation 2.40.

$$-\frac{1}{2}\hat{\nabla}_r^2 - \sum_A \frac{Z_A}{r_{ai}} + \sum_{k \neq j} \langle \varphi_k(\tau_k, t)|\hat{v}_{jk} + \hat{u}(\tau, t) + \hat{\alpha}_{xc}|\varphi_j(\tau_j, t)\rangle = i\frac{\partial}{\partial t}|\varphi_j(\tau_j, t)\rangle$$

$$\hat{A}(t)|\varphi_j(\tau_j, t)\rangle = i\frac{\partial}{\partial t}|\varphi_j(\tau_j, t)\rangle. \qquad (2.40)$$

Here, $\hat{A}(t)$ is the time-dependent action function operator and it contains the time-dependent action exchange-correlation function operator, $\hat{\alpha}_{xc}$. As with the exchange-correlation operator in time-independent DFT, the form of this operator is unknown. A common approach to this problem is to use the exchange-correlation function operator from time-independent DFT, making the assumption that the exchange-correlation function reacts instantaneously to the time evolution of ϱ.

There are a variety of approximations to TD-DFT of varying complexity. One simple one is Linear-Response adiabatic TD-DFT (LR-TD-DFT) which stems from time-dependent Hartree-Fock theory. It makes use of the polarisability of the ground state of the system to garner information about the singly-excited electronic states. Its implementation makes use of the Casida equations, a subject beyond the scope of this work, which give the linear response of a system to a perturbation.[157–159] LR-TD-DFT makes use of the KS orbitals from time-dependent KS-DFT to simplify the Casida equations down to the following:

$$(\mathbf{A} - \mathbf{B})^{\frac{1}{2}}(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B})^{\frac{1}{2}}\mathbf{Z} = \vec{\omega}^2\mathbf{Z}, \qquad (2.41)$$

in which $\mathbf{Z} = (\mathbf{A} - \mathbf{B})^{\frac{1}{2}}(\mathbf{X} + \mathbf{Y})$; $\mathbf{A}$ and $\mathbf{B}$ are orbital rotation Hessian matrices; $\vec{\omega}$ is a vector containing excitation energies and $\mathbf{X}$ and $\mathbf{Y}$ are coefficient matrices. Solving these equations means that $(\mathbf{A} - \mathbf{B})$ needs to be diagonalised. This is not necessarily always computational tenable. Additionally, it is not always necessary to solve the full Casida equations as in LR-TD-DFT. Approximations can be made to 2.41. For instance, $\mathbf{B}$ can be neglected to give equation 2.42. However, this is

quite a severe approximation to make.

$$\mathbf{A}\mathbf{X}_{TDA} = \vec{\omega}_{TDA}\mathbf{X}_{TDA}. \tag{2.42}$$

This is the Tamm-Dancoff Approximation (TDA) in which $\mathbf{X}_{TDA}$ is the LR-TD-DFT coefficient matrix and $\vec{\omega}_{TDA}$ is a vector of excitation energies within the TDA approximation to LR-TD-DFT.[160] TDA has been found to reproduce band shapes for the emission and absorption spectra for various conjugated molecules however, it does underestimate the relative peak intensities, making it a reasonable approximation.[161]

2.6 Solid state density functional theory

Up until this point all the theory discussed is primarily used on single molecules or small clusters in the gas-phase. While it is possible to include solvent in these systems either explicitly or implicitly it is not so feasible to use these methods to treat, for example, crystalline solids.[162] A crystalline solid consists of an infinite number of repeating units in two or three dimensions. These units are called unit cells. The vectors along which the repeating unit cells are placed to recreate the infinite system are the unit, or primitive, vectors and are represented as a_i where $i = 1, 2, 3, ...$ It is combinations of these unit vectors that make up what is known as the Bravais Lattice.[163] The Bravais lattice and the symmetry point group of the unit cell allow us to classify crystalline solids into one of 230 space groups.[164]

As stated above, it is not tractable to study the infinite number of electrons contained within an infinite, repeating structure, and so this problem must be circumvented. This can be done using Bloch's theorem, which makes links between the properties of the electrons in an infinite periodic system with those in the unit cell.[165] More exactly, it states that the wavefunction of an electron in an external periodic potential is:

$$\psi_k(\mathbf{r}) = e^{i\mathbf{k}.\mathbf{r}}u_k(\mathbf{r}), \tag{2.43}$$

where,

$$u_k(r) = u_k(r + a_i). \tag{2.44}$$

Thus the wavefunction is the product of a function with the same periodicity as the potential and a complex phase factor that stems from the translational symmetry of the system. This then means that the wavefunction at some location one unit vector away from the original point becomes that given in equation 2.45.

$$\psi_k(r + a_i) = e^{i\mathbf{k}.a_i}\psi_k(r). \tag{2.45}$$

Therefore, the probability density is exactly the same at both points as the phase factor cancels out.[163] Additionally, if certain vectors, **k**, are chosen then $e^{i\mathbf{k}\cdot a_i}$ will equal one meaning the wavefunction will be in phase in all periodic repeats of the unit cell.

A set of the three smallest vectors that satisfy this are sufficient to describe all of the reciprocal lattice vectors. These are an analogue to the unit lattice vectors in real space, with reciprocal lattice being the Fourier Transform of the lattice in real space. The primitive vectors in reciprocal space are defined by equation 2.46.

$$a_i \cdot b_i = 2\pi \delta_{ij}. \tag{2.46}$$

This leads to a set of primitive reciprocal lattice vectors which are defined by;

$$b_1 = 2\pi \frac{a_2 \times a_3}{\Omega}, b_2 = 2\pi \frac{a_3 \times a_1}{\Omega}, b_3 = 2\pi \frac{a_1 \times a_2}{\Omega}, \tag{2.47}$$

in which Ω is the volume of the unit cell. These primitive reciprocal vectors can then be used to define the volume of a cell in reciprocal space which is referred to as the first Brillouin zone.[166] It is possible, then, to calculate the wavefunction for an infinite system based on the wavefunction for the electrons in the finite unit cell using an, in principle, infinite number of primitive reciprocal, or k, vectors in the first Brillouin zone.[164] However using a infinite number of k vectors is not necessary. Nearby k vectors contain nearly identical information so it is possible to construct a suitable wavefunction using only a finite number of k vectors.

Another tool used in solid state DFT is periodic boundary conditions. These allow for the macroscopic properties of a system to be calculated from a much smaller collection of atoms. In essence, tiles or boxes of a small repeating unit are arranged in two or three dimensions to replicate the entire, infinite system. Should a particle leave the box or tile at one edge it is replaced by an equivalent entering the box or tile from the opposite edge to maintain the total number of particles within the box. The periodic boxes must, however, be of the correct size. Boxes that are too small can lead to unphysical interactions between the system and its own image in a neighbouring box.[167]

The way in which the wavefunction of a unit cell can be extended to describe a large repeating system has been outlined above, but as in non-periodic DFT, some basis is needed to construct the orbitals. However, unlike non-periodic DFT which uses localised basis sets, solid state calculations make use of plane waves, which are simply sine and cosine functions. Each orbital wavefunction is made up of a linear combinations of these plane waves each with different reciprocal

lattice vectors.[168]

$$\Psi_i^k(r) = \sum_{\mathbf{G}} a_{i,k+\mathbf{G}} e^{i(k+\mathbf{G})\cdot r}. \tag{2.48}$$

In this, k is the wavevector that characterises the wavefunction, $\mathbf{G}$ is the reciprocal lattice vector and r is the general position vector. Taking this into account the KS equation may be written in the following form:

$$\sum_{G'}\{\frac{\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|\delta_{GG'} + V_{ion}(\mathbf{G}-\mathbf{G}') + V_{elec}(\mathbf{G}-\mathbf{G}') + V_{XC}(\mathbf{G}-\mathbf{G}')\}a_{ik+\mathbf{G}'} = \epsilon_i a_{i,k+\mathbf{G}'}, \tag{2.49}$$

where V_{ion} is the electron-nuclear functional, V_{elec} is the electron-electron functional and V_{XC} is the exchange-correllation functional.

Plane waves have several benefits, including the ease with which one can change from real space representation to reciprocal space representation *via* fast Fourier transforms. Additionally, basis set superposition errors that exist when using localised basis sets are avoided. Basis set superposition errors are due to the mismatch in energy that occurs between two molecules calculated at short and long range as the basis functions of the two halves overlap at short range. However, the description of the electron density around the nucleus is not done well by plane waves. To properly describe the density in these regions various approaches exist. One of these is the use of pseudopotentials. These potential functions represent the way the valence electrons interact with the nucleus and core electrons.[168] Various forms are available for pseudopotentials depending on how many plane waves they are constructed from. Soft pseudopotentials, which use fewer plane waves, decrease computational cost, and even ultra- and super-soft pseudopotentials exists. However, the softer the potential the less transferable between atomic environments it is. Therefore a balance must be struck.

An alternative approach is that of the projector augmented wave (PAW) which is more accurate than using ultrasoft pseudopotentials.[169] PAW methods use different basis functions for different regions of the system. Atomic regions use basis functions made from atom-like partial waves whereas the basis functions in bonding regions are a set of envelope functions, the smooth curve outlining the extremes in amplitude of a rapidly oscillating signal. The partial solutions for the different regions are then matched at the interface. Essentially, PAW methods combine augmented wave methods and pseudopotentials to give a unified electronic structure approach.[170]

2.7 Multi-configurational self-consistent field theory

Up until this point the majority of the discussion has covered various kinds of DFT which, like HF theory, is a single-determinant method. This means that only one electronic configuration is taken into account during the calculations. There are systems, however, that require more than one configuration to be included in order to describe the chemistry occurring properly. Such processes include the breaking and forming of bonds, or reactions involving diradicals and first row transition metals. Multi-configurational self-consistent field theory (MCSCF) is one way to approach calculations on multi-reference systems.[171,172]

Part of the basis of MCSCF is Configuration Interaction (CI). In CI methods the non-relativistic Schrödinger equation is solved within the BO approximation. In CI an N electron basis function can be written as substitutions or excitations from the HF reference determinant as in equation 2.50

$$|\Psi\rangle = A_0|\Phi_0\rangle + \sum_{ra} A_a^r|\Phi_a^r\rangle + \sum_{a<b,r<s} A_{ab}^{rs}|\Phi_{ab}^{rs}\rangle + \sum_{a<b<c,r<s<t} A_{abc}^{rst}|\Phi_{abc}^{rst}\rangle + ..., \quad (2.50)$$

where A is some variational coefficient, $|\Phi_0\rangle$ is the HF reference determinant and $|\Phi_a^r\rangle$ is the Slater determinant formed by replacing spin orbital a in $|\Phi_0\rangle$ with spin orbital r. This constitutes a single excitation from the ground state wavefunction and terms corresponding to double, triple and higher excitations can be included into the CI wavefunction. Every N-electron Slater determinant can be described using the set of N spin orbitals it is formed from. This set of orbital occupancies is called a configuration interaction, or configuration state function (CSF), hence the name of this method.[173] By constructing the wavefunction as a linear combination of these CSFs, the CI method accounts for electron correlation by expanding the wavefunction beyond a single determinant.

MCSCF uses a combination of HF and Configuration Interaction (CI). In MCSCF one constructs the wavefunction from a linear combination of Slater determinants as in equation 2.51.

$$\begin{aligned}\Psi_{MCSCF} &= \sum_K A_K \Phi_K \\ &= \sum_K A_K A \prod_{i \subset K} \sum_\mu \chi_\mu c_{\mu i}.\end{aligned} \quad (2.51)$$

This is a linear combination of CSFs, much as is done in CI, which differ in how the electrons are arranged in the MOs, ϕ_i. Unlike CI, both the A_K coefficients and the $c_{\mu i}$ coefficients are optimised variationally. If all possible CSFs within a given active space are used then what is known as complete active space SCF

(CASSCF) is being used.[172] Alternatively, if only some of the possible CSFs are selected by the user then this is restricted active space SCF (RASSCF).[174]

A good, demonstrative example of the need for a multi-configurational wavefunction is when looking at how the MOs of CH_2 change as a function of the HCH angle. Firstly, let the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) be a_1 and b_1 respectively. Now, when the HCH angle is small the a_1 orbital is a combination of the carbon $2s$ and $2p_z$ while the b_1 orbital is purely carbon $2p_x$. This changes as the angle approaches linearity the $2s$ contribution to the a_1 orbital decreased to zero and so the HOMO energy rises. Finally, at an HCH of $180°$ the two MOs merge to form a π MO.[175]

Thus, the ground state singlet wavefunction can be described by equation 2.52 when the structure is bent.

$$\Psi = a_1(\alpha)a_1(\beta).$$

(2.52)

However, as the angle increases towards linear the configuration with two electrons in the b_1 MO becomes increasingly important such that when linear the singlet wavefunction becomes:

$$\Psi = 2^{\frac{1}{2}}\{a_1(\alpha)a_1(\beta) - b_1(\alpha)b_1(\beta)\}.$$

(2.53)

Thus, to describe the bend potential of the CH_2 singlet state at all angles the wavefunction needs to be

$$\Psi = A\{a_1(\alpha)a_1(\beta)\} + B\{b_1(\alpha)b_1(\beta)\}.$$

(2.54)

This is one of the simplest possible MCSCF wavefunctions, the two-configuration SCF wavefunction.

It is possible to include perturbation theory into multi-configurational calculations, termed CASPTn calculations where the *n* indicates which order of perturbation theory is employed. The order of the perturbation is the point at which the power series describing the perturbation is curtailed. Perturbation theory allows for the approximate solution of a complicated quantum chemical system by solving the problem for a simpler system and then applying a correction in the form of a small perturbation.[176] As such it can allow for the addition of electron correlation into a system.[177] CASPTn methods, in particular, take a multi-configurational reference function and so can be applied to radical or transition metal systems which benefit from a multi-configurational approach.

2.8   Molecular Mechanics

Unlike the previous sections, molecular mechanics (MM) does not use quantum mechanics, but rather classical models to predict the energy of a system as a function of its geometry. Such approaches are often employed for geometry optimisations of systems that are too large to feasibly tackle with quantum methods, such as those seen in chapter 4. In MM the energy of the system is constructed from potential energy terms for bond stretches, bends, torsional energy, van der Waals energy, electrostatics and cross terms, see equation 2.55

$$E = E_{str} + E_{bend} + E_{tors} + E_{vdW} + E_{el} + E_{cross}. \tag{2.55}$$

The stretching and bending potentials of bonds are treated similarly using truncated polynomial series to describe the increases in energy as the bond or angle deviates from the equilibrium geometry. The torsional energy term deals with the interactions resulting from changes in dihedral angles and is modelled using a Fourier series as the angle can deviate from equilibrium quite substantially. For atom pairs that are further than two atoms apart van der Waals interactions need to be calculated.[178] At short range such interactions are highly repulsive before becoming attractive at intermediate range. If the separation is too great then no interactions are felt. A computationally economical way to handle these forces is to use a Lennard-Jones potential, equation 2.56,

$$E_{vdW}(R^{AB}) = \epsilon \left[(\frac{R_0}{R})^{12} - 2(\frac{R_0}{R})^6 \right]. \tag{2.56}$$

The electrostatic terms, giving the interactions between partially charged atoms, use a Coulombic potential. Finally, the cross terms account for how some interactions have an effect on others, for instance how hydrogen bonding affects the stretching in OH bonds.[179]

MM methods can be used in conjunction with quantum mechanical methods through implementations like 'our Own N-layered Integrated molecular Orbital molecular Mechanics' (ONIOM).[180] ONIOM allows one to treat the region of the system where the chemistry under investigation is occurring using higher level methods and the rest of the system with a computationally cheaper approach.[181] Commonly quantum mechanics is used for the chemically interesting region and teamed with MM to describe the surrounding system.

The methods and theories laid out here have been used in the following chapters to follow astrochemical systems of interest. UV-Vis spectra of a variety of molecules have been generated using TD-DFT while radical systems have been

tackled using MCSCF. In addition to this, solid state optimised geometries and spectral properties have been found using plane wave methods. More detail on the exact implementation of the methods used in each chapter are given within the chapter.

3. Benzene and carbon dioxide ice irradiation

The chemistry that occurs on icy dust grain surfaces is a key linking feature of the chapters in this work. Some of the processes and molecules that result from the irradiation of such ices have already been discussed in the introduction. In this chapter the identities of the products of irradiation of a CO_2 and benzene ice mix by electrons are determined.

3.1   The biological relevance of carboxylic acids

The synthesis of biologically relevant molecules in the interstellar medium (ISM) is a topic of much interest. Firstly, to understand how the molecules that have been detected in the ISM were formed and then, secondly and as already discussed during the introduction, by elucidating the formation of these molecules in the ISM the origins of life on Earth may also be understood. Additionally, it could shed light on the possibility for similar biological chemistry occurring on other planets. In section 1.1.4 of the introduction it was briefly discussed how planets form from smaller entities in space, and thus, how the chemistry available and occurring on those smaller parts influence the chemistry that takes place on the planet as a whole. So, by understanding ISM chemistry light can be shed on the chemistry of early Earth as well as other planets.

One such small, biologically relevant, group of molecules are carboxylic acids. These are organic species whixh contain a hydrocarbon section bonded to the carboxylic acid group COOH. Carboxylic acids can act as biomarkers as well as intermediates to the formation of amino acids, proteins and lipids.[182–187] Thus far, both formic and acetic acids have been detected in the ISM.[87,188–191]

The biological relevance and successful detection of carboxylic acid species in the ISM has led to a number of studies from the astrochemical community. It has been shown that carboxylic acids can be formed *via* the irradiation of model ISM ices composed of suitable reactant materials. For instance, water and CO ice has been shown to yield formic acid, HCOOH, when irradiated with energetic (5 keV) electrons at 10 K.[192] Similar ice irradiation studies using methane and carbon dioxide have led to acetic acid and higher order acids, while CO_2 and a hydrocarbon ice also yields a mix of acids.[83,193,194]

One particular variety of carboxylic acids that are of interest are aromatic carboxylic acids. Benzoic acid has been identified in the Murchison and Orgueil meteorites and can be used as an intermediate in forming secondary metabolites.[195] In fact, the Murchison, Orgueil and Tagish Lake meteorites have been found to contain an array of more complex carboxylic acids ranging from succinic to glutaric acid (figure 3.1).[41] However, despite both the identification of these species

and the results of the laboratory ice studies, the mechanism for the formation of carboxylic acids in the ISM is not fully understood.
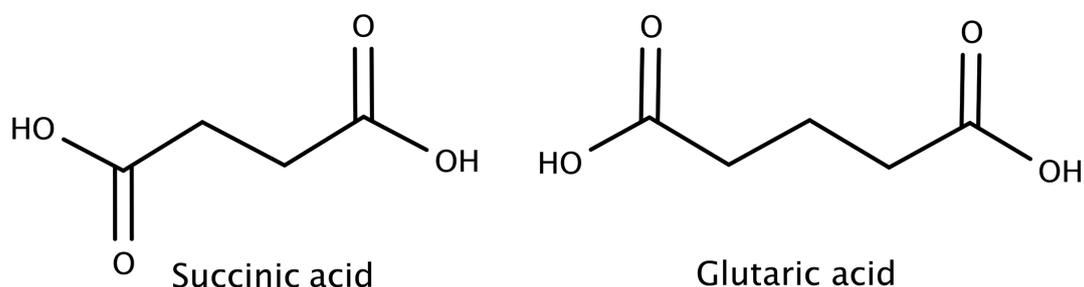
Figure 3.1: The structures of succinic and glutaric acids, both of which have been identified on meterorites.[41]

Suggestions have been made as to how these molecules form in the ISM, such as those by Irvine *et al.* which made use of the following gas-phase reactions:

$$CH_4 + O_2{}^+ \longrightarrow HCOOH_2{}^+ + H \tag{3.1}$$

$$HCOOH_2{}^+ + e^- \longrightarrow HCOOH + H \tag{3.2}$$
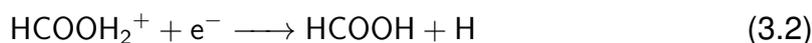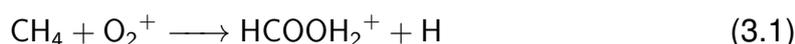
which is an ion molecule reaction followed by dissociative recombination.[98] For acetic acid specifically the following gas-phase formation route using protonated methanol has been posited,

$$H_3COH_2{}^+ + HCOOH \longrightarrow H_3CCOOH_2{}^+ + H_2O \tag{3.3}$$

$$H_3CCOOH_2{}^+ + e^- \longrightarrow H_3CCOOH + H \tag{3.4}$$

which is an alkyl cation transfer and then a dissociative recombination.[17] Both methanol and formic acid have been detected so, provided that protonated methanol can be formed, this could be a plausible formation route.[188,196] However, gas-phase formation routes have proved to not be as favourable as a grain surface based reaction pathway. Gas-phase formation routes give an average abundance of 2.5×10^{-11} relative to hydrogen while ice surface based routes give an average of 4.9×10^{-11}.[68,69] While the difference is not particularly severe it is clear that surface chemistry has its place in generating organics in the ISM, particularly as laboratory experiments on processed ices have yielded carboxylic acids alongside formyl and hydroxycarbonyl radicals.[193] Thus, a potential radical-based mechanism for the formation of carboxylic acids is outlined below.

$$H_2O + \longrightarrow HO\cdot + H\cdot \tag{3.5}$$

$$CH_4 \longrightarrow CH_3 \cdot + H \cdot \tag{3.6}$$

$$RH \longrightarrow R \cdot + H \cdot \tag{3.7}$$

$$H \cdot + CO \longrightarrow HCO \cdot \tag{3.8}$$

$$H \cdot + CO_2 \longrightarrow HOCO \cdot \tag{3.9}$$

$$HCO \cdot + OH \cdot \longrightarrow HCOOH \tag{3.10}$$

$$HOCO \cdot + CH_3 \cdot \longrightarrow CH_3COOH \tag{3.11}$$

$$HOOCH \cdot + R \cdot \longrightarrow RCOOH \tag{3.12}$$
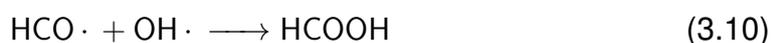
Energetic electrons cause the bond breaking reactions 3.5, 3.6 and 3.7 leading to radical products. The excess kinetic energy from these reactions then allow the addition of a hydrogen radical to either CO or CO_2, reactions 3.8 and 3.9, to give the formyl and hydroxycarbonyl radical respectively. There can be barrierless radical recombination like those described in reactions 3.10, 3.11 and 3.12 to create the carboxylic acid species.

It is possible for benzene to undergo the kinds of reactions laid out above to lead to benzoic acid and thus this may be the source of the benzoic acid found on meteorites. Indeed, studies on the irradiation of a 1:50 benzene and CO_2 ice mix at 10 K have identified benzoic acid and *meta-* and *para*-diacids. In these experiments mid-IR data was used to confirm the identity of the products.[197] Experimental work on a 1:100 benzene to carbon dioxide ice yields the UV-vis spectrum in figure 3.2 upon annealing to 200 K.[198] James *et al.* suggest that the residue left after annealing the ice is formed from a 'benzene-like' structure. However, the composition of the residue could not be unequivocally assigned. The previous work of McMurtry *et al.*, in which mid-IR data was used to identify the *meta-* and *para*-diacids, implies that these diacid structures could be present in addition to benzoic acid.[197]

As the exact nature of the residue remains unknown, we aim to use solid-state periodic DFT calculations to confirm which of the diacids are formed. The diacids being: phthalic acid (1,2-benzenedioic acid, figure 3.3b); isophthalic acid (1,3-benzenedioic acid, figure 3.3c) and terephthalic acid (1,4-benzenedioic acid, figure 3.3d). The aim of this work is to suggest the relative proportions thereof by comparison of calculated UV-vis spectra to experiment.
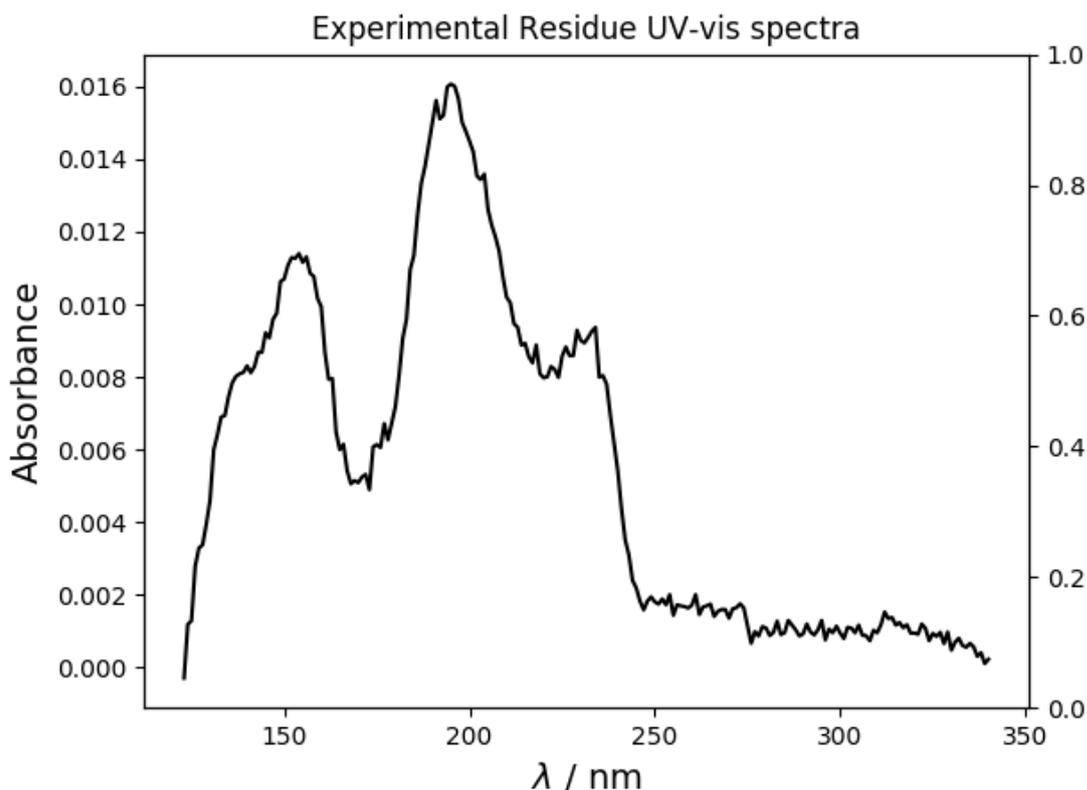
Figure 3.2: The UV-vis spectrum of the residue of an 1:100 benzene (C_6H_6): carbon dioxide (CO_2) ice processed with 1 keV electrons for 180 minutes and then annealed to 200 K. Reproduced with permission from James *et al.* 2019.[198]
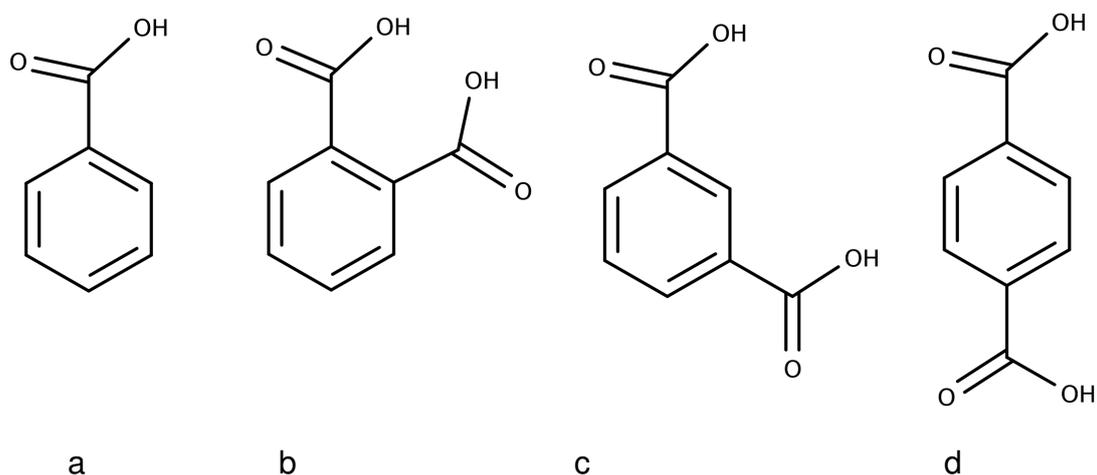
Figure 3.3: The structures of a) benzoic acid, b) phthalic acid, c) isophthalic acid and d) terephthalic acid.

3.2 Methodology

All calculations have been performed using the GAUSSIAN 09 suite of electronic structure programmes[†] using DFT with the M06 functional[199] using the

---

[†]GUASSIAN 09, revision D.01, M. J. Frisch et al., Gaussian, Inc., Wallingford, CT, 2009

def2-TZVP basis set.[200,201] In addition to this, the Vienna Ab-initio Simulation Package (VASP)‡ using DFT with the projector-augmented wave method[170,202] and the PBE0 functional.[203]

The ground electronic states of unit cells, as described in literature, of benzoic acid, phthalic acid, isophthalic acid and terephthalic acid in the gas-phase were optimised to a minimum in Gaussian 09.[204–207] Each cell contained four molecules of each acid arranged in two layers with hydrogen bonding interactions between the molecules in each layer, see figure 3.4. The IR and UV-vis spectra, calculated using time-dependent DFT (TD-DFT), for all systems were generated in Gaussian 09 at the M06/def2-TZVP level.[208,209] Periodic solid-state systems of each acid were optimised using VASP using their respective crystal structures[207,210–212] from which the solid-state IR and UV-vis spectra of the optimised systems were calculated at the PAW/PBE0 level (see appendix A.3 and A.4 for example inputs).[213]

All experimental data were gathered by David Ashworth of the Foster Group. Experimental FTIR measurements of powder samples of the individual acids for benchmarking purposes were taken on a Perkin Elmer Spectrum 2 FTIR spectrometer, equipped with a diamond UATR Two accessory. Sixteen scans were collected from 400 - 4000 cm^{-1}, with a resolution of 1 cm^{-1}. Terephthalic acid was analysed as supplied. Larger crystals of benzoic acid and phthalic acid were lightly ground with a marble pestle and mortar prior to analysis. Benzoic acid/terephthalic acid mixtures were tumbled prior to analysis to ensure homogeneity.

Experimental solid UV-vis measurements of ground samples of the individual acids for benchmarking purposes were taken on a Cary5000 spectrometer between 400 - 175 nm, with a resolution of 1 nm and a scan speed of 600 nm s^{-1}. Spectra were collected using a PrayingMantis[TM] Solid State sample stage reflectance accessory. The sample holder was coated in Teflon tape to minimise holder absorbance and a background recorded. Ground samples were sprinkled onto the tape. overlaid
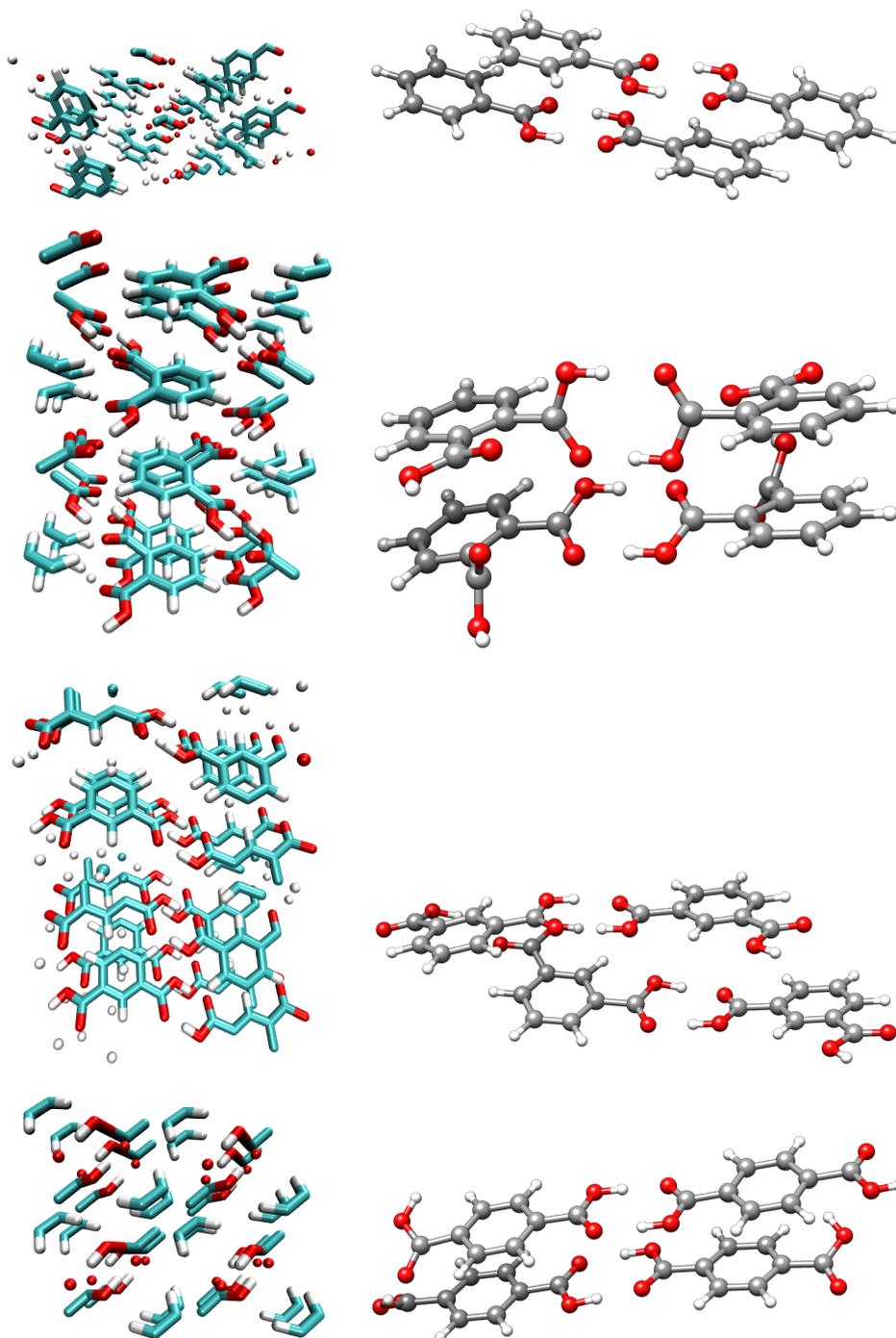
Figure 3.4: The optimised structures of the periodic unit from VASP(left) and the DFT gas-phase tetramer from Gaussian 09 (right) for top: benzoic acid, second row: phthalic acid, third row: isophthalic acid and bottom: terephthalic acid.

3.3 Using computational means to understand the experimental findings

3.3.1 IR benchmarking

Prior to performing any calculations to generate spectra, periodic solid-state DFT geometry optimisations were carried out on all of the acid systems using crystal structures from the Cambridge Crystallographic Data Centre database (CCDC)*. The optimised periodic cells can be seen in figure 3.4. In each case the anticipated hydrogen bonding and $\pi-$stacking interactions are evident in the alignment of the acid molecules. The gas-phase optimised unit cells can also be seen in figure 3.4. Comparison of the structures makes it obvious that not all of the long range order of the system is replicated in the gas-phase. In particular the alternating directions of the stacked sheets in the benzoic acid system is not seen in the gas-phase due to the limited scale of the calculation. The optimised structures in both the gas- and solid-phase have been used going forward to generate spectroscopic data.

Before undertaking the solid state calculations to generate the UV-vis spectra for the acids, non-periodic DFT and solid state DFT IR calculations were carried out and compared to experimental IR data to confirm the validity of the solid state approach. The experimental IR for each of the acids were obtained in the solid phase. Figure 3.5 shows the calculated and experimental spectra for each acid. Both the calculated spectra show the strong CO and OH absorbances (at 1600 and 2600 cm^{-1} respectively) one would associate with carboxylic acids at similar wavelengths to those in the experimental spectra. No shift has been applied to the VASP spectra. The gas-phase spectra have been scaled by 0.94, 0.91, 0.92 and 0.91, respectively for benzoic, phthalic, isopthalic and terephthalic acid, to account for the lack of anharmonicity in the calculations. The scaling factors were chosen so as to be the smallest possible to bring the CO, OH and CH_3 peaks into position.

The substantial broadening of the OH peak of the COOH group seen in the experimental and solid-phase DFT calculations arises from the extent of hydrogen bonding present in these systems. This also shifts the peaks to a lower wavelength than expected in the gas-phase due to the weakening of the OH bond.[214] As a consequence of this the peaks for the CH and OH stretches merge into one very broad peak. Examination of the individual vibrations show the alcohol stretches fall between 3210 and 3300 cm^{-1}. The same bond weakening effect

---

‡VASP 5.3.2 September 2012: https://www.vasp.at
*https://www.ccdc.cam.ac.uk
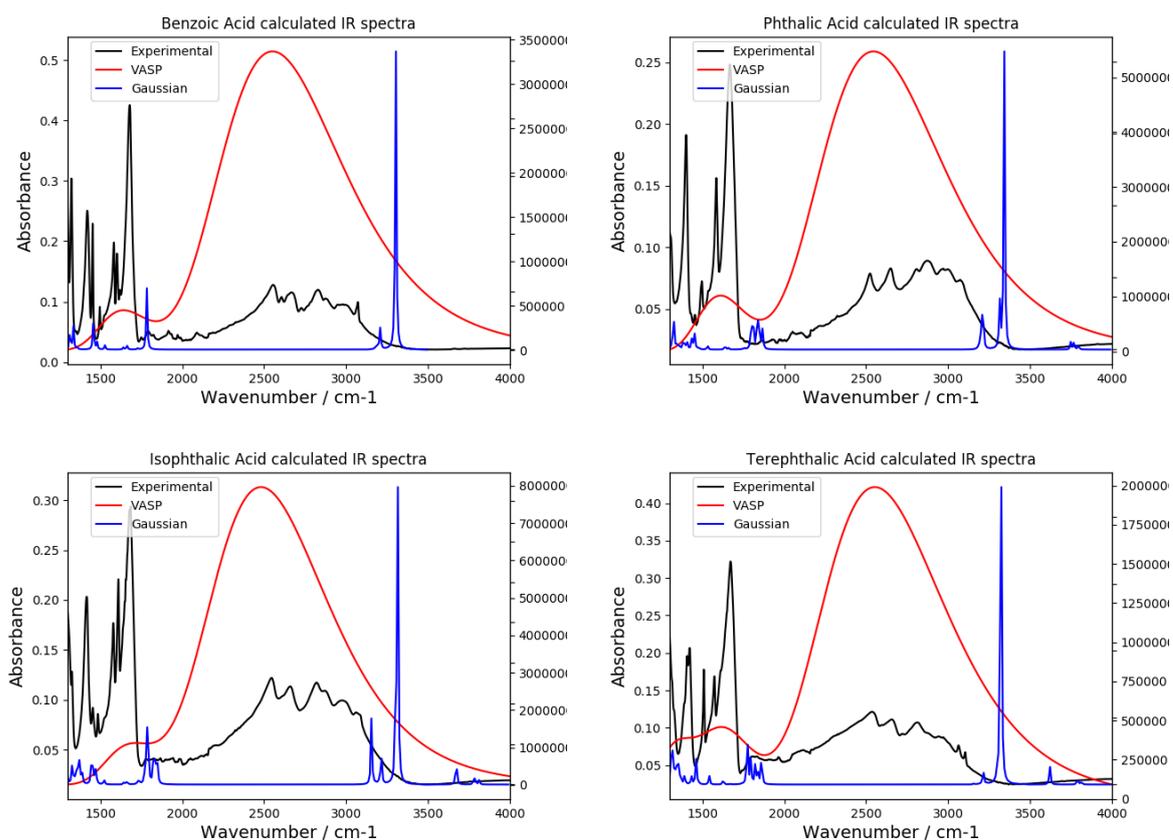
Figure 3.5: The experimental solid phase IR spectrum (black), the non-periodic gas-phase calculated spectrum (blue) and the periodic, solid phase spectrum (red) for top left: benzoic acid, top right: phthalic acid, bottom left: isophthalic acid and bottom right: terephthalic acid.

accounts for the shift of the CO double bond peak at roughly 1700 cm^{-1}. The fingerprint region has not been calculated in the solid state spectra as lower lying modes are not predicted as accurately and no spectral assignment is being made from the fingerprint region, hence it is not shown in the spectra. The rough agreement between the solid state DFT IR spectra and those generated experimentally for the key functional group peaks (CO and OH) suggests that the geometries generated through this methodology will be suitably representative of the system and can be used to calculate UV-Vis spectra to attempt to assign the identities of the residue components. It is also clear that the use of gas-phase methods are not sufficient to model the system as, while the peaks occur at a reasonable wavenumber, the peak shape poorly matches that from experiment.

3.3.2 Single acid systems

For each of the optimised solid-state acid systems VASP UV-vis calculations using TD-DFT on the previously optimised unit cells for each acid were performed. As with the IR calculations these were compared with experimental solid state UV-Vis measurements. The results are shown in figure 3.6, in this instance none of the spectra have been shifted. What is clear here is that the peaks require shifting in order to bring them into line with the experimental spectrum. The shifted spectra can also be seen in figure 3.6 and a shift of 0.5 eV has been applied. Going forward all calculated UV-Vis spectra have been shifted. The match between the experimental and calculated spectra is not strong. This is thought to stem from the fact that powder experimental samples were used whereas the calculations were on highly crystalline model systems and so show more distinct peaks.[215] As the experimental ice systems was annealed prior to taking its spectrum it is expected that the calculations will provide much better agreement.

Each calculated spectrum was plotted against the UV-Vis spectrum of the model benzene and carbon dioxide ice after it had been processed with high energy electrons. All the calculated spectra plotted against the residue spectra have been shifted by 0.25 eV to lower wavelengths to correct for overestimation of UV-Vis excitations by DFT.[216] Figure 3.7 shows each of these plots. It is worth noting that, as anticipated, the match between the residue spectra and the calculated solid-state spectra is much better than between the calculated spectra and the experimental powder UV-Vis. This is due to the residue having under gone heating which will have caused a transition from an amorphous ice to a more crystalline residue. Thus, the greater crystallinity of the residue will be better replicated by the crystalline calculated systems than the powder experimental samples were. It is clear that no single acid can account for the absorbance peak
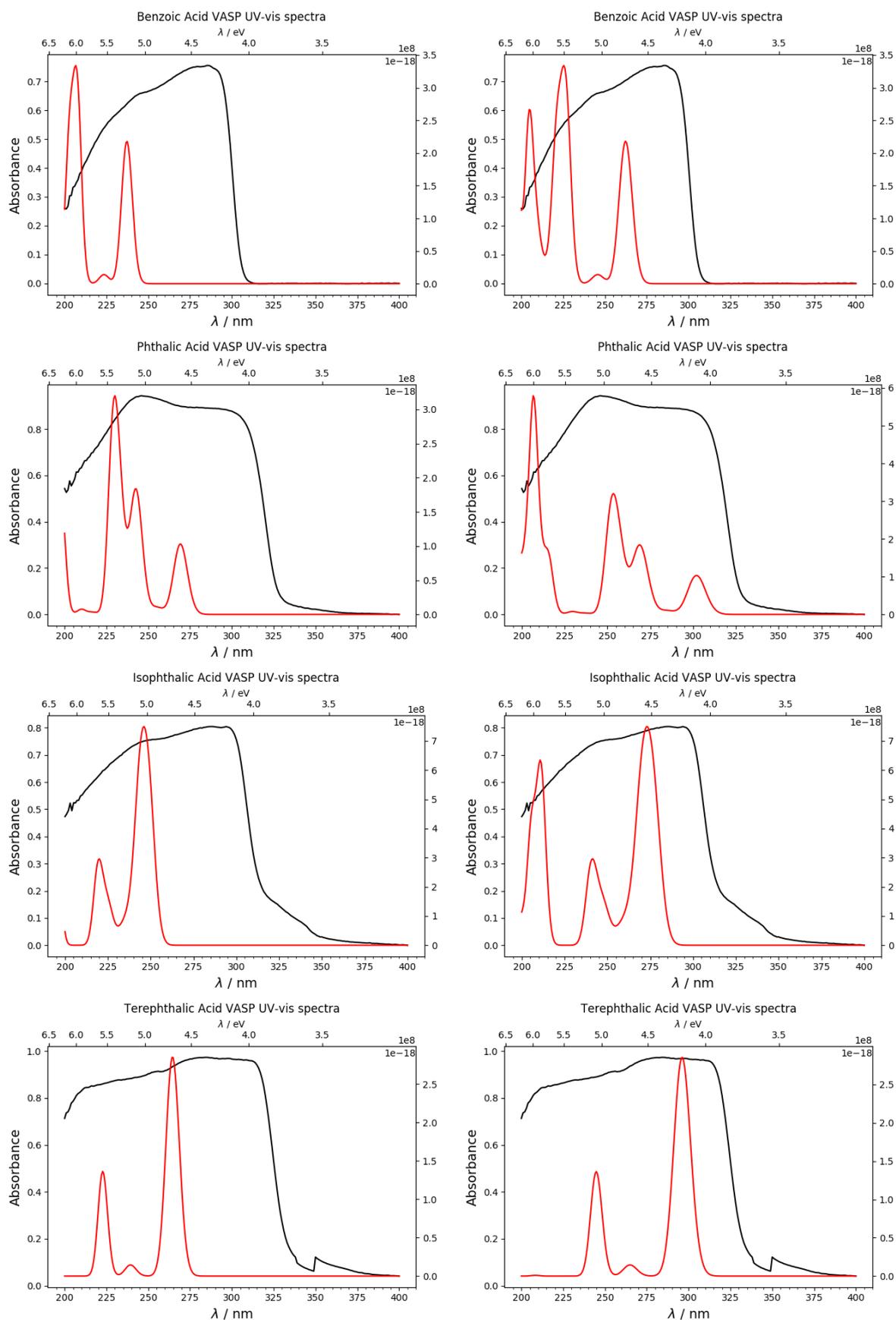
Figure 3.6: The experimental powder (black) and VASP calculated (red) UV-Vis spectra both unshifted (left column) and shifted by 0.5 eV (right column) for top: benzoic acid, second row: phthalic acid, third row: isophthalic acid and bottom: terephthalic acid.

61
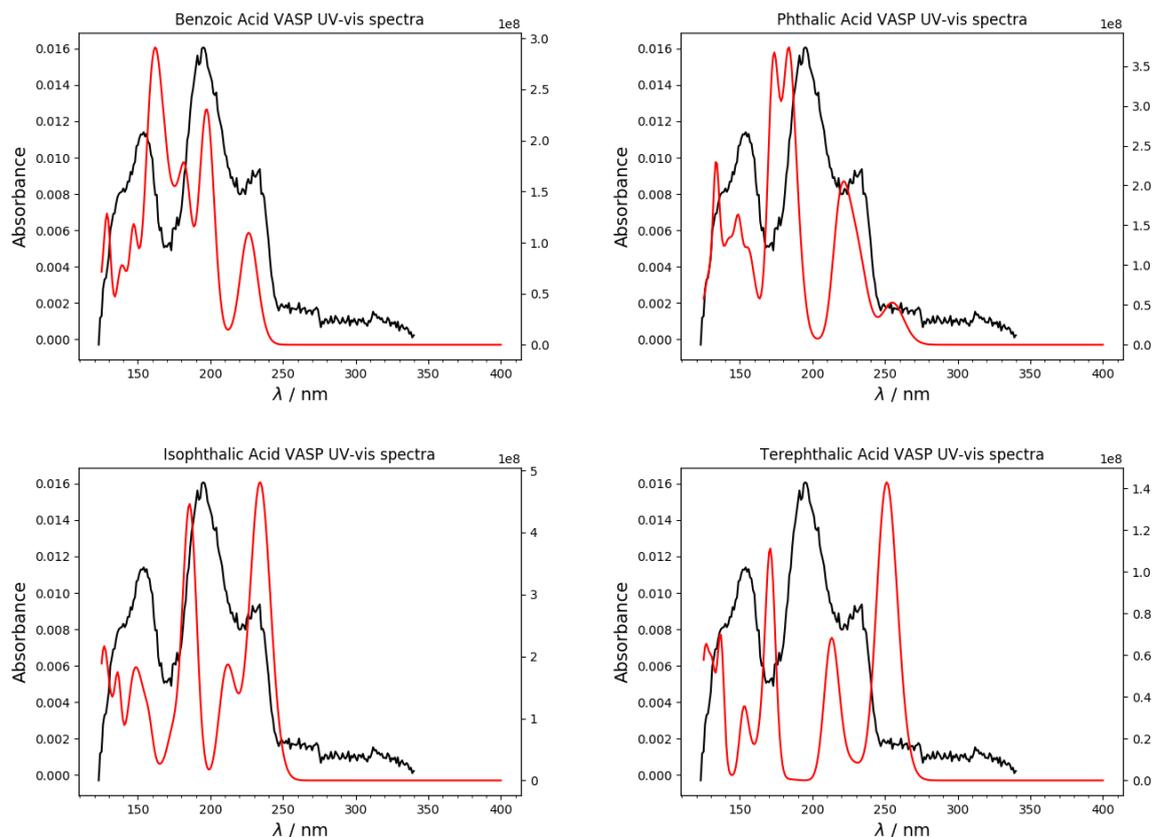
Figure 3.7: The model processed benzene and CO_2 ice experimental (black) and VASP calculated (red) UV-vis spectra for top left: benzoic acid, top right: phthalic acid, bottom left: isophthalic acid and bottom right: terephthalic acid.

shape of the irradiated ice. In particular the peak at 175 nm is not well represented in all samples. This finding is consistent with the literature in which both benzoic acid and *meta*- and *para*-diacids have been detected as the product of processing this kind of ice.[197]

3.3.3 Mixed acid systems

The proposed mechanism for the addition of CO_2 to benzene in ISM ices is based on radical-radical reactions. For any diacid to be produced a first addition of CO_2 must occur. There is also no certainty that any one benzene ring will undergo two additions. Additionally the kinetics for the first and second addition of a carboxylic group has been elucidated. The first addition has a rate constant of $(2.8 \pm 1.1) \times 10^{-4}$ s^{-1} while the second addition has a much lower rate constant of $(9.3 \pm 8.9) \times 10^{-6}$ s^{-1}.[197] Given that it is not certain that all benzene rings will undergo two additions and that the second addition proceeds much slower than the first it seems intuitive that the monoacid has been detected in these ices.
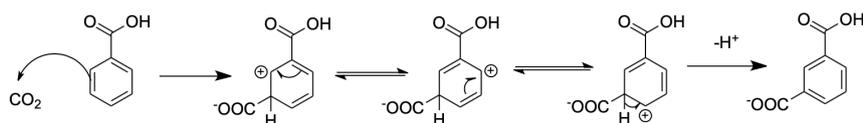
Figure 3.8: How a carboxylic acid group stabilises a radical on the benzene ring leading to *meta*-directed functional group addition.

The location of the second added carboxylic acid group can be influenced by the first carboxylic acid group. Carboxylic acid groups are known to possess *meta*- directing character due to charge stabilisation, see figure 3.8. However, given the suggested mechanism proceeds *via* radical chemistry, which does not affect the π system, this stabilisation effect may not be seen. There are steric arguments against the formation of phthalic acid, *ortho*-benzene diacid, as this structure places the carboxylic acid groups in the closest proximity and thus would experience the largest amount of steric hindrance. To confirm that phthalic acid is not part of the ice a composite spectrum of benzoic and phthalic acid in a 1:1 mix was made. Figure 3.9 shows the spectrum and as can be seen the peak at roughly 180 nm which is not present in the residue spectrum indicates that phthalic acid is not part of the products. The *meta*- and *para*- acids would not experience the same level of steric hindrance and so both species could form, likely in statistical mixtures.
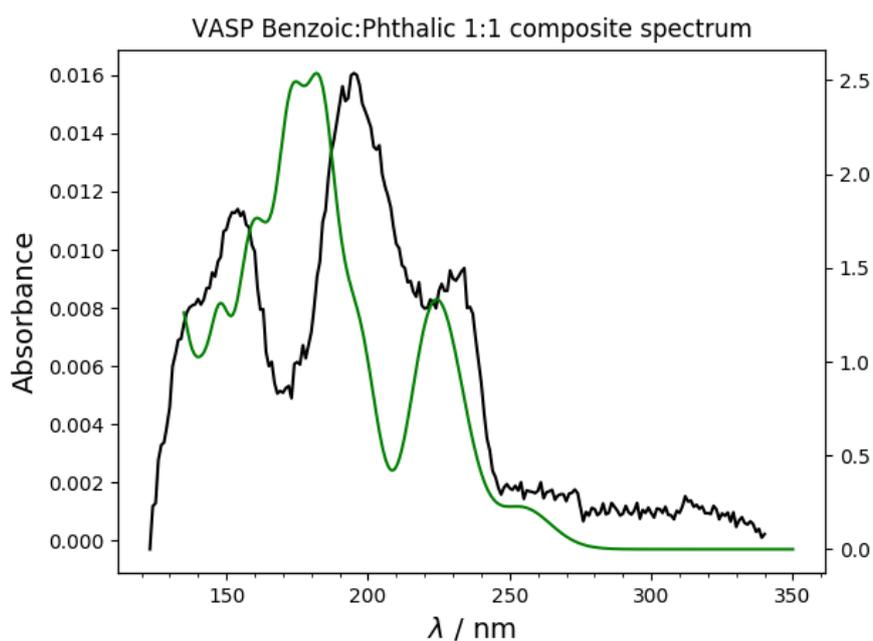
Figure 3.9: The composite calculated UV-Vis spectrum for a 1:1 benzoic acid:phthalic acid mix (green) compared with the experimental residue UV-Vis spectrum (black).

The *para-* acid, terephthalic acid, and the *meta-* acid, isophthalic acid, have both been proposed to have been detected. These detections were made on the basis of IR data from the processed ice. Figure 3.10 shows the experimental, gas-phase DFT and solid-state periodic spectra for both isophthalic and terephthalic acid overlaid. As is clear from these overlays both compounds have very similar IR spectra. The high degree of similarity would cause difficulties in precisely identifying one over the other from IR data alone. Hence, the IR data is inconclusive and what can be said is that both acids could be present. The calculated UV-Vis spectra of terephthalic acid and isophthalic acid, however, are more distinct and may allow for a much more confident identification, see figure 3.7.

In order to begin attempting to replicate the experimental data some thought needs to be had over what other likely residue components could be. It is unlikely that any benzene or CO_2 remain in the ice after the annealing process as both species desorb at temperatures lower than 200 K.[198] Consequently the residue will be comprised of a mixture of benzoic, isophthalic and terephthalic acids. Additionally, as the first addition of a COOH group is faster than the second it has also been anticipated that benzoic acid will comprise a larger part of the residue than either of the two diacids. Composite UV-Vis spectra have been made from the solid-state periodic DFT calculated spectra in which variable proportions of each neat spectrum have been summed together as Gaussian distributions to yield an overall multi-acid spectrum. These composite spectra (shifted by 0.25 eV to lower wavelengths for the same reasons as the neat acid spectra in section 3.3.2) have then been compared with the experimental residue spectrum.

Figure 3.11 shows the composite spectra for a mixture of benzoic and terephthalic acid (green trace) overlaid with the experimental residue spectrum (black). The mixtures used are a 3:1, 1:1 and 1:3 mix of benzoic acid to terephthalic acid. As becomes obvious in the mixtures with an equal or majority proportion of terephthalic acid there is a peak at roughly 250 nm which is not seen in the residue. Given that the first addition of CO_2 to benzene occurs faster than the second, this result is consistent with the idea that a larger quantity of the monoacid would be present. The closest shape match is that of the 3:1 benzoic acid to the terephthalic acid spectrum although this still reverses the relative intensities of the two peaks at the lowest wavelength, roughly 190 and 175 nm. Additionally, the shape of the peak at the lowest wavelength (150 nm) in the residue spectrum is not well matched by that in the calculated spectrum. Both benzoic and terephthalic acids have a peak at roughly 175 nm in their UV-Vis spectra which combine to give a signal in this region that is too large compared to the residue spectrum. This suggests that, should any terephthalic acid be formed it is a minor constituent of the residue.
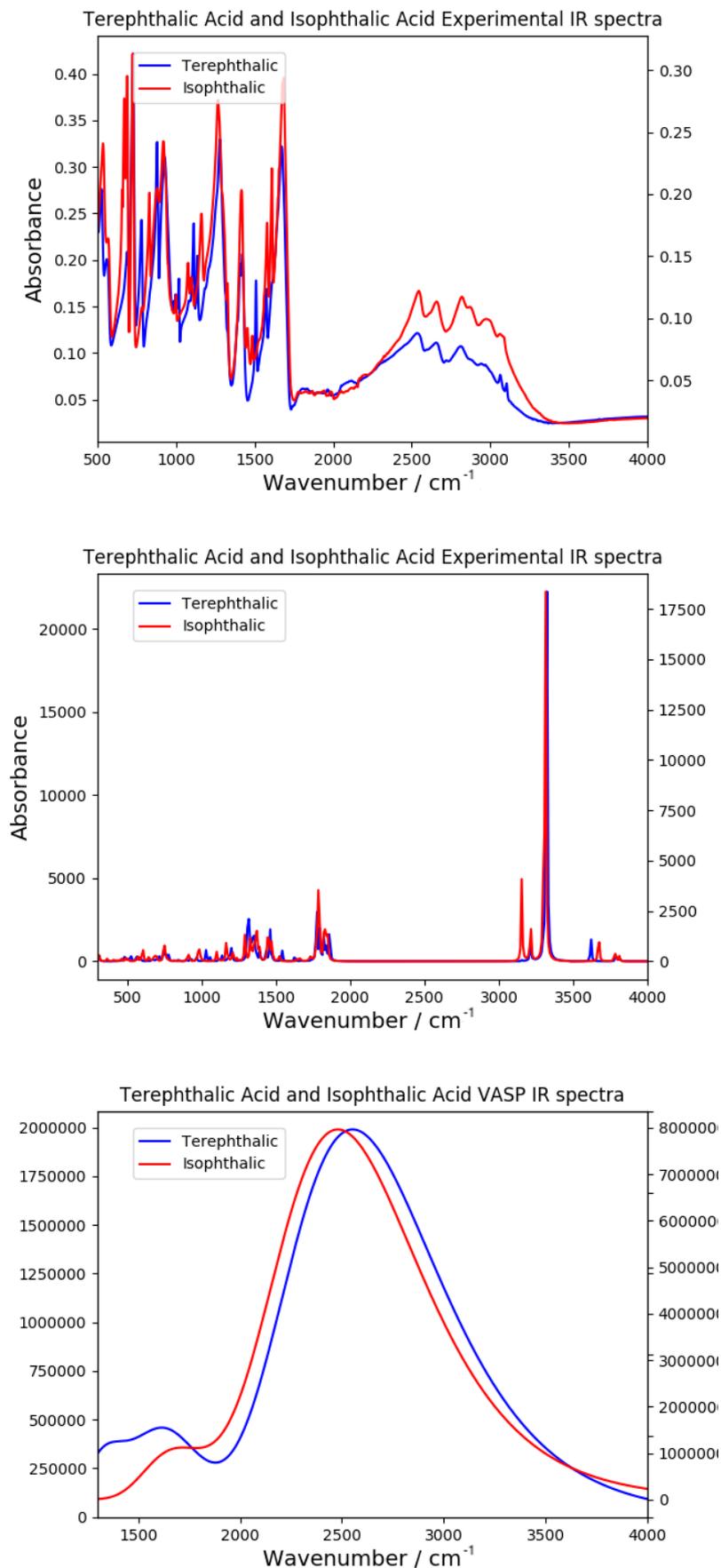
Figure 3.10: The overlaid IR spectra for terephthalic (black) and isophthalic (red) acids; top: experimental, middle: calculated gas-phase and bottom: calculated solid-state periodic.
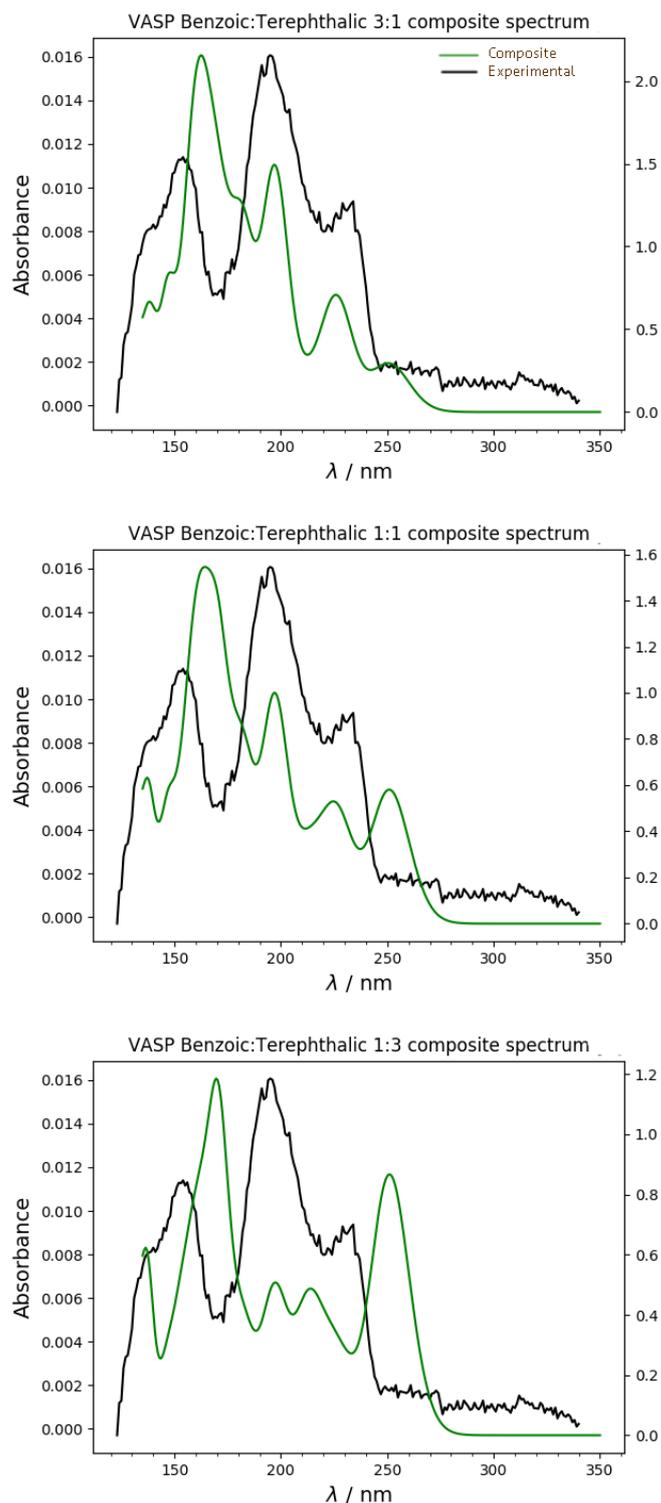
Figure 3.11: The composite spectra for top: a 3:1 mixture of benzoic acid and terephthalic acid, centre: a 1:1 mixture of benzoic acid and terephthalic acid and bottom: a 1:3 mixture of benzoic acid and terephthalic acid compared with the experimental residue spectrum (black trace).
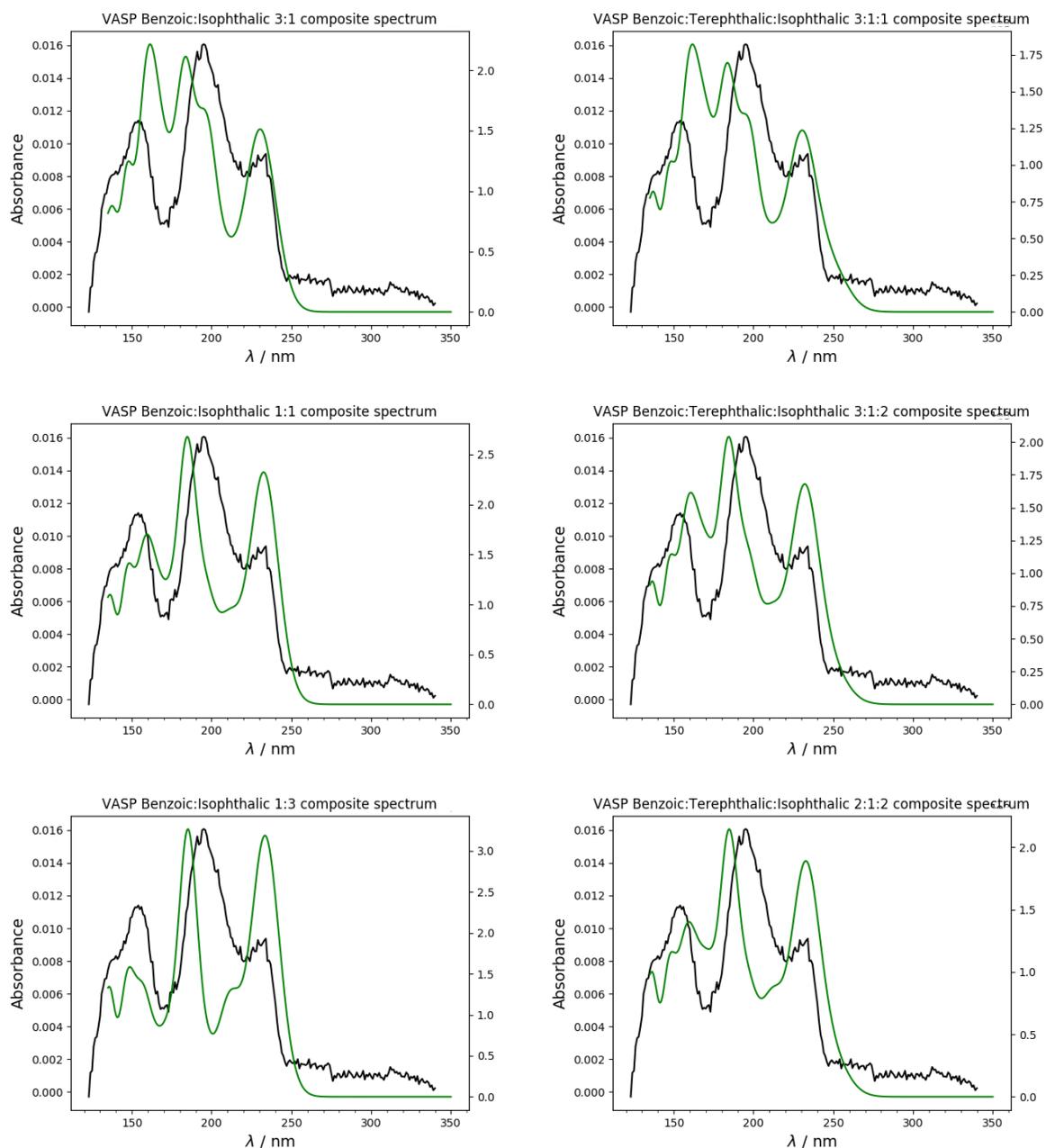
Figure 3.12: Left: the composite spectra for top: a 3:1 mixture of benzoic acid and isophthalic acid, centre: a 1:1 mixture of benzoic acid and isophthalic acid and bottom: a 1:3 mixture of benzoic acid and isophthalic acid compared with the experimental residue spectrum (black trace). Right: the composite spectra for top: a 3:1:1 mixture of benzoic acid, terephthalic acid and isophthalic acid, centre: a 3:1:2 mixture of benzoic acid, terephthalic acid and isophthalic acid and bottom: a 2:1:2 mixture of benzoic acid, terephthalic acid and isophthalic acid compared with the experimental residue spectrum (black trace).

Figure 3.12 shows the equivalent composite spectra for a mixture of benzoic acid and isophthalic acid in the same 3:1, 1:1 and 1:3 proportions compared with the residue spectrum. The shape match between these is much better, due to the lack of the large peak at 175 nm, than that of the terephthalic acid containing composite spectra. In particular the 1:1 mixture shows a great degree of overlap with the residue spectrum, although the intensity of the peak at roughly 240 nm is overestimated. Nonetheless this mixture of acids presents a much more likely residue composition than benzoic acid and terephthalic acid. This is somewhat at odds with the first addition of the carboxylic acid group proceeding at a quicker rate than the second, however, due to the CO_2 being vastly in excess compared to the benzene in the original ice (a 100:1 mix) a high degree of conversion to the diacid is not unlikely.

As previously stated it is not a simple task to distinguish between isophthalic acid and terephthalic acid and thus, there is a possibility that both species are present in the residue. In order to explore this possibility composite spectra including all three acids have been plotted (figure 3.12). 3:1:1, 3:1:2 and a 2:1:2 mixtures of benzoic acid : terephthalic acid : isophthalic acid have all been considered. The 3:1:1 mixture (top right in figure 3.12) shows the least successful shape match to the residue spectrum with the two lowest wavelength peaks blending together. As with the benzoic acid and isophthalic acid mixtures, the three acid mix with equal proportions of benzoic acid and isophthalic acid (bottom right in figure 3.12) matches the residue spectrum most closely. Including terephthalic acid in this mixture makes little visible change from the 1:1 benzoic acid to isophthalic acid spectrum (centre left figure 3.12) further suggesting that it may be present in the ice in minor amounts but would take further investigation to identify conclusively.

What this data does suggest is that the directing nature of the COOH group may be affecting the system. Clearly isophthalic acid containing ice mixes provide a better match to the experimental data than those with terephthalic acid which implies the formation of the *meta* acid is more favourable. If this is the case then there is a possibility that the reaction might not proceed through purely radical chemistry but charged species instead. Ascertaining whether this supposition is correct would require further work however, it is not obvious how this could be achieved .

3.4   Conclusions

After confirming the validity of the method chosen using IR data, and based on the experimental work of James *et al.*, it is clear by inspection of the com-

posite calculated spectra that the most likely composition of the residue from a benzene:CO_2 1:100 ice is a 1:1 mixture of benzoic acid and isophthalic (*meta*-benzene diacid) acid.[198] No benzene or CO_2 is anticipated to remain in the ice after annealing and the excess of CO_2 in the ice would lead to a large percentage of conversion to the diacid. Some singly substituted acid remains due to the first addition to the benzene ring possessing a larger rate constant than the second addition.[197]

The previously reported detection of isophthalic and terephthalic acid, in addition to benzoic acid, in such a residue is not fully supported by the UV-vis spectra calculated here. Mixtures of benzoic acid and terephthalic acid fail to reproduce the residue spectrum. Additionally, including minor quantities of terephthalic acid in a 1:1 mixture of benzoic and isophthalic acid does not visibly change the shape of the composite calculated spectrum and so some terephthalic acid may be present but only in small amounts. IR data for isophthalic acid and terephthalic acid do not differ substantially enough to conclusively distinguish between the two.

4. Irradiation of polycyclic aromatic hydrocarbons in water ice

As already outlined in the introduction it has become apparent that dust grain surface reactions play a key part in the formation of COMs in the interstellar medium.[64,65] They not only act as surface catalysts, but can also mediate the dissociation of newly formed products by serving as an energy sink. The introduction also briefly covered the make-up of these grains, with carbonaceous material being suggested as a major grain core component. This chapter examines the behaviour of polycyclic aromatic hydrocarbons, a possible grain component, when subjected to radiation, and how water ice affects that behaviour.

4.1 Polycyclic aromatic hydrocarbons and their place in the ISM

Dust is commonplace in the interstellar medium (ISM). In fact, it accounts for approximately 1% by mass of the material in interstellar dust clouds.[217] This dust is believed to be either silicaceous or carbonaceous in nature; see figure 1.3 in the introduction.[19,218] Polycyclic aromatic hydrocarbons (PAHs) are believed to be a constituent of carbonaceous dust grains and are themselves widespread in interstellar environments.[21,71,217,219,220] Steps towards understanding the formation of PAHs from acetylene in the ISM have also been made.[221,222]

Ices are as present in the ISM as dust, the latter frequently being coated in the former.[3] In particular, water ice is a primary component of the icy mantles found on dust grains.[76] PAHs have been shown to act as catalysts to water formation in interstellar space by providing a surface upon which the relevant reactions can take place.[22,23,70] Icy dust grain and PAH surface catalysis has proved to be significant in the formation of organic molecules in the ISM.[69,123,223] These range from species such as hydrogen peroxide, which can be involved in water formation, to molecules of biological relevance and importance like ribose.[65,224]

The involvement of PAHs coated by water ice mantles in the formation of COMs means that the nature and behaviour of such species in ices warrants further investigation. PAHs embedded in water ice have been shown to be easily ionised to PAH cations.[225–228] Computational studies have observed the ionisation energy of PAHs in ice to be reduced by as much as 2 eV compared to the gas-phase.[229] This reduction means that PAHs frozen in ice could be ionised by near-UV radiation as opposed to vacuum UV allowing PAH cations to be involved in chemistry taking place in regions of the ISM classified as non-ionising.

In an experimental study involving quarterrylene in water ice (in a 1:500 ratio) Gudipati and Allamandola observed, upon irradiation with 266 nm light, the growth of a peak corresponding to the positively charged quarterrylene ion (see
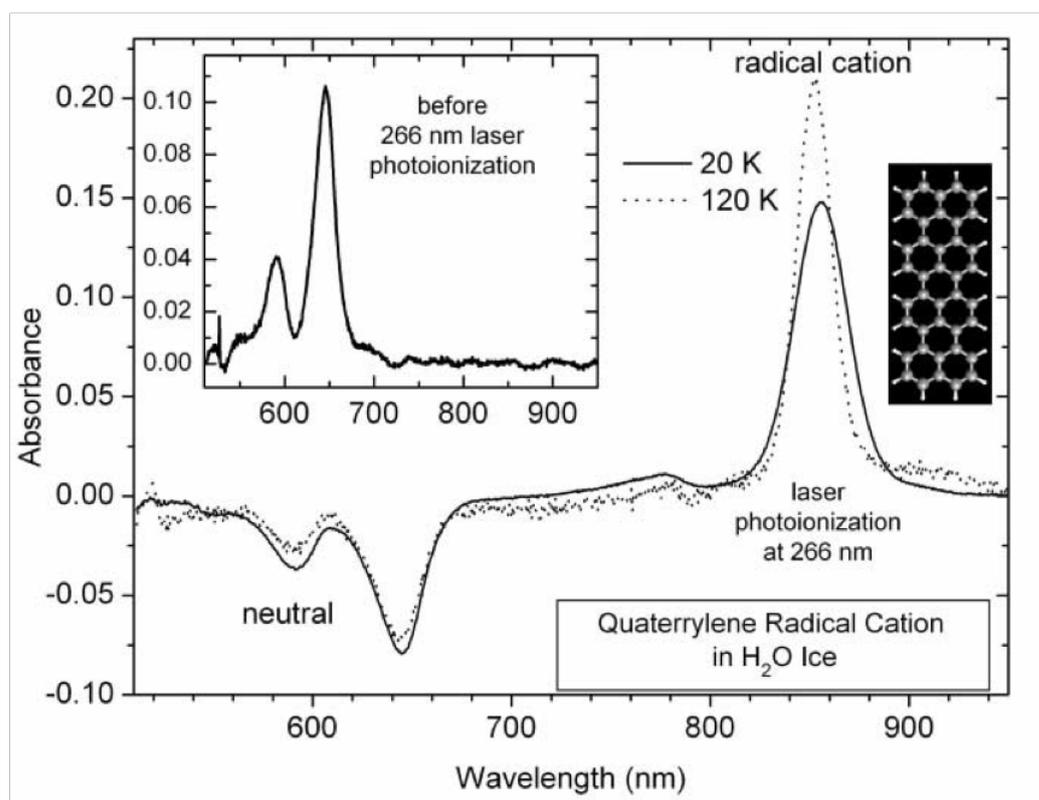
Figure 4.1: The absorption spectrum of desposited quarterrylene in water ice at 20 K showing the decrease in absorbance of the neutral species and an increase in the absorbance of the cationic species upon irradiation with 266 nm laser light for 45 minutes. Inset: The absorbance spectrum before irradiation. Reproduced with permission from Gudipati and Allamandola, Astrophys. Journal, 2006.[230]

figure 4.1).[230] This growth is also accompanied by a negative absorbance in the region of the parent neutral species which amply shows the conversion of the neutral species into the cation. Gudipati *et al.* posit that the quarterrylene occupies a 'special site' in the ice matrix that promotes ionisation at much reduced wavelengths (2.1 eV lower in energy) compared to the gas-phase. They also put forward the possibility that the lowering of the ionisation energy is due to the ability of the surrounding water ice to solvate the electron.[231,232] However, the experimental data is not able to confirm this.

In this work we have studied the behaviour of a range of PAHs (benzene, napthalene, anthracene, pyrene, perylene, coronene and quarterrylene) using time-dependent density functional theory (TD-DFT) on a model PAH/ice system. Calculations were performed both to attempt to observe the effect of water ice on the ionisation energy of the PAH and also to investigate whether there is any evidence of the water ice solvating the excited electron.

4.2 Methodology

All calculations have been performed using the GAUSSIAN 09 suite of electronic structure programmes [†] using DFT at the CAM-B3LYP level[148,149,233] and the 6-311G** basis with the D3 version of Grimme's dispersion with Becke-Johnson damping.[234,235] Additional calculations were carried out using the Maestro 2014-3 release molecular modelling software package [‡] using the AMBER and OPLS force fields.[236–238]

The ground electronic states of neutral and some ionised PAH species with and without surrounding water ice clusters were optimised. The selected number of water molecules were placed, pre-geometry optimisation, to maximise coverage of the PAH. For small clusters this was done using DFT in Gaussian 09, and second derivatives were calculated to ensure the structures were stationary points, while larger systems were optimised to a minimum using molecular modelling in Maestro 2014-3. The UV-Vis spectra for all systems were calculated using time-dependent DFT (TD-DFT) in Gaussian 09 at the CAM-B3LYP/6-311G**/GD3BJ level.[208,209] The Tamm-Dancoff approximation calculations to find the transition oscillator strengths between excited states were done at the same level of theory while the electron density calculations were done at the CAM-B3LYP/6-311G(d,p) level.[239–243] The cube files generated during the electron density calculations were then used to create density difference images to display where the electron density had moved to during the excitation.

4.3 The electronic behaviour of PAHs in ice

4.3.1 Ionisation

To ensure comparability between the experimental and computational results the gas-phase spectrum of benzene only was benchmarked against that generated experimentally. The location of the maximum in the calculated spectrum for gas-phase benzene is consistent with that determined via experiment (168 nm for the calculated spectrum and 175 nm for the experimental spectrum) and thus no scaling has been applied to the calculated spectra, see figure 4.2 [*]. Additionally, no shifts in peak positions are observed between the calculations for the isolated, gas-phase PAH and that embedded in ice which is consistent with experimental observations.[229,244] We have also calculated the spectrum for the benzene cation

---

[†]Gaussian 09, revision D.01, M. J. Frisch et al., Gaussian, Inc., Wallingford, CT, 2009

[‡]Maestro Release 2014-3: MS Jaguar, Schrödinger, LLC, New York, NY, 2014

[*]Spectrum of gas-phase benzene obtained from the NIST WebBook, see https://webbook.nist.gov/cgi/cbook.cgi?ID=71-43-2

and pyrene (neutral and cationic) as an example of a larger system. The intent behind this was to ascertain whether the calculations gave the same changes in the spectra when the cations were treated with TD-DFT. The cationic systems so show an analogous shift in the wavelength of the largest peak to that seen in the experimental work of Gudipati *et al.*, see figure 4.3, indicating that the method is treating the system correctly. A calculation for the quarterrylene cation was attempted however this proved to be too computationally intensive to be feasible.

In order to understand the mechanism behind the reported decrease in ionisation energy of PAHs in water ice a series of calculations based on Koopmans' theorem were carried out. Koopmans' theorem posits that the energy of the HOMO is equivalent to the negative of the ionisation potential, and that the optical gap is given by the HOMO-LUMO energy gap. HF calculations on benzene in increasing amounts of water were carried out to outline the behaviour of the HOMO-LUMO gap, equal to the negative of the ionisation energy, with increasing quantities of water present. The HOMO-LUMO gaps calculated however, show no change
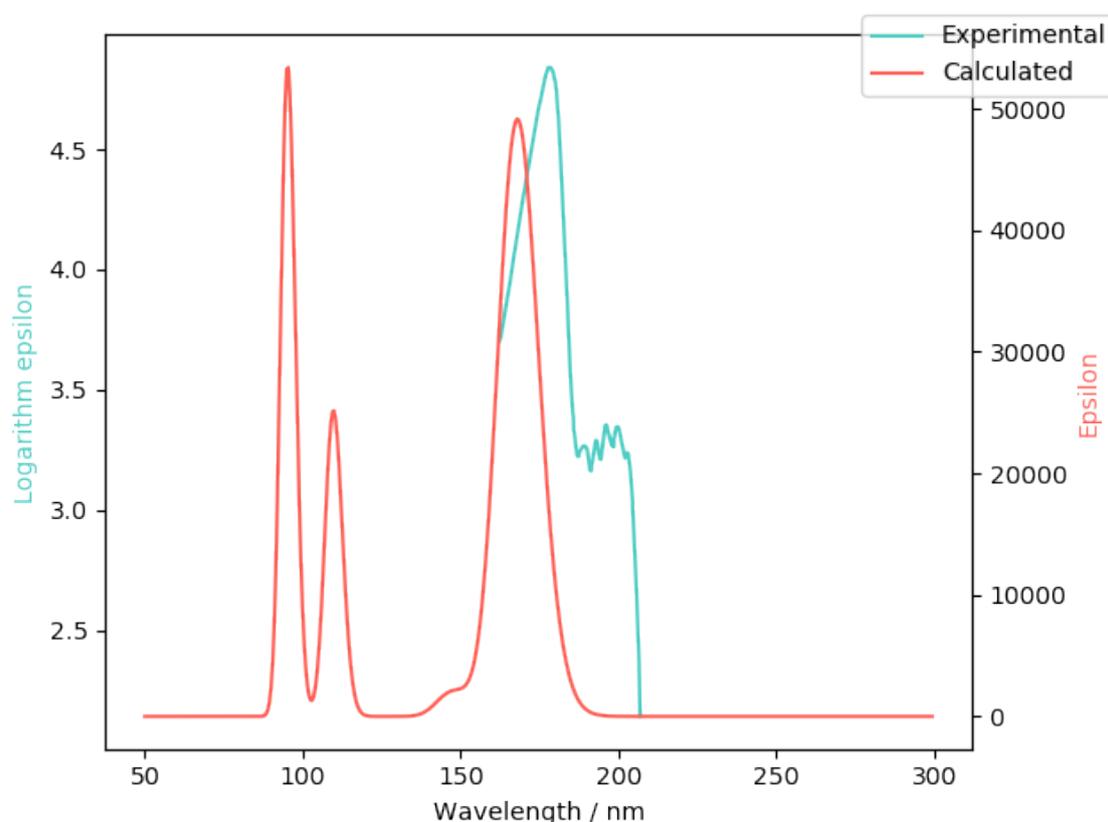
Figure 4.2: The experimental and calculated UV-Vis spectra of benzene in the gas-phase. The spectrum of gas-phase benzene was obtained from the NIST WebBook, see http://webbook.nist.gov. Epsilon here is simply the molar aborsbtion coefficient with units of $dm^{-3}mol^{-1}cm^{-1}$.
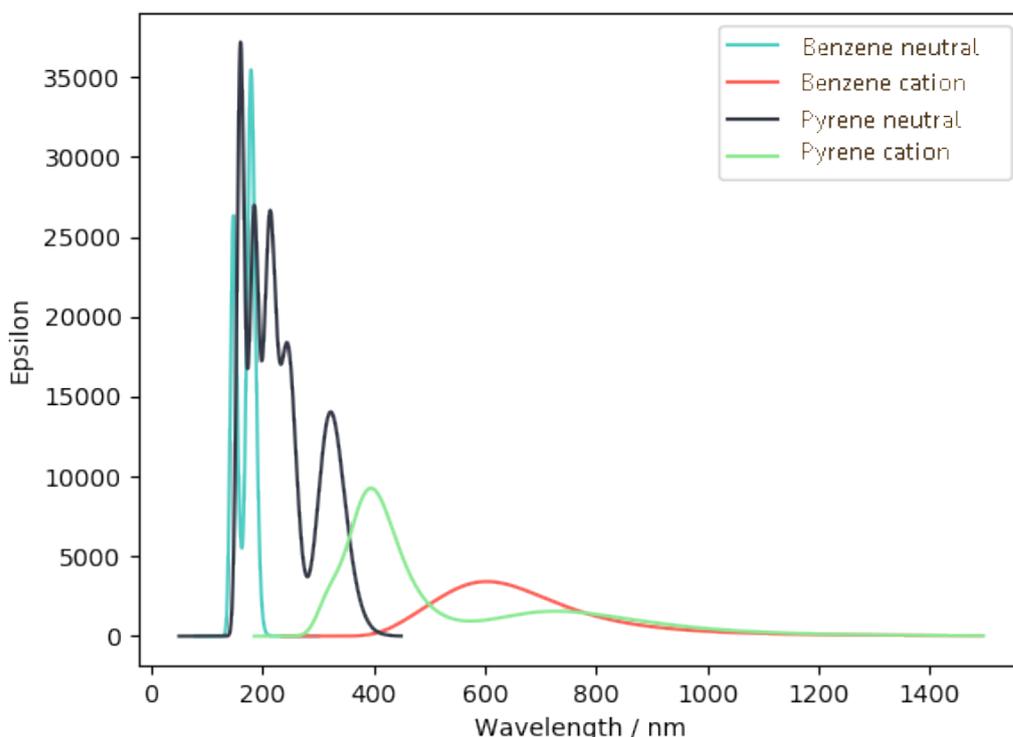
Figure 4.3: The UV-vis spectra for benzene, the benzene cation, pyrene and the pyrene cation showing the shift in the UV-Vis spectra between the neutral and the charged species in ice.

with increasing amounts of ice as shown in figure 4.4. Not only is the energy gap remaining consistent, but the energies of the HOMO and LUMO themselves show little to no change with increasing quantities of water. This demonstrates that there is, electronically, no change in the system. Thus, no ionisation can be said to be taking place.

In contrast, the UV-vis spectrum of benzene in water ice does change with increasing size of the ice as shown in Figure 4.5. The peak height at 170 nm clearly shifts to lower energies and decreases in magnitude with increasing numbers of water molecules in the ice. The inset shows the difference between the gas-phase ionisation energy and the excitation energy against the number of waters displaying how the excitation energy steadily drops away from the gas-phase value. These results suggest that rather than an ionisation process occurring there is simply an excitation as the signals in the UV-Vis spectra correspond to the promotion of an electron from a lower lying state to a higher one. It also shows that the energy input to achieve this excitation is reduced by including increasing quantities of water in the system, thus, going forward excitation energies have been used rather than ionisation energies.
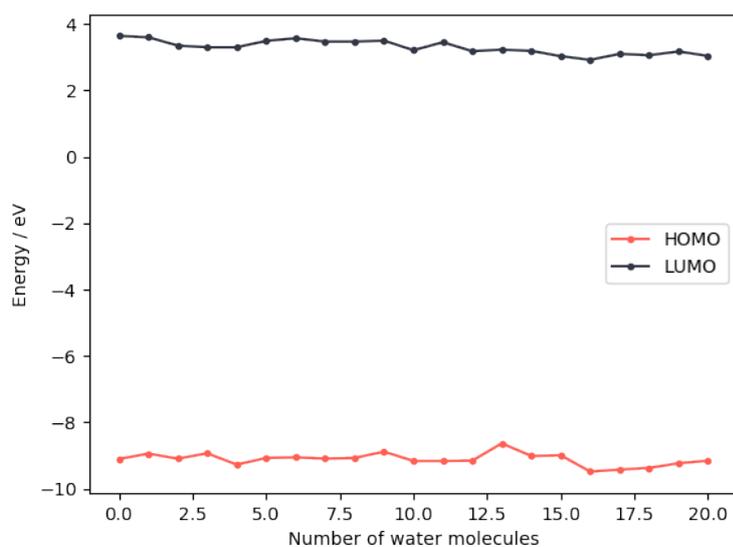
Figure 4.4: The HOMO and LUMO energies of a series of systems of benzene in increasingly large quantities of water ice showing that the ionisation energy, as approximated by Koopmans' theorem, do not appreciably change.

Park and Woon report a decrease in the ionisation of 2.1 eV (from the gas phase value of 9.24 eV) when modelling benzene in ice using the isodensity polarised continuum model (IPCM) to apply the water ice to the system.[229,245] We see an analogous decrease in the excitation energies in our calculations but to a much lesser degree, a maximum of 0.5 eV with 100 water molecules. A greater amount of water placed around the benzene does result in a larger excitation energy decrease such that the size of the decrease does appear to be slowly converging on that previously reported (see also Gudipati *et al*).[230] Were it computationally tenable to continue to a much more extensive explicit water clusters it is believed that the decrease could reach 2 eV.

Thus, there is an effect of the ice on the electronic structure of the PAH. However, it is observed in the UV-vis and not the HOMO-LUMO gap. This suggests that describing what is occurring as ionisation in the traditional sense is misleading. The electron is clearly not removed to infinity and so ionisation is incomplete. It is however, highly possible that the electron is promoted off the PAH and onto the ice to become a solvated electron.[231] This process would still lead to the generation of a charged PAH in the ice, in line with what has been observed even though it is not technically a full ionisation process. Therefore, in line with these findings, excitation energies have been used to investigate the behaviour of larger PAHs to confirm that this behaviour is observed in other, related systems.
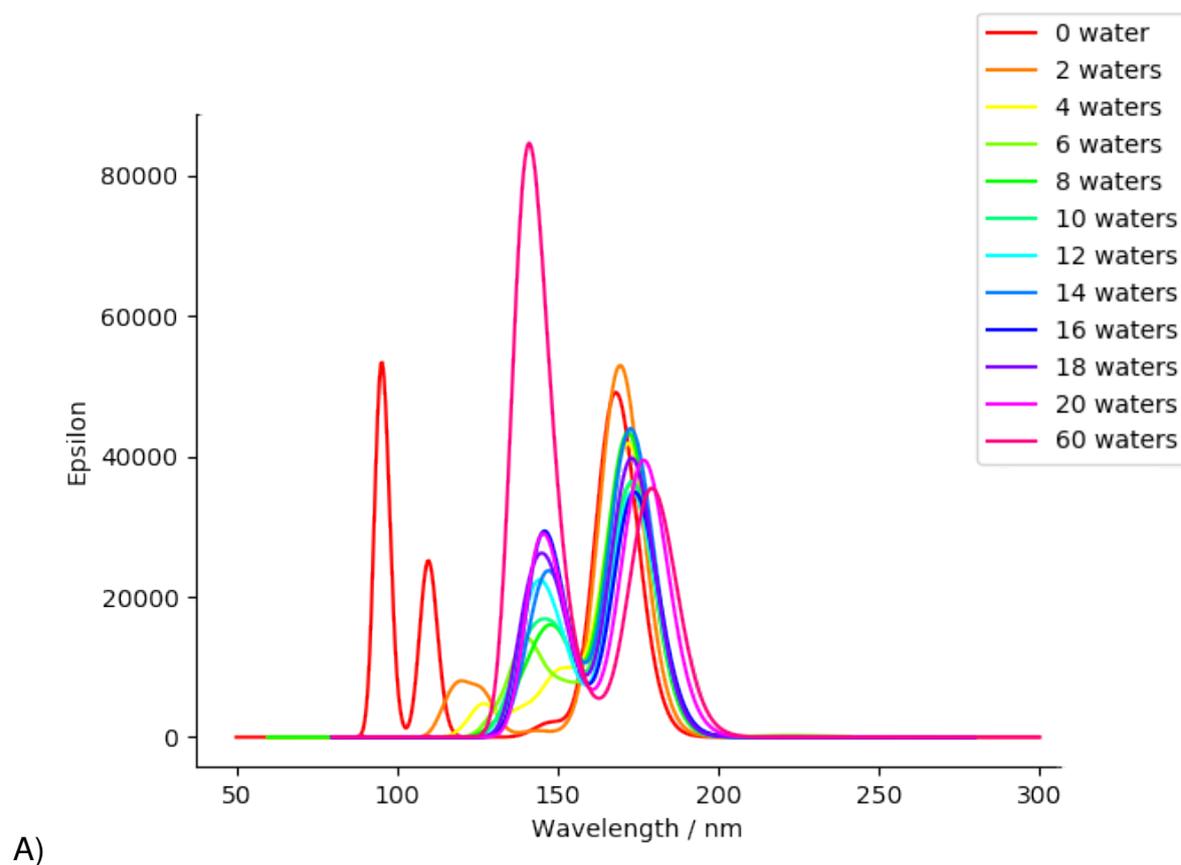
A)
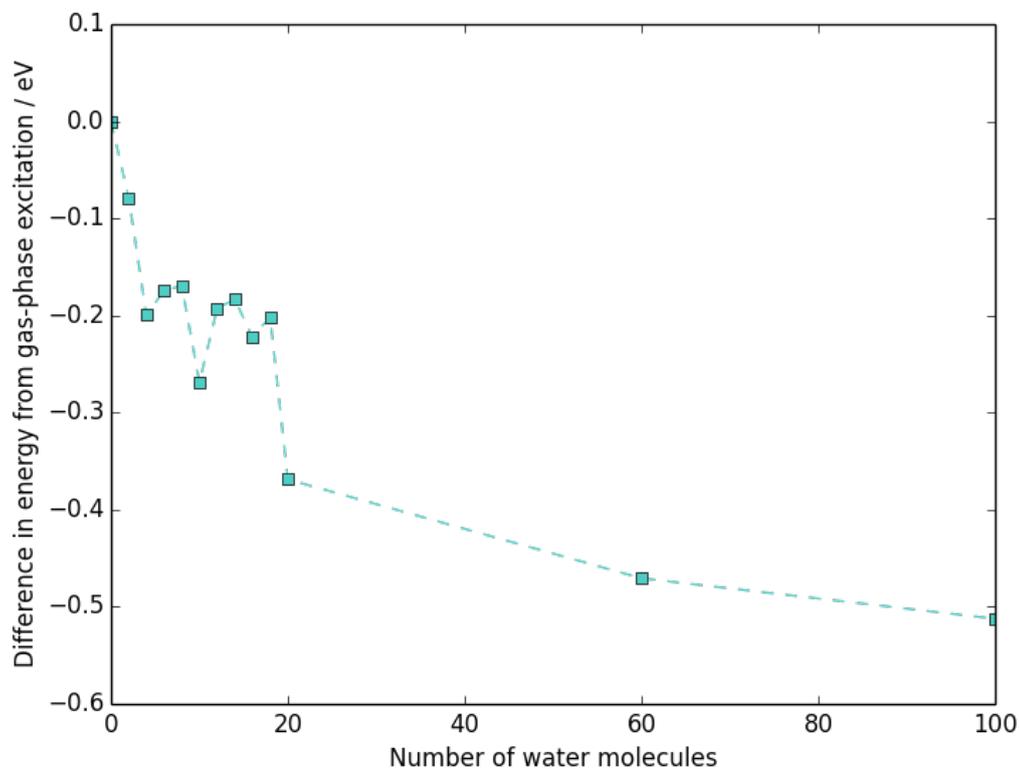
B)

Figure 4.5: A: The UV-vis spectra for benzene with increasing numbers of water showing the overall decrease in the size of the excitation at 170 nm as the amount of water increases. B: the magnitude of the decrease in the excitation energy of the maximum absorption plotted against the amount of water.

4.3.2 The effect of water ice on excitation energies

As can be seen in figure 4.5 the decrease in excitation energy does not decrease uniformly with the addition of more and more water ice. Moreover, it was found that it does have some small dependence on the exact position of the water molecules around the benzene. To investigate this further, five different starting geometries for benzene surrounded by twenty water molecules were optimised. Each of these resulted in a different energy minimum. These optimised benzene-water clusters were then treated with TD-DFT to yield the UV-vis spectra to obtain the excitation energies. The greatest variation in excitation energy maximum between structures was 0.1 eV as can be seen in figure 4.6. It is therefore possible that small structural discrepancies could cause the discontinuities within the
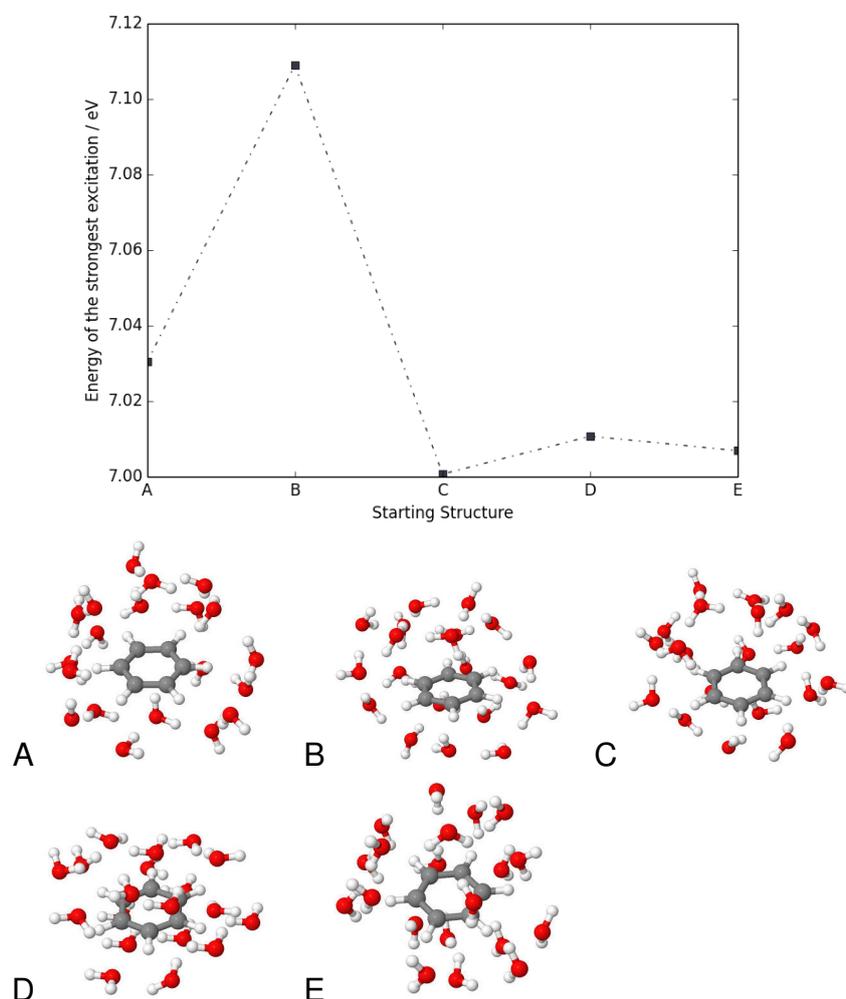
Figure 4.6: How the excitation energy of benzene surrounded by 20 water molecules changes depending on the arrangement of water around the benzene where A, B, C, D and E are the different starting structures used.

data. However, they are not large enough to invalidate any conclusions that can be drawn.

As it is believed that the water ice is solvating the excited electron, the amount of coverage of the PAH by the ice could affect the decrease in the excitation energy. This stems from the belief that a greater degree of coverage by the ice allows for more opportunity for interactions between the ice and the PAH species. This effect is indeed seen in figure 4.5 where larger amounts of water offering better coverage leads to a lower excitation energy, it is also worth noting that the degree of coverage has only a very small effect of the position of the peaks (a maximum movement of only 6.5 eV to lower energy). However, given the organic nature of the PAHs they mix poorly with a polar solvent. This can lead to some difficulties in maintaining coverage over the PAH regardless of the quantity of water used. This separating of the two can be seen in figure 4.7, in which benzene has moved out of a hexagonal water ice structure. As the crystalline ice structure did not retain the PAH molecule and as the fluctuations caused by different starting structures are only small, amorphous ice structures have been used as these provided better coverage. Additionally, in order to probe the maximum extent to which the ice reduces the excitation energy the UV-Vis spectrum of a system was only calculated if the PAH was well covered by the ice as poorly covered ices would give a higher excitation energy.

The excitation energy for benzene in the gas-phase determined here is lower than that expected, benchmark data for the ionisation energy of benzene and other PAHs can be found in Deleuze *et al.*[246] This is partly a result of the method utilised, TD-DFT is known to underestimate excitation energies to a degree, however not to such an extent that the excitation energies provide no useful information.[247] Addition of the zero point energy correction causes the energy to be over-estimated and as it has no effect on the trends being investigated has not been applied in this case. It is also a factor of using the excitation energy as opposed to the ionisation energy. As the electron is not fully being removed the electron the energy input is likely to be lower and so these results are consistent with this.

This decrease in excitation energy is not only observed for benzene. Napthalene, anthracene, pyrene, perylene, coronene and quarterrylene all display the same trend of the excitation energy decreasing when placed in ice although the decrease in energy is smaller than that of benzene (see figure 4.8). The gas-phase reference ionisation energies for napthalene, anthracene, pyrene, perylene and coronene were taken from the NIST WebBook and the references therein. Ionisation energies have been used as the reference, despite the excitation energies being those under investigation, for ease of comparison as existing work in
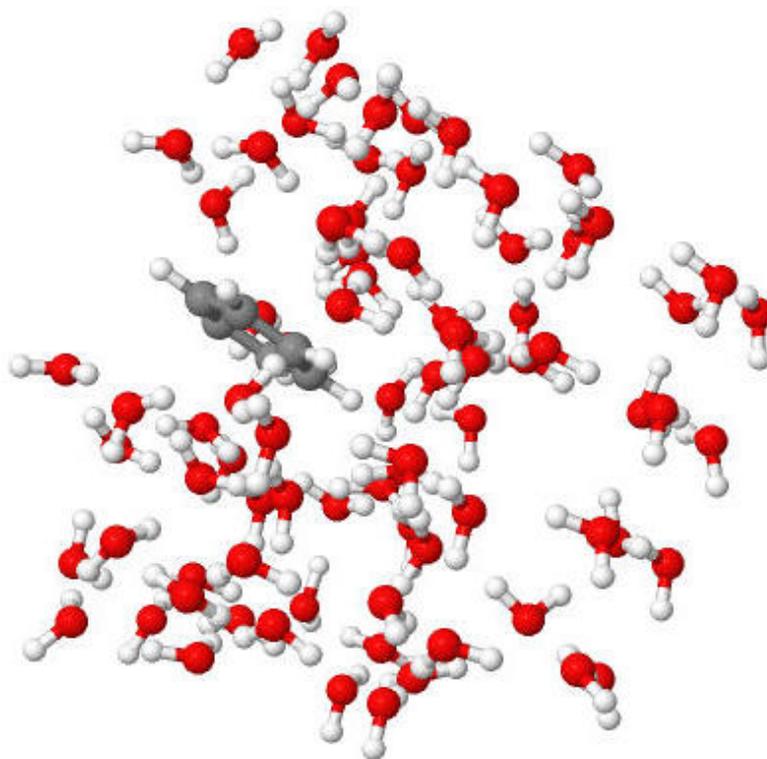
Figure 4.7: The optimised structure of benzene in 80 water molecules where the benzene has migrated partially out of the ice structure.

the literature is based on ionisation energies.[248–251] The reference ionisation energies for benzene and quarterrylene are taken from Woon and Park and Gudipati and the relevant references contained within respectively.[229,230]

Again, the observed decrease in excitation energy is consistent with the calculations of Woon and Park who report progressively smaller decreases in ionisation energy for larger PAHs (their work includes napthalene, anthracene and pyrene in addition to benzene). As with benzene, the decreases seen here using explicit water molecules are not as significant as the 1.5 eV reported in the literature and this is likely a feature of the limits on the amount of water that can possibly be included in the calculation. ONIOM (our Own N-layered Integrated molecular Orbital Mechanics[180,252]) method calculations were attempted in an effort to increase the scope of the water cluster. In these calculations the PAH was treated with quantum mechanical methods while the water was in a molecular mechanics layer. However, this proved not to be the most efficient way to perform calculations on these systems. Any TD-DFT calculation performed on these systems gave results as if the PAH was in the gas-phase and it became apparent that, for generating the UV-Vis spectra, the entire system needed to be treated quantum mechanically. The most effective method proved to be optimising the PAH
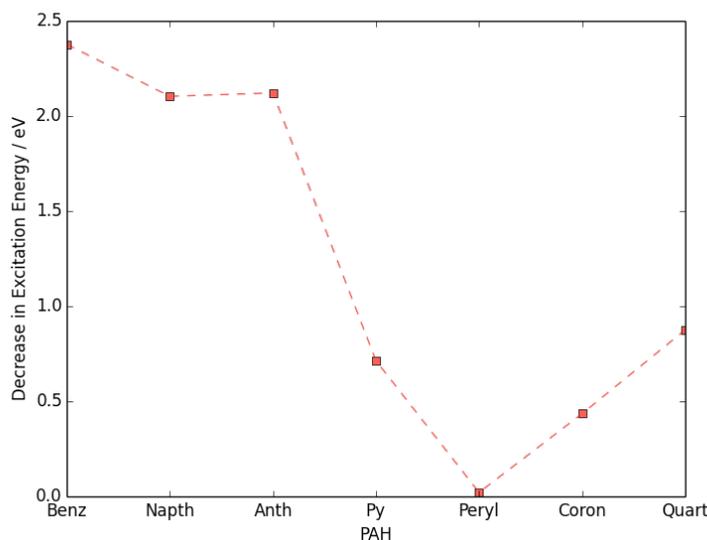
Figure 4.8: The maximum decrease in the excitation energy compared to the literature gas-phase ionisation energy observed in a series of PAHs (benzene (Benz), napthalene (Napth), anthracene (Anth), pyrene (Pyr), perylene (Peryl), coronene (Coron) and quarterrylene (Quart)) in water ice compared to the gas-phase.

and water cluster using molecular mechanics and then performing the TD-DFT portion of the calculations quantum mechanically.

This trend, that the decrease seen in the excitation energy becomes smaller as the size of the PAH increases, reverses for coronene and quarterrylene. Here the decrease is between 0.8 and 0.9 eV, inverting the trend in the size of the excitation energy prior to this point. Gudipati *et al* posit that the decreasing trend in ionisation energy is linked to the induced polarization of the PAH, with smaller PAHs having larger induced polarizations leading to greater decreases in ionisation energy.[230] Using equation 4.1 which has been taken from Gudipati *et al.*, the ionisation energy of the PAH in the ice is simply the difference between the gas-phase ionisation and the sum of the polarisation (P_+) and the electron affinity.[230]

$$IE_{PAH-Ice} = IE_{PAH-Gas} - (1 - \frac{1}{\epsilon_\infty})\frac{e^2}{2R} - EA_{Ice} \qquad (4.1)$$

Here, $IE_{PAH-Ice}$ is the ionisation energy of the PAH in the ice, while $IE_{PAH-Gas}$ is the ionisation energy of the PAH in the gas-phase and EA_{Ice} is the electron affinity of the ice. The second term corresponds to P_+ where ϵ_∞ is the optical dielectric constant of the medium, e is the electron charge and R is the effective radius of the ion. Given that the polarisation depends on the effective radius of the ion it is possible that PAHs that strongly deviate from a circular shape, perylene and quarterrylene for example, may display results that do not align as

well as other PAHs to the model. It is also possible that, as we have found that it is the excitation energy rather than the ionisation energy that changes that this model's predictions do not truly apply to the system. In that case then there is no prediction that suggests a steady decrease in the lowering of the excitation energy is expected.

Equally, as previously stated, there is some geometry dependence on the size of the decrease and as exhaustive geometry sampling has not been carried out in this particular instance it is highly possible that other cluster arrangements could yield larger decreases. What is clear, though, is that in all cases the ice causes the excitation energy to drop below that of the gas-phase and so the facility of cation generation would increase. The optimised geometries used for all systems can be seen in figure 4.9.

4.3.3 Electron density

In order to ascertain if it is the ice surrounding the PAHs that allows for the decrease in excitation energy, electron density calculations were undertaken and the density differences between the ground and excited states calculated. It was thought that the lost electron is solvated by the ice, making it easier to remove from the PAH system.[231] In all instances electron density is observed on the water in a number of the excited states for each PAH. In all cases the excited state that corresponds to the strongest transition in the UV-vis spectrum also shows electron density on the surrounding water (see figure 4.10 and 4.11). This heavily implies that the ease of ionisation seen in PAHs in interstellar ice analogues is linked to the ability of the water molecules to accept the excited electron from the PAH.

In addition to investigating where the electron density lies, the transition strengths between the excited states for benzene have been calculated. This was done to ascertain whether excitations into higher states would decay down into the states that were found to have electron density on the water ice. In the case for benzene, several higher states have transition oscillator strengths greater than 0.1 between them and the excited states with electron density on the water, see table 4.1. For benzene the states that correspond to the strongest UV-Vis transition are excited states 3 and 4. Table 4.1 shows that a variety of higher lying states have strong decay pathways down into excited states 3 and 4. Examining the electron density of these states (top row in figure 4.10) shows both of these states to have electron density on the water ice. This implies that, should an electron from the PAH be excited into one of these higher lying states it would decay back down into the state where the electron density lies on the water. It was hoped that the transition
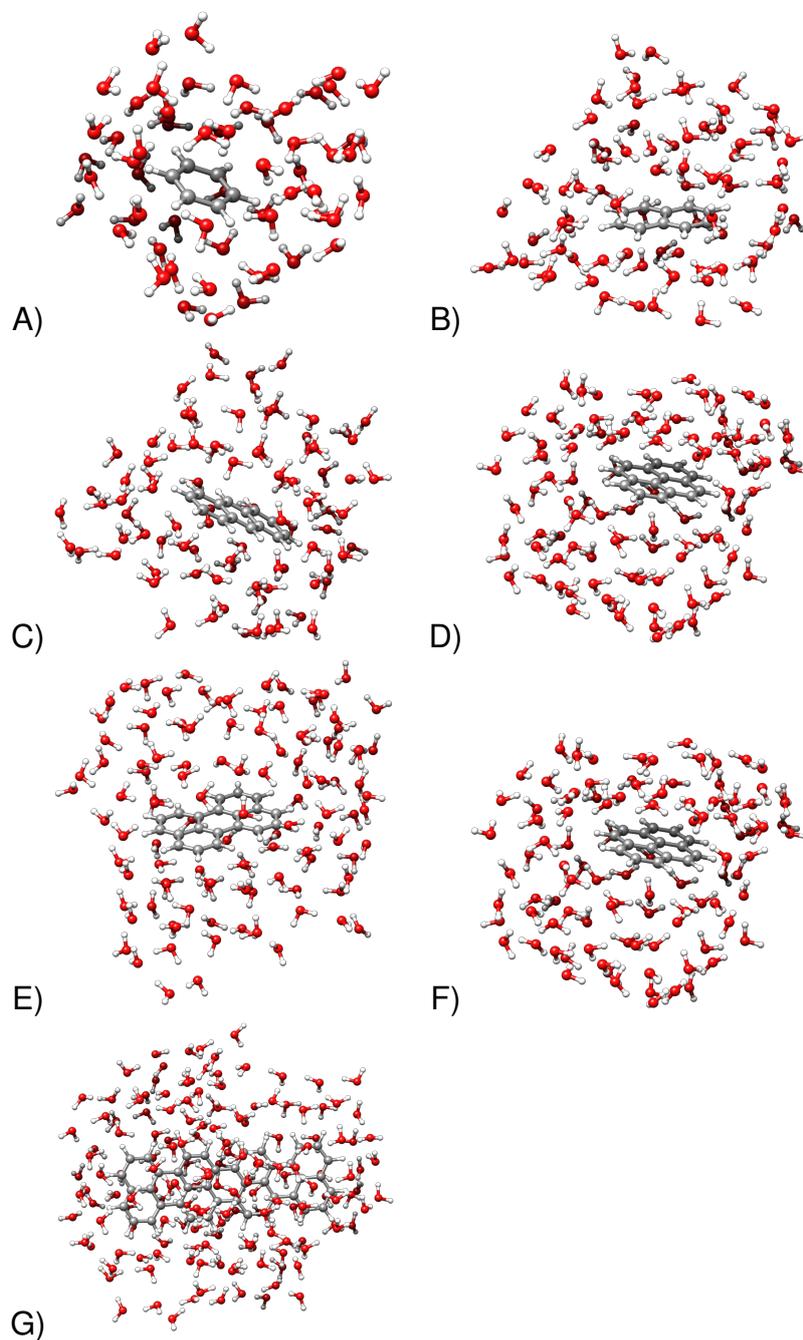
Figure 4.9: The optimised structures of the PAHs (A: benzene, B: napthalene, C: antracene, D: pyrene, E: perylene, F: coronene and G: quarterrylene) in water ice used to generate the UV-Vis spectra from which the decrease in the excitation energy was ascertained.
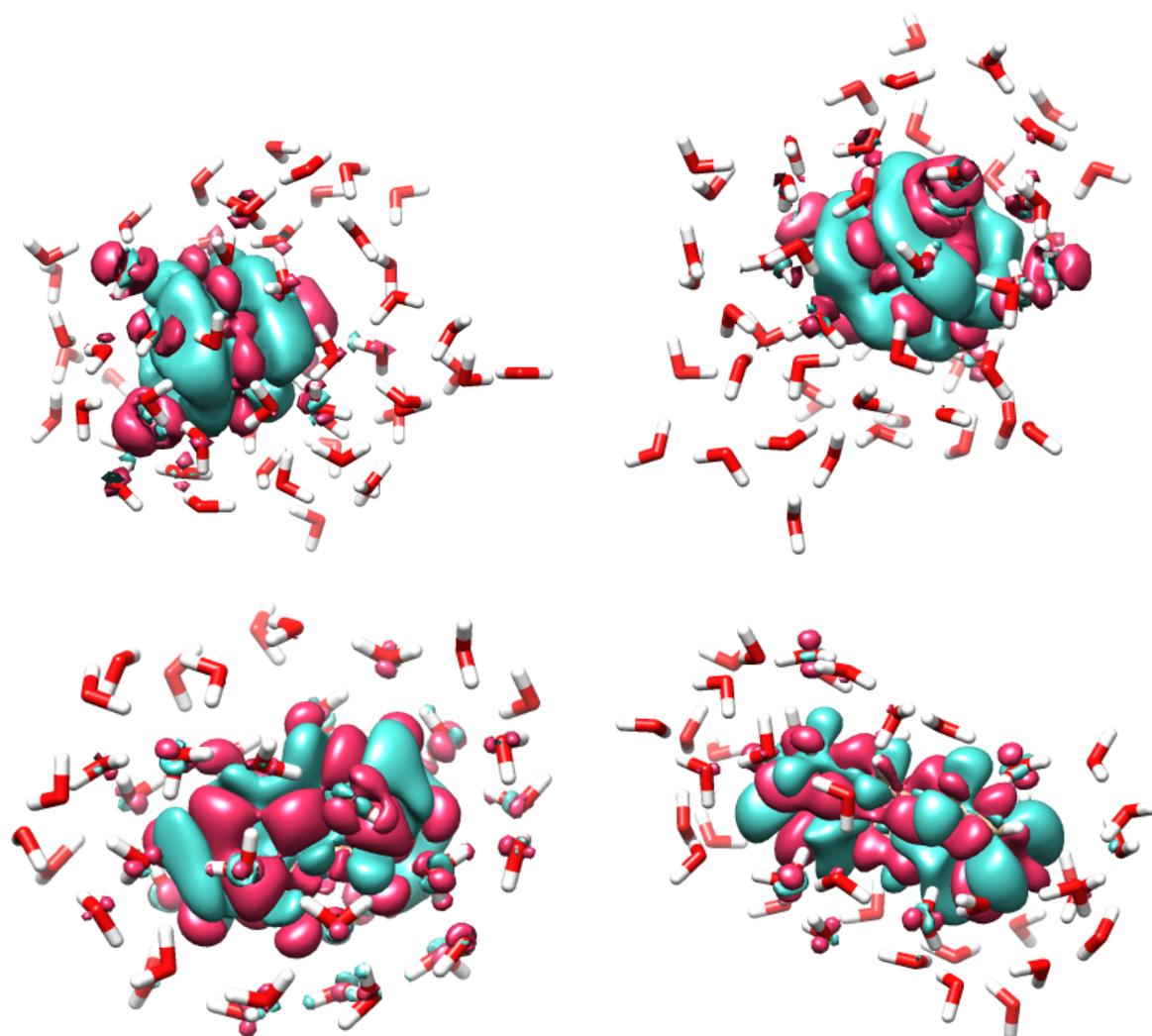
Figure 4.10: The electron density difference between the ground state and the excited state corresponding to the strongest transition for various PAHs in water ice (top row: both benzene; bottom row: napthalene and anthracene) in the excited state that corresponds to the strongest transition in the calculated UV-vis spectrum. Red areas correspond to a gain in electron density and blue areas to a loss of electron density.
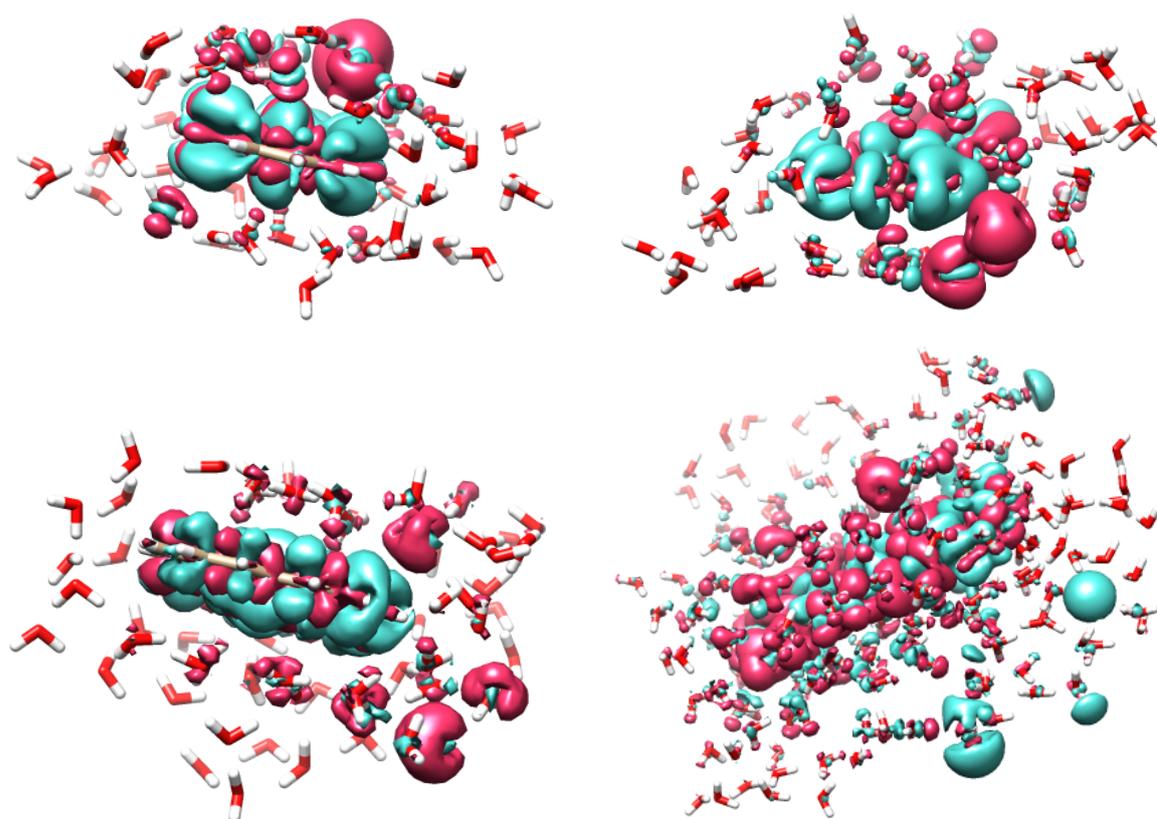
Figure 4.11: The electron density for various PAHs in water ice (top row: pyrene and perylene; bottom row: coronene and quarterrylene) in the excited state that corresponds to the strongest transition in the calculated UV-vis spectrum. Red areas correspond to a gain in electron density and blue areas to a loss of electron density.

Table 4.1: All the calculated electronic transition oscillator strengths between the excited states of benzene or pyrene in water ice with a value greater than 0.1.

Benzene			Pyrene		
State i	State j	Oscillator strength	State i	State j	Oscillator strength
7	3	0.3	10	2	0.4
8	3	0.4	12	2	0.2
9	3	0.1	13	2	0.1
9	4	0.3	16	7	0.1
10	4	0.4	18	7	0.1
16	3	0.3	19	1	0.1
18	4	0.3	22	7	0.2

strengths between the excited states of quarterrylene could be calculated but this calculation proved to be too computationally expensive to be tenable. Therefore, in order to examine the behaviour of a larger system, pyrene was chosen. Here, the excited state that corresponds to the strongest UV-Vis transition is excited state 21 which, like benzene, shows electron density on the water ice. However, due to this being a much higher energy excited state, there are no higher lying states that relax down into it. Nonetheless, this state does not have any strong transition oscillator values into any lower energy excited states, suggesting that the electron, once excited, would remain in place. Overall, the data firmly suggests that the electron can easily be excited away from the PAH system and onto the surrounding ice where it would remain. This then effectively creates a PAH cation within the ice.

4.4   Conclusions

It has been shown that the previously reported decrease in the ionisation energy of PAHs in water ice is potentially due to the ability of the water ice to accept the excited electron and solvate it away from the PAH.[229,253] All of the PAHs tested here, benzene, napthalene, anthracene, pyrene, perylene, coronene and quarterrylene showed a decrease in the strongest excitation in the calculated TD-DFT spectrum compared to the gas-phase ionisation energy. The largest decrease was seen in benzene, 2.1 eV, and the smallest in perylene, 0.1 eV. The extent of the decrease in excitation energy is dependent on the coverage of the PAH by the ice. Greater coverages by larger quantities of water ice lead to larger decreases in the excitation energy.

While the trend in the decrease in the excitation energy displays a downward slope with larger PAHs showing a smaller decrease than smaller ones, it is not smooth. This is thought to be partly due to the small geometry dependence seen in the size of the excitation energy decrease. It may also result from the shape of the PAH itself, with those being the least circular showing smaller decreases than expected when compared with the polarisation based model proposed in Gudipati *et al* 2006.[230]

Electron density calculations show that all of the PAHs in ice investigated in this paper have excited states with electron density evident on the ice. These states correspond to those contributing to the strongest excitation found in the gas-phase spectrum for the PAH. In addition to this the transition oscillator strengths for benzene and pyrene were calculated which suggested that once the electron was excited into the ice it was unlikely to decay back to a state with the electron returned to the PAH.

5. Methyl formate and glycolaldehyde formation

As has been previously discussed a wide range of organic molecules can be formed on dust grain surfaces in the ISM. In this chapter the focus falls on the isomers methyl formate and glycolaldehyde whose observed abundances, as mentioned in the introduction, have yet to be fully replicated by simulations.

5.1 Molecules in the ISM

Methyl formate (in both its *cis-* and *trans-* forms) and its structural isomers glycolaldehyde and acetic acid (see figure 5.1) have been established to exist in the interstellar medium (ISM) for some years now. Methyl formate was the first to be observed towards Sagittarius B2 in 1975 by Brown *et al.*[254] This was followed by the detections of glycolaldehyde and acetic acid in 1997 and 2000 respectively (both also detected towards Sagittarius B2).[125,255] Of the three isomers, methyl formate is the most abundant, then acetic acid, with glycolaldehyde being the least abundant. The abundance ratios for the isomers were found to be 26:1:0.5.[256] Formation routes for each of these species have been included in various chemical networks. However, the resulting abundances have yet to fully replicate those found in the ISM. The fractional abundance with respect to H_2 of methyl formate found in hot cores is of the order of magnitude 10^{-8} whereas models only yield an abundance on the order of 10^{-9}.[68,69]
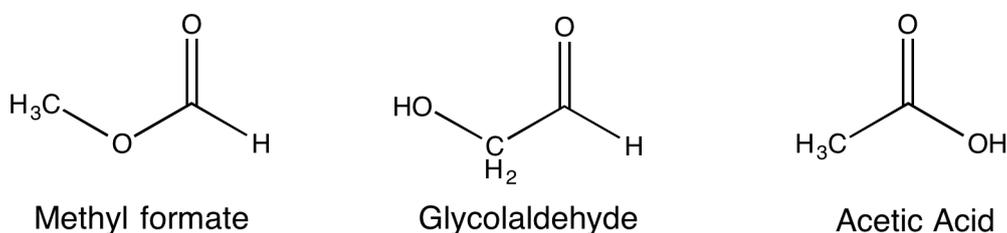
Figure 5.1: The $C_2H_4O_2$ isomers methyl formate, glycolaldehyde and acetic acid.

After its successful detection, methyl formate was initially thought to be formed in the gas-phase. This formation route involved H_2CO and MeOH, both of which are formed on dust grain surfaces.[122,257] A subsequent dissociative recombination, as shown in reactions 5.1 and 5.2, yields the desired product. There are alternative channels, but these were found to be less efficient than the one shown.[123]

$$[CH_3OH_2]^+ + H_2CO \longrightarrow [HC(OH)OCH_3]^+ + H_2 \tag{5.1}$$

$$[HC(OH)OCH_3]^+ + e^- \longrightarrow HCOOCH_3 + H \tag{5.2}$$
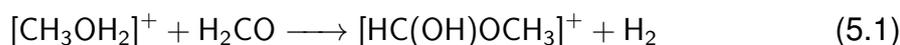
If one makes the assumption that these processes are efficient, then significant amounts of methyl formate can be produced in a 10^4 - 10^5 year time span.[258]

However, these proposed gas-phase reactions were later modelled computationally and it was found that reaction 5.1 has an energy barrier of 128 kJ mol^{-1} which is insurmountable under astrochemical conditions.[123] Alternative gas-phase formation routes for methyl formate have been explored to fully explore whether or not forming methyl formate under these conditions is possible. Table 5.1 shows those routes that were investigated.[123] However, these reactions were not found to be viable under astrochemical conditions either due to the presence of an energy barrier (reactions 1-3) or by requiring three-body reactions (reactions 4-5).[123] In these latter cases, the third-body is required to radiate away any energy generated during the reaction to prevent the product decomposing immediately back to the reactants.

Table 5.1: Alternative reactions for the formation of methyl formate as proposed by Horn *et al.*[123]

Reaction Pathway	
$H_2CO + [H_2COH]^+ \longrightarrow [HC(OH)OCH_3]^+ + h\nu$	(1)
$[CH_3OH_2]^+ + CO \longrightarrow [HC(OH)OCH_3]^+ + h\nu$	(2)
$[H_2COCH_2OH]^+ + H_2CO \longrightarrow [HC(OH)OCH_3]^+ + H_2CO$	(3)
$CH_3OH + [HCO]^+ \longrightarrow [HC(OH)OCH_3]^+ + h\nu$	(4)
$CH_3^+ + HCOOH \longrightarrow [HC(OH)OCH_3]^+ + h\nu$	(5)

Thus far all the mentioned reactions have led to *cis*-methyl formate as the product. *Trans*-methyl formate (where the carbonyl group and the CH_3 group are *trans* to one another) is the higher energy conformer. However, it may be that there are formation routes to it which are more energetically accesible than those leading to the *cis*-isomer. Thus, two further reactions for the formation of methyl formate were investigated, (5.3) and (5.4), which were used to generate both isomers of methyl formate.[63]

$$CH_3OH + [HC(OH)_2]^+ \longrightarrow [HC(OH)OCH_3]^+ + H_2O \qquad (5.3)$$

$$HCOOH + [CH_3OH_2]^+ \longrightarrow [HC(OH)OCH_3]^+ + H_2O \qquad (5.4)$$
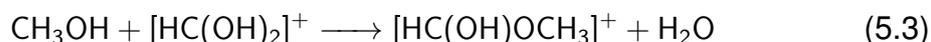
Reaction 5.3 was found to have a barrier to the formation of both *cis*- and *trans*-methyl formate (17 and 21 kJ mol^{-1} respectively), while reaction 5.4 produces *trans*-methyl formate *via* a barrierless reaction.[63] Given the presence of an energy barrier to interconversion between *trans* and *cis* of about 35 kJ mol^{-1}, this indicates that *trans*-methyl formate should be present to some extent in interstellar environments. When these new routes for the production of methyl formate were added to existing chemical networks, the results did indeed suggest that *trans*-methyl formate should be an abundant molecule in hot cores.[259] A tentative

detection of *trans*-methyl formate in the Sagittarius B2(N) molecular cloud has been made based on a laboratory rotational spectrum.[260]

However, even with the inclusion of formation routes for both isomers, the observed abundance of methyl formate cannot be replicated by gas-phase chemistry alone, with the abundance of *cis*-methyl formate being 3 orders of magnitude too low and the abundance of *trans*-methyl formate not sufficient to close the gap.[259] Therefore, it has become increasingly clear that gas-phase reactions in the ISM cannot produce methyl formate in the amounts that have been measured. Equally, the abundances of glycolaldehyde and acetic acid, while qualitatively in agreement, do not match those observed.[69] Consequently, proposed formation routes for methyl formate have shifted to be those taking place on the surfaces of icy dust grains as ices contain potential building blocks for methyl formate and its isomers.[81,261–264] In particular, interstellar ices are typically composed of H_2O, CO and CO_2, with MeOH, CH_4, H_2CO and NH_3 having also been observed.[69] These molecules can photodissociate to produce radicals which lead to further, more complex, moieties.

This chapter aims to use computational means to establish whether the radicals HCO, OCH_3 and $HOCH_2$, generated from ice mantle species, can be used to make methyl formate and glycolaldehyde. Both gas-phase and ice surface reactions describing the dissociation of methanol to give $HOCH_2$ and OCH_3 were carried out followed by reactions between HCO and both dissociation products of methanol to investigate the formation of methyl formate and glycolaldehyde from them. The effect of the presence of water ice molecules on these radical-radical reactions was also tested.

5.1.1 HCO and $HOCH_2$

In order to explore the formation of methyl formate, and to a lesser extent glycolaldehyde, an initial selection of radicals to use was made. This was done by simply examining the structure of methyl formate and glycolaldehyde and cleaving along a suitable bond. The natural choice for methyl formate is along the central CO bond and the analogous CC bond for glycolaldehyde (see figure 5.2). As can be seen in figure 5.2 the resulting radical pairs are HCO plus either OCH_3 or $HOCH_2$ depending on whether methyl formate or glycolaldehyde is the target. Both OCH_3 and $HOCH_2$ can be generated *via* dissociation of a hydrogen from methanol in the ISM.[265] However, given the conditions of the ISM and the relative strengths of these bonds, with a CH bond being easier to dissociate than OH by 91 kJ mol^{-1}, then $HOCH_2$ would be more likely to form in interstellar environments.[266] The reaction of HCO and $HOCH_2$ to give both isomeric products
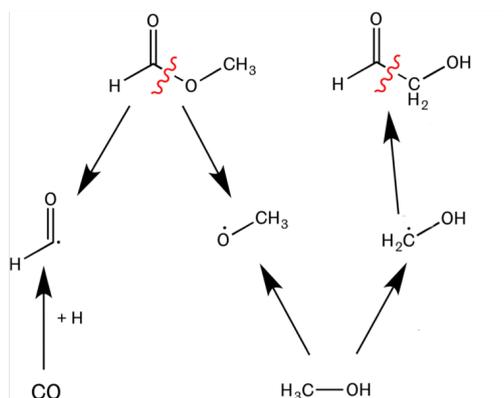
Figure 5.2: The bonds in methyl formate (left) and glycolaldehyde (right) which can be cleaved to yield a suitable radical pair.

(methyl formate and glycolaldehyde) will also be probed to discover if competing reactions might lead to the difference in abundance between the two isomers.

5.2 Methodology

Initial calculations were performed using the GAUSSIAN 09 suite of electronic structure programmes[*] using DFT at the B3LYP level[148,149,233] using a 6-311G** basis. These were followed by calculations carried out using the MOLPRO suite of *ab initio* programmes[†].

The ground state structures of methyl formate, glycolaldehyde and methanol were optimised in GAUSSIAN 09 using DFT at the uB3LYP/6-311G** level and these optimised structures were then used to construct z-matrices for scan type calculations. The scans were set up such that the $HOCH_2$ and HCO radicals were oriented to form either of the target products (mimicking the internal angles and dihedral angels found in either methyl formate or glycolaldehyde) and placed at an initial distance of 5 Å. The inter-radical distance was decreased in steps of 0.25 Å, with the structures being re-optimised, except for the inter-radical distance, at each step. The angles and dihedral angels of the carbon and oxygen backbone of the structure were held fixed at the values found in the optimised product structure to maintain structure of the radicals during the approach. This scan procedure was then repeated with two water molecules added to the system to simulate a partial solvation shell.

---

[*]GAUSSIAN 09, revision D.01, M. J. Frisch et al., Gaussian, Inc., Wallingford, CT, 2009

[†]MOLPRO, version 2015.1, a package of ab initio programs, H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schtz, and others , see http://www.molpro.net.

Similar scan type calculations were also performed in MOLPRO using the MC-SCF/RASSCF programmes and the cc-pVDZ basis set of Dunning and coworkers to afford a more accurate portrayal of the radical nature of the system (see appendix A1 for an example input).[267–270] For these calculations the structure of methanol was used as the starting geometry. The bond of the dissociating hydrogen was increased in 0.05 Å steps and at each step the remainder of the structure was optimised at the HF level. This procedure was then repeated with two water molecules in the system. The radicals pairs ($HOCH_2$ + HCO or OCH_3 + HCO) were oriented to form either glycolaldehyde or methyl formate, as in the DFT scans, and set at a distance of 5 Å apart and the distance then decreased in 0.05 Å steps. As previously, the geometries were re-optimised at the Hartree-Fock level at each step. Scans were performed with and without explicit water molecules present. This process was not carried out for the dissociation of H_2CO as there is only one possible product unlike with methanol where the aim was to discover the dominant product.

For both the methanol and methyl formate/glycolaldehyde MCSCF calculations, the active space was selected by first obtaining the orbital occupancies *via* an MP2 calculation. All orbitals with an occupancy greater than 1.98 were closed so that they were all doubly occupied and those with occupancies lower than 0.2 were restricted.[271,272] The restricted orbitals were limited to a total of no more than two electrons within those orbitals. The remaining orbitals were left active. The number of Configuration State Functions (CSFs) in each calculation are shown in Table 5.2.

Table 5.2: The number of CSFs in each RASSCF calculation

Reaction	Without water	With water
MeOH $\longrightarrow$ MeO$^\bullet$ + H$^\bullet$	1176	325
MeOH $\longrightarrow$ CH$_2$OH$^\bullet$ + H$^\bullet$	66	6
HCO$^\bullet$ + OMe$^\bullet$ $\longrightarrow$ Methyl formate	13965	276
HCO$^\bullet$ + CH$_2$OH$^\bullet$ $\longrightarrow$ Glycolaldehyde	5715	1891
HCO$^\bullet$ + CH$_2$OH$^\bullet$ $\longrightarrow$ Methyl formate	4186	10585

5.3 Testing the proposed formation route

5.3.1 Methanol dissociation

In order to approach understanding the system, the feasibility of forming the reactant radicals was first investigated. Initial work was conducted in the gas-

phase and then repeated with water present to act as a minimal ice surface. To ascertain whether the proposed radicals are viable options in the ISM restricted active space multi-configurational (RASSCF) scan calculations were carried out to create potential energy curves for the dissociation of methanol. For this either the alcohol hydrogen or one of the methyl group hydrogens was removed in incremental steps. At each step of the calculation the geometry is optimised as in all previous cases. The result of this without any water included in the system can be seen in figure 5.3. As can be seen in the curves for the OH dissociation the difference between the HF and RASSCF curves at long range is quite substantial. This shows the need for treating these systems in a multi-configurational manner to ensure the correct behaviour is modelled. The difference between HF and RASSCF is less for the CH dissociation, implying the active space is not fully correct. This has been sought to be corrected below.
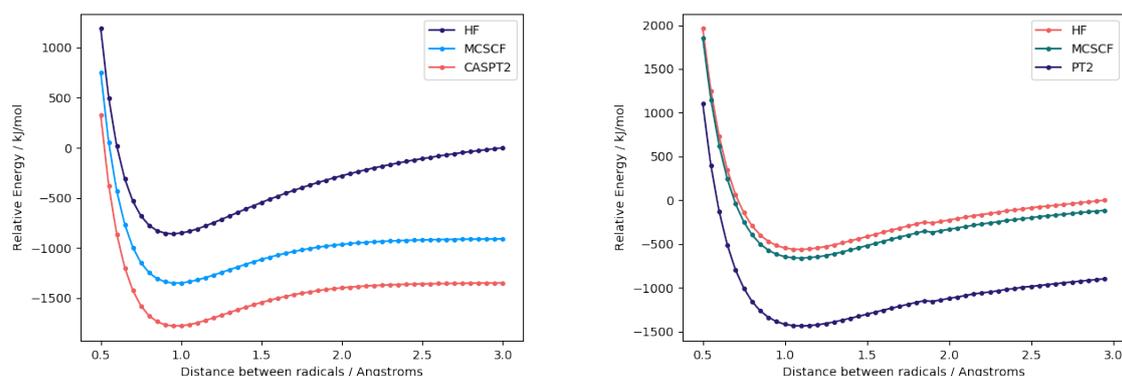
Figure 5.3: The dissociation curves for the OH bond (left) and CH bond (right) in methanol using HF, MCSCF and CASPT2 methods.
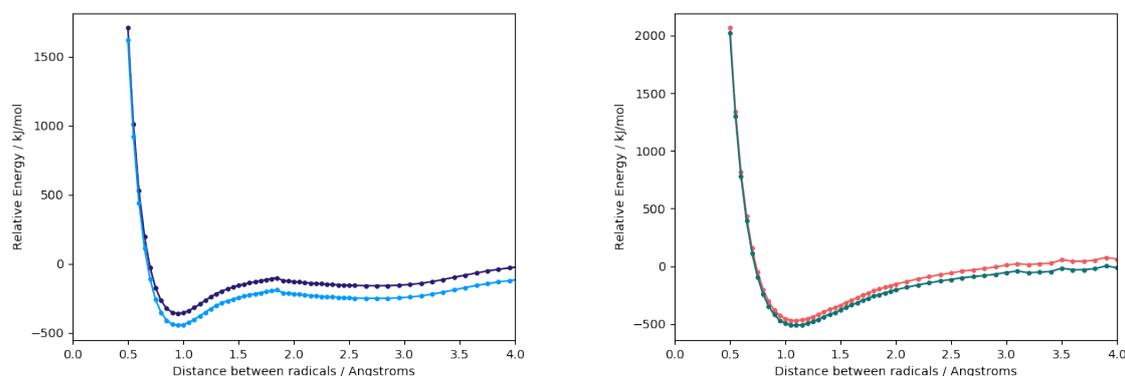
Figure 5.4: The dissociation curves for the OH bond (left) and CH bond (right) in methanol using HF (dark blue and orange respectively) and MCSCF (light blue and green) methods with two specific water molecules included in the system.

The resulting dissociation energies are 5.6 eV (129.1 kcal/mol) for the CH bond and 4.5 eV (103.8 kcal/mol) for the OH bond. In the case of OH, this is in reasonable agreement with previous computational and experimental work (105 kcal/mol) however, the CH dissociation is overestimated compared to the previously reported 96.2 kcal/mol.[273] This result suggests that the initial idea to use $HOCH_2$ in the radical pair with HCO due to the easier dissociation of the CH bond was incorrect and that OCH_3 is more likely to be feasible in an astrochemical environment than first thought. Additionally it is clear from the potential energy curves that using CASPT2 does not alter the shape of the curve or the dissociation but merely shifts it. In light of this further calculations were only done using RASSCF.

As this is to investigate the formation of methyl formate on an ice grain surface the same type of scan calculation was repeated but with two water molecules included in the system. Figure 5.4 shows the dissociation curves now with these two water molecules included. Such a small scale ice system was used as a larger system would not have been feasible to treat with MCSCF. Here, the similarity between the HF and RASSCF curves is thought to stem from the interactions between the water and the dissociating hydrogen radical allowing HF to treat the system better than in the non-solvated case. The dissociation energy for the CH bond shows little change being 5.5 eV (126.8 kcal/mol) while that for the OH bond drops significantly to 3.1 eV (71.5 kcal/mol). The undulations in the curve for the OH dissociation are caused by changes in the geometry of the system which are addressed in more detail below. The CH curve remains much smoother during the dissociation as no large structural changes are seen. It is clear from figure 5.4 that in this case the dissociation of the CH bond is described much better than in the bare case. Therefore, it was decided not to pursue the bare calculations any further.

A selection of the optimised geometries along the curve for each bond were visualised and showed why including water in the system only seems to affect the OH dissociation case. The dissociation energy for the CH bond remains largely unchanged from that found for the gas-phase dissociation and this stems from the lack of interaction between the removed hydrogen and the water. In contrast for the OH bond dissociation the water included in the system readily picks up the removed hydrogen. This can be seen in figure 5.5. There is a possibility that the way in which the water has arranged itself around the methanol affects the outcome. For instance in the OH dissociation case there is an oxygen atom directed at the out-going hydrogen ready to receive it while this is not the case for the CH bond. Given that the systems were optimised with regards to the geometries for this quantity of water the behaviour is representative. Larger ice systems may

show other behaviour as there would be more available ice molecules to accept the dissociating hydrogen.
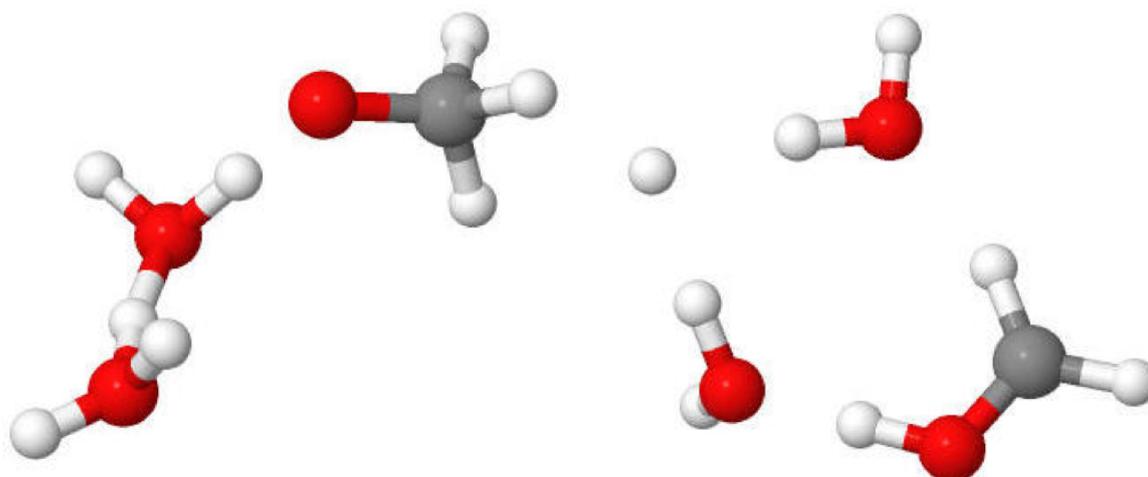
Figure 5.5: The resulting structures from the RASSCF calculations for the dissociation of the OH (left) and CH (right) bond in methanol with two nearby water molecules.

In order to further test this, a DFT scan calculation was carried out on the CH bond dissociation system with five water molecules. This amount of water was chosen as this was the smallest amount that showed a water oxygen atom placed in such a way to receive the dissociating hydrogen. DFT as opposed to RASSCF was used in this case as the previous HF curves showed good behaviour in the long range and because the increased system size made the use of multiconfigurational methods untenable (see figure 5.6). However, even with this larger amount of ice the water still does not take up the hydrogen radical and it remains largely unsolvated. It causes disruption of the hydrogen bonding network between the ice water molecules (see inset in figure 5.6). Obviously, in a more rigid ice lattice structure, the rearrangement of the water hydrogen bonding would be less feasible. Thus, this CH dissociation could be less unfavourable than portrayed here, but one would expect OCH_3 would still be the major dissociation product. This was also found to be consistent with observational findings in which OCH_3 has been detected towards the dense core B1-b confirming the viability of forming this species in the ISM.[274]

Consequently, if methyl formate is being made on icy grain surfaces then OCH_3 becomes an even more viable starting radical than $HOCH_2$. The significant difference in dissociation energies between the two bonds when able to interact with water ice could lead to a greater abundance of OCH_3 compared to $HOCH_2$. As a consequence of these results it would seem that the initial choice of radicals does not reflect what would occur in an interstellar environment and
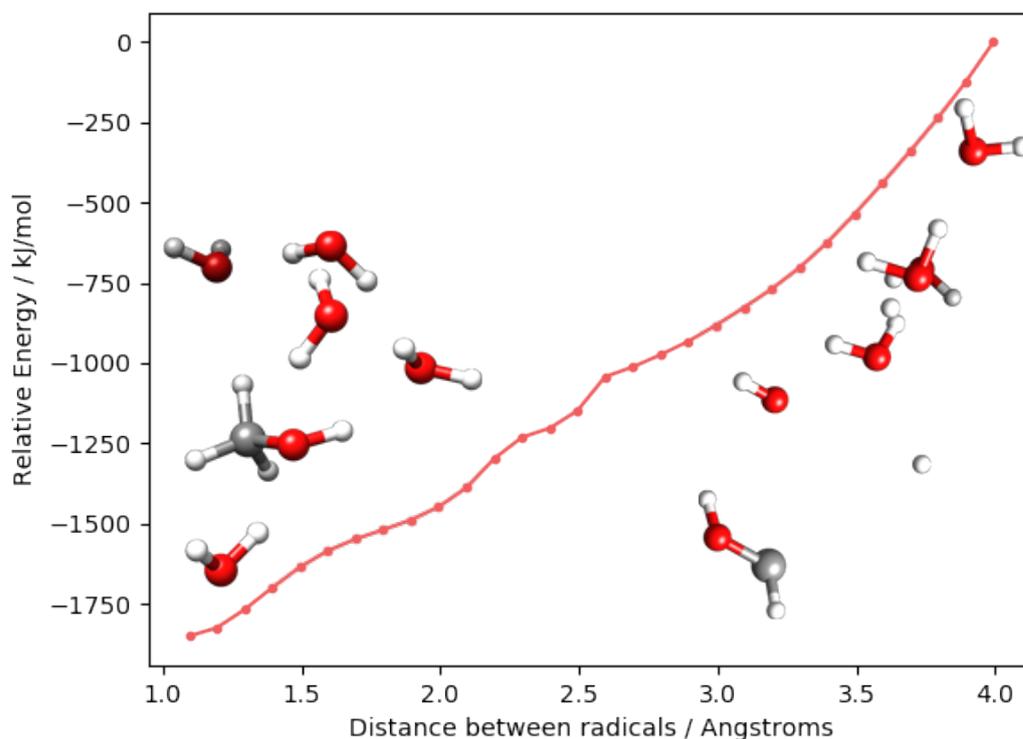
Figure 5.6: The potential energy curve as calculated by DFT for the dissociation of the methanol CH bond in the presence of five water molecules. Inset left: the initial structure of the system and inset right: the final structure of the system.

the distribution of the two isomers, methyl formate and glycolaldehyde, may stem from the ease of formation of OCH_3 over $HOCH_2$. In order to confirm this, though, the potential formation routes for methyl formate and glycolaldehyde from these radicals must be tested with regards to their viability in an astrochemical setting through further calculations.

5.3.2 DFT calculations

In order to explore the potential for using $HOCH_2$ and HCO to form methyl formate instead of glycolaldehyde, initial DFT scan calculations were carried out. The scope of these calculations was to give a low cost, short time scale indication of the systems behaviour. Unrestricted DFT was used to try and describe the radicals as best as possible. Hereby, we note that DFT does not treat open-shell systems as well as other methods used below. Thus, our DFT calculations were intended to give initial indications and inform our further approach.

As outlined in the previous section $HOCH_2$ is likely to be the minor product of methanol dissociation. To rule out competing reactions to form methyl formate or

97
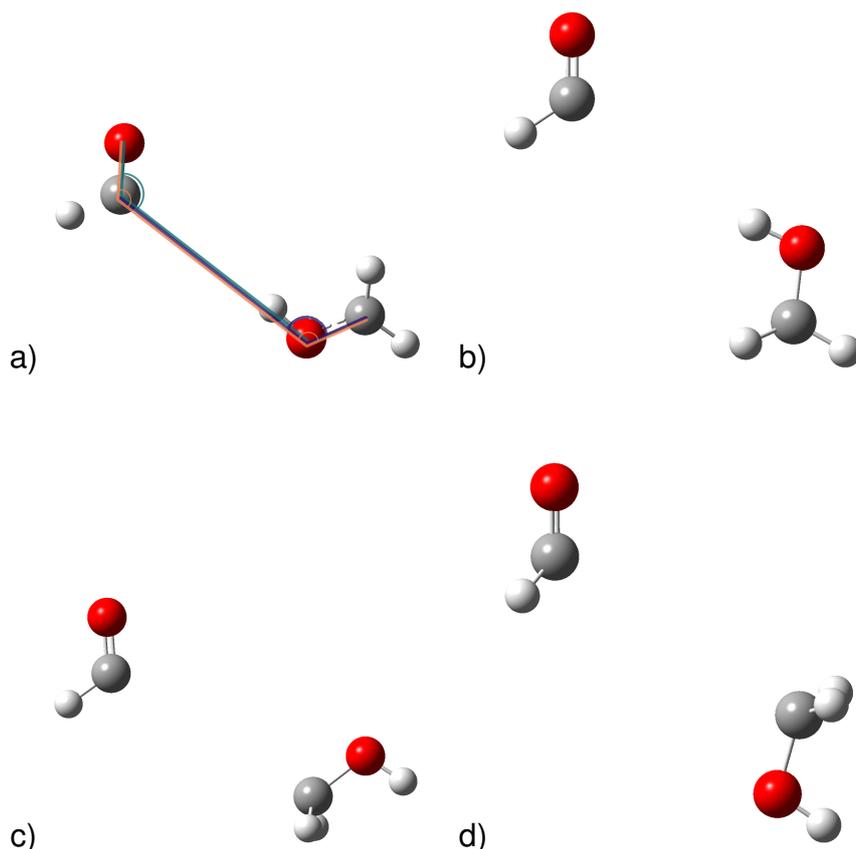
Figure 5.7: The input geometries for the formation of *cis*- (a) and *trans*-methyl formate (b) and *cis*- (c) and *trans*-glycolaldehyde (d) from the radicals HCO and HOCH$_2$. The highlighted OCO (green) and COC (dark blue) angles and the OCOC dihedreal (orange) in (a) are the additional structural elements kept fixed during the scans.

glycolaldehyde from HCO and HOCH$_2$ being a factor in the greater abundance of methyl formate over glycolaldehyde, the route to methyl formate formation from this radical pair has been tested. In these calculations the two radicals were brought together in a geometry conducive to forming the target product. Examples of the input geometries used can be seen in figure 5.7. In order to maintain the intended direction of approach for the incoming radicals a select combination of angles and dihedrals within the system were frozen in the optimisation in addition to the distance between the radicals. The restricted geometric variables are highlighted on the input structure for *cis*-methyl formate in figure 5.7 as an example.

When these scans, see figure 5.8, were performed in the absence of any water molecules neither methyl formate nor glycolaldehyde were observed. As is clear from the figure there are signifcant discontinuities in the graph. In these scans, arrangements of radicals that could lead to both the *cis*- and *trans*- isomers were tested in light of the possible existence of *trans*-methyl formate in the

98

ISM.[63] Investigation of the product structures showed that the ultimate products of these calculations were either two molecules of formaldehyde, in the case of methyl formate, or CO + CH_3OH when trying to form glycolaldehyde. This occurred for both *cis-* and *trans-* orientations. The point at which this unwanted product formation occurs can be seen on the potential energy curves marked by the sharp discontinuity at 3.25 Å or 2.75 Å for glycolaldehyde and methyl formate respectively in figure 5.8. Because we are only investigating a one-dimensional cut of a multidimensional surface, the optimisation that happens at each point is not stable leading to unintended side product formation. Smaller scan steps were attempted however these had no effect on the shape of the curve.

In the case of methyl formate, the alcohol hydrogen must move onto the carbon to generate the methyl group. Therefore, it is not so surprising that the product is not formed. However, to form glycolaldehyde a carbon-carbon bond simply needs to be made. That, therefore, glycolaldehyde is not seen then is more unexpected. Despite this, the result here is consistent with previous work by Woods *et al.*, which indicates that gas-phase formation routes for glycolaldehyde only begin to contribute to observed abundances at core temperatures of at least 100 K.[275,276] It would appear that prior to any carbon-carbon bond formation the hydrogen from the HCO radical transfers onto the carbon of the $HOCH_2$ to form the methanol observed. An analogous hydrogen transfer, although in this instance from the alcohol of the $HOCH_2$ onto the HCO, is occurring during the methyl formate scans to yield the two formaldehyde molecules seen. After this point, as the initial radicals are no longer present the desired products can no longer form. Moreover, as there are alternative reaction channels that are more favourable it is unlikely that there is any chance that methyl formate or glycolaldehyde would be produced in the gas-phase.

As a result of the absence of methyl formate or glycolaldehyde in the reaction, it suggests that ice is crucial to the success of the reaction. Therefore, an initial attempt at including ice in the system was made by running a set of scans with two water molecules present. Two water molecules were chosen as this is the minimum amount of water required to perform the hydrogen transfer needed to form methyl formate from HCO and H_2COH. This water was intended to act as a minimal ice surface as a grain surface reaction is thought to be the source of methyl formate in the ISM. It was hoped that it would both prevent the hydrogen transfer that leads to the unwanted side products and facilitate the formation of methyl formate. Figure 5.9 contains the potential energy curves and products for the calculations including water.

By including water in the system the loss of the radicals via formaldehyde formation is negated and both *cis-* and *trans*-glycloaldehyde are observed. Ad-
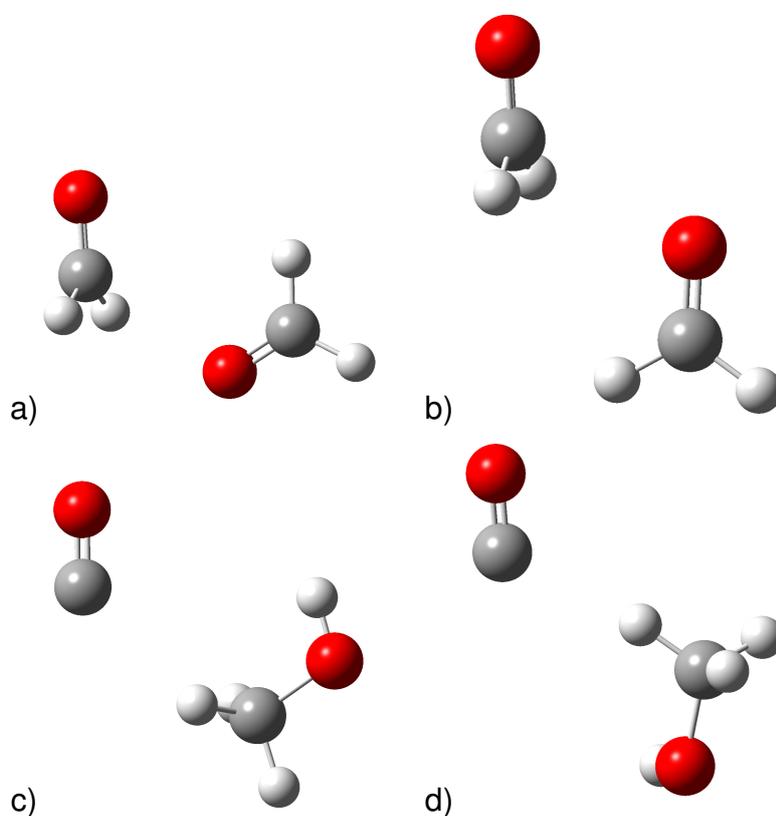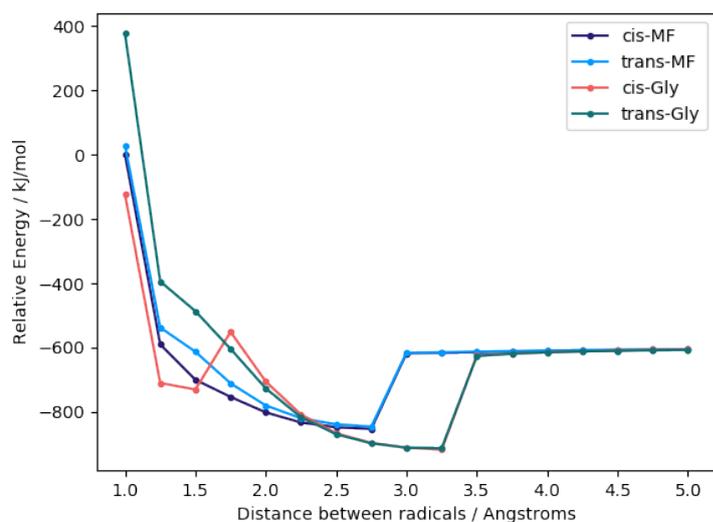
Figure 5.8: Upper: the potential energy curves for the attempted formation of *cis*-methyl formate (dark blue), *trans*-methyl formate (light blue), *cis*-glycolaldehyde (orange) and *trans*-glycolaldehyde (green). Lower: the lowest energy geometries resulting from each scan, *cis*-methyl formate (a), *trans*-methyl formate (b), *cis*-glycolaldehyde (c) and *trans*-glycolaldehyde (d).

ditionally the discontinuities that were present in the gas-phase curves are no longer present, see figure 5.9. This is likely a result of the ability of the water molecules to hydrogen bond with the hydrogens present on the radicals thus pre-
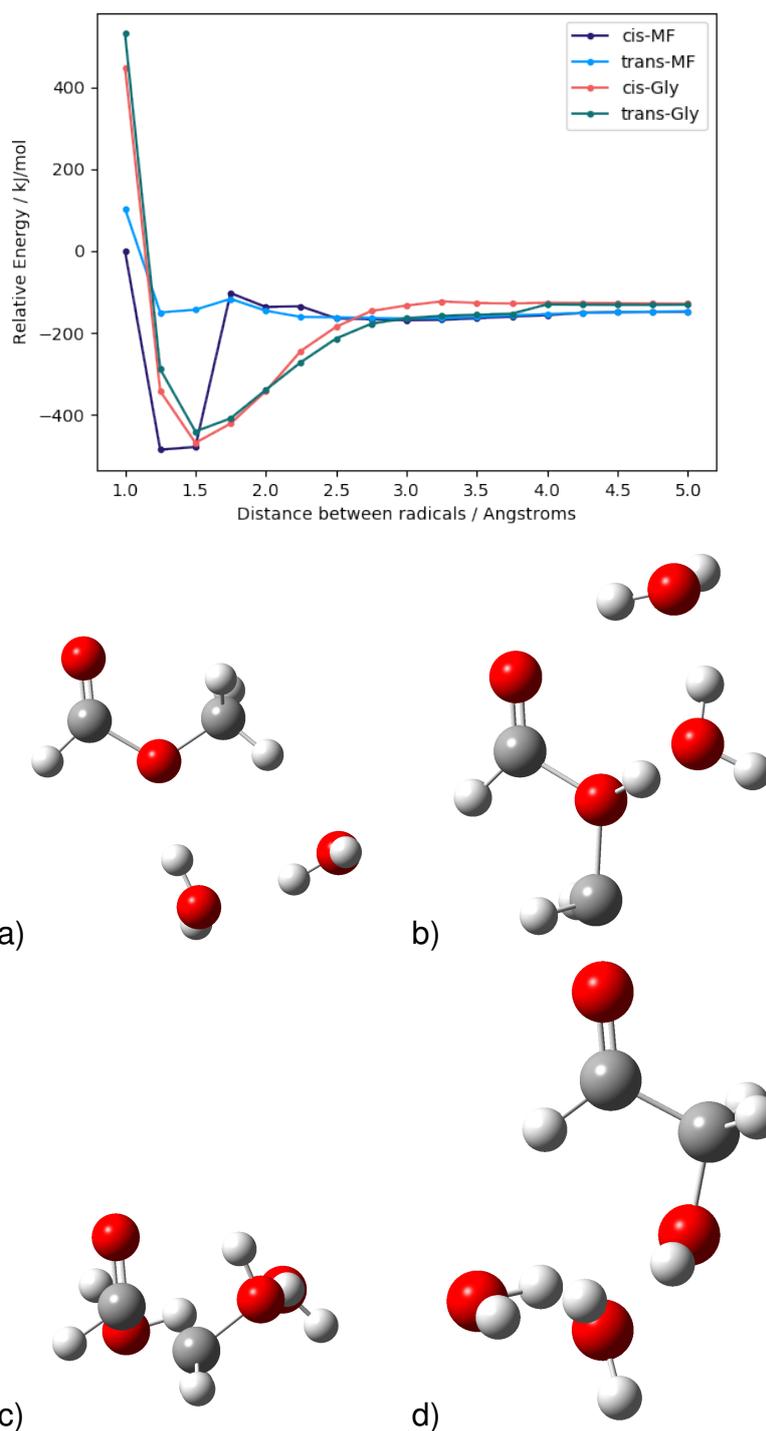
Figure 5.9: Upper: the potential energy curves for the formation of *cis*-methyl formate (dark blue), *trans*-methyl formate (light blue), *cis*-glycolaldehyde (orange) and *trans*-glycolaldehyde (green) with two water molecules included in the system. Lower: the lowest energy geometries resulting from each scan *cis*-methyl formate (a), *trans*-methyl formate (b), *cis*-glycolaldehyde (c) and *trans*-glycolaldehyde (d).

serving them until they are close enough to form a carbon-carbon bond. As shown by figure 5.9, of the two methyl formate isomers only *cis*-methyl formate is seen as a product. Again this shows that water acts as a means of maintaining the radicals until they are within a suitable reaction distance. It suggests that water can act as a hydrogen shuttle, or bridge, allowing the alcohol hydrogen to 'move' from the oxygen and onto the carbon to form the methyl group of methyl formate. At present it is thought that *trans*-methyl formate could equally be formed in this manner but is not observed here due to the placement of the water molecules. As opposed to spanning the distance between the $HOCH_2$ oxygen and carbon the water molecules are arrayed around the carbonyl of the HCO. The lack of *trans*-methyl formate is therefore likely to be a result of this difference in water placement.

While *cis*-methyl formate is able to be formed from $HOCH_2$ + HCO there is a barrier to this reaction of 72.3 kJ/mol. The barrier corresponds to the movement of the hydrogen atoms from the alcohol group and water molecules leading to CH_3 group formation. Given ISM conditions the presence of a barrier is too great to surmount. A transition state calculation followed by an internal reaction coordinate (IRC) scan were carried out to further explore the barrier. The geometry of the calculated transition state can be seen in figure 5.10. Tunnelling of light atoms
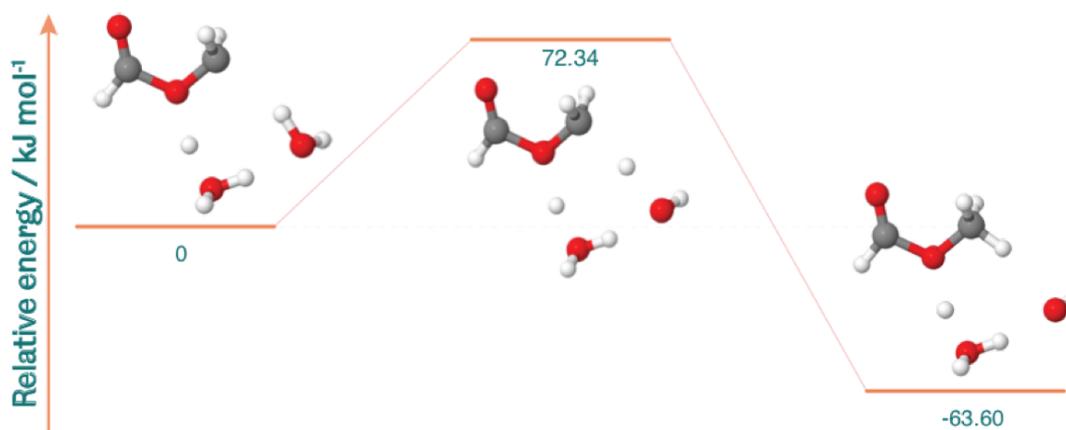
Figure 5.10: The calculated transition state for the formation of methyl formate from HCO and $HOCH_2$ *via* hydrogen transfer between the radicals and the included water molecules.

has been shown to contribute to COM abundances in the ISM and therefore, before discounting the reaction, the potential for hydrogen to tunnel through the barrier was explored.[60] By fitting a gaussian curve to the IRC data it allowed for a tunnelling probability to be calculated using the Wentzel-Kramers-Brillouin (WKB)

method, which is a method for approximating the solution to a time-independent one-dimensional differential equation.[277] The calculated tunnelling probability was 4.25×10^{-17}. As lifetimes in the ISM can be lengthy the formation of methyl formate from these radicals may still be a feasible, although minor, route *via* hydrogen tunnelling. However, due to the difficulties of treating radical systems with DFT it is possible this result is simply an artefact of the method giving the system ionic character which is not representative of the actual processes taking place. In light of this, we explored this reaction further with MCSCF which deals with radicals much more accurately, as detailed in the next section.

5.3.3 Multi-configurational calculations

As DFT is a single determinant method, which often results in long range interactions that are incorrectly ionic in nature, and as radicals are often multi-reference problems, multi-configurational self-consistent field calculations (MC-SCF) were performed to confirm the results of the DFT scans. As before the HCO and $HOCH_2$ radicals were used to form both methyl formate and glycolaldehyde, both in the *cis* arrangement, as well as using HCO and OCH_3 to generate methyl formate for comparison. The potential energy curves for the glycolaldehyde and methyl formate from HCO + OCH_3 scans can be seen in figure 5.11. Both of these scans result in the intended products with no barrier to formation (see figure 5.11). Additionally, treating the problem in a multi-configuration manner yields curves displaying much better long range behaviour when compared to the HF curve.

It is therefore reasonable to believe that these radical pairs could represent formation routes for glycolaldehyde and methyl formate in the ISM. Radical formation routes have already been found to be viable in the ISM. For example the dimerisation of the formyl radical can barrierlessly lead to glycolaldehyde, so this results is consistent with other work in the field.[276] In figure 5.12 it is shown that when the same type of scan was performed fo HCO + $HOCH_2$ to form methyl formate the product is not observed at any point during the scan. Repeating the scan with water included in the system, as with those done using DFT, results in the same lack of product. There is no noticeable product minimum to be found in the curve (shown in figure 5.12). A sample of the optimised geometries along the scan path show both the two formaldehyde molecules seen in the waterless DFT scans as well as formic acid plus CH_2 and then $CO_2 + CH_2 + H_2$ at shorter inter-radical distances. These can all be seen in figure 5.12.

Given that, even with water included, the RASSCF calculations on HCO + $HOCH_2$ to form methyl formate do not generate the target COM suggests that
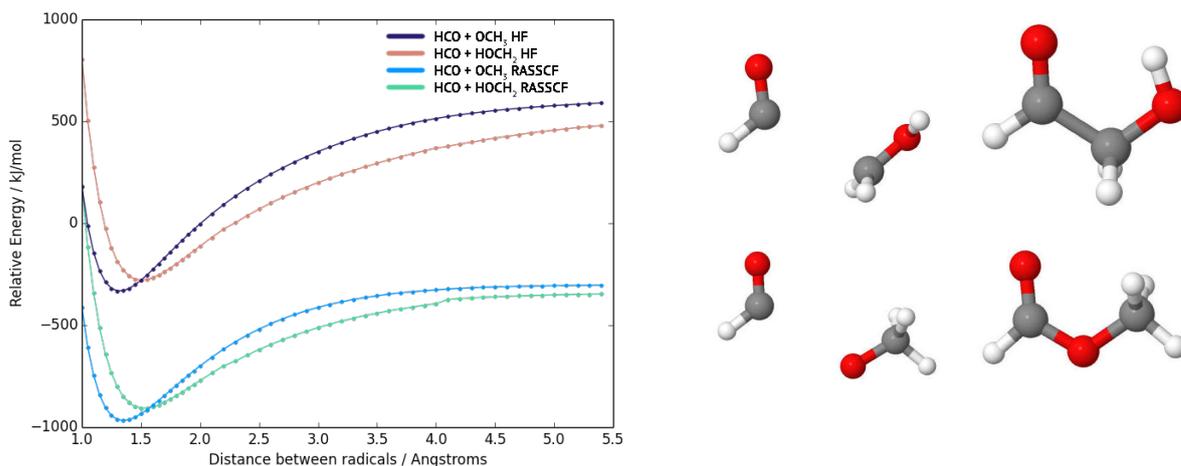
Figure 5.11: Left: the potential energy curves for the formation of *cis*-glycolaldehyde from HCO + HOCH$_2$ using HF (orange) and using RASSCF (green) and for the formation of *cis*-methyl formate from HCO + OCH$_3$ using HF (dark blue) and RASSCF (light blue). Right: the input geometries and the geometry of the lowest energy point on the potential energy curves for the formation of *cis*-glycolaldehyde (upper) from HCO + HOCH$_2$ and *cis*-methyl formate (lower) from HCO + OCH$_3$.

this may not be a feasible formation route. It is thought that in the DFT calculations the radicals are behaving in an unwanted ionic way making the hydrogen removal appear more facile than it would actually be. In light of this and the results of the methanol dissociation calculations it would appear that HOCH$_2$ + HCO are not a suitable radical pair for forming methyl formate and competition between formation methyl formate and glycolaldehyde from this radical pair does not contribute to their relative isomeric abundances. As can be seen in figure 5.11 forming methyl formate from the more intuitive radical pair HCO + OCH$_3$ is barrierless and does not require any additional water molecules to facilitate the reaction. The same can be said for the formation of glycolaldehyde from HOCH$_2$ + HCO and thus, should either radical pair encounter one another on the ice the relevant COM is likely to form. Therefore, the conclusion must be drawn that the relative ease of OCH$_3$ formation compared to HOCH$_2$ formation is the more strongly contributing factor in the difference in abundance between the two isomers.

5.4 Conclusions

The gas-phase dissociation of methanol gave energies of 5.6 eV for the CH bond and 4.5 eV for the OH bond. However, when water ice is included in the calculations the difference becomes significantly more pronounced. The dissociation energy for the CH bond was found to change very little and is 5.5 eV. The
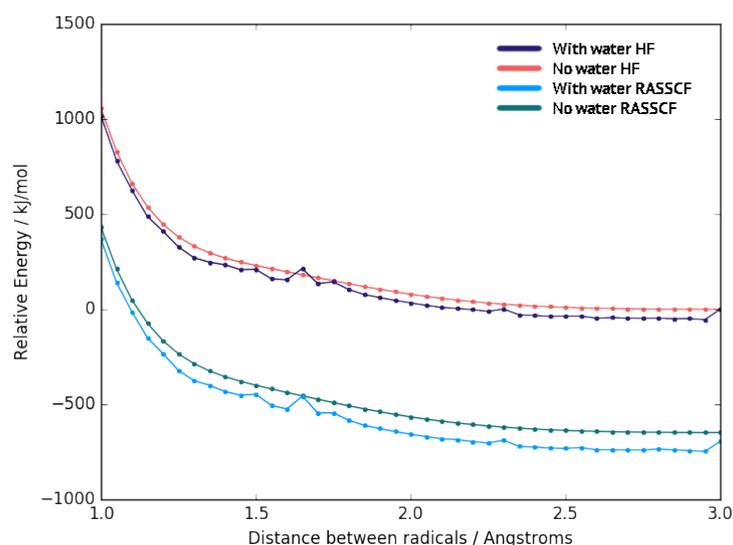
Figure 5.12: Upper: the potential energy curve for forming *cis*-methyl formate from HCO and HOCH$_2$ with water (HF in dark blue and RASSCF in light blue) and without water included (HF in orange and RASSCF in green) Lower: the optimised structures from using HCO + HOCH$_2$ to form methyl formate as calculated using RASSCF methods. Where a is the initial geometry, b is the point at which formic acid is observed and c is the final products at very small CO distances.

OH bond dissociation energy, in contrast, decreases by one and half electron volts to 3.1 eV. By examining a selection of optimised structures for both bonds along the dissociation curve this difference stems from how much the removed hydrogen interacts with the water. In the CH bond case there is little interaction with the water. For the OH bond the water picks up the dissociating hydrogen making the process more facile

The results of the DFT calculations on the formation of methyl formate and glycolaldehyde from HOCH$_2$ + HCO show no product when performed in the gas-phase. Instead either two formaldehyde molecules or methanol and CO are seen. These products were found to be the result of hydrogen transfer between the two radicals at distances longer than that at which they would react to form the in-

tended product. When water is included in the system both intended products are seen as the water helps to preserve the radicals until they are at a suitable distance at which to react. In the methyl formate calculations the water performs an additional role in facilitating the movement of the alcohol hydrogen onto the CH_2 group to form the methyl group of the methyl formate molecule. This hydrogen movement has an associated barrier of 72.3 kJ/mol. Tunnelling calculations give a tunnelling probability of 4.25×10^{-17}. In general, it was found that DFT dealt poorly with the radical based problems contained in this chapter.

Analogous multi-configurational calculations on the formation of methyl formate from both OCH_3 + HCO and $HOCH_2$ + HCO as well as glycolaldehyde from $HOCH_2$ + HCO. These show that both glycolaldehyde and methyl formate will form without barriers from $HOCH_2$ + HCO and OCH_3 + HCO respectively. Using $HOCH_2$ + HCO to form methyl formate does not yield the desired product even in the presence of water. This is in contradiction with the DFT results. However, as multi-configurational methods are much better suited to radical systems this result is taken to be the more representative of the chemistry taking place.

As it was found that dissociating the methanol OH with water present in place of an icy grain mantle requires less energy than dissociating the CH bond, and that the recombination with OCH_3 + HCO is a barrierless reaction route. Then methyl formate formation using OCH_3 and HCO is potentially viable in the ISM.

6. Urea formation from radical and charged species

In this chapter the formation of another organic molecule in the ISM is discussed. In particular urea, which has been briefly covered already in section 1.1.2 of the introduction. The following sections will outline the biological importance and context of urea as well as several possible formation pathways using neutral, radical and charged reactant species.

6.1   Urea as an astrobiological molecule

Urea, $CO(NH_2)_2$, is the simplest known dipeptide. Terrestrially it is the waste product of a variety of biological functions and often used as a fertiliser. It was also the first biological molecule to be produced entirely abiotically.[278] As such, it is viewed as a potentially important molecule in the abiotic origin of life, and was one of the identified products of the now famous Urey-Miller experiment in the 1950s.[111] A subsequent experiment, for the prebiotic formation of pyrimidines, by Robertson and Miller further showed the relevance of urea to the origins of life field.[279] They showed that urea acts as a carbon source in the formation of cytosine and uracil via the scheme given in figure 6.1.[280] The biological relevance of these two nucleotides is clear and fits in with the proposed RNA world hypothesis (which was discussed in section 1.1.2 in the introduction) for the Origins of Life on Earth.[281]
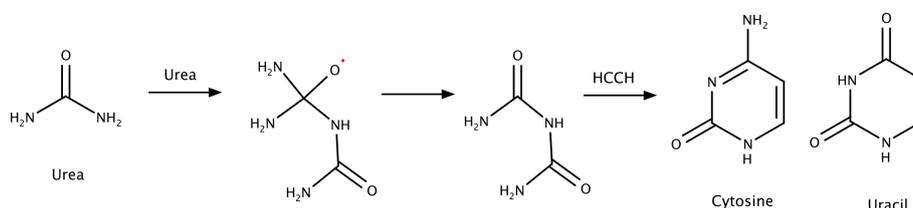
Figure 6.1: The reaction of urea with itself to ultimately yield the nuclear bases cytosine and uracil.

Thus, in the current models for the formation of biologically relevant molecules in the ISM, urea is key. However, there are some concerns over the proposed terrestrial means for the production of urea. It was thought that a Wöhler-type synthesis mechanism would account for the presence of urea on early Earth. A Wöhler type synthesis starts from $NH_4(OCN)$ which dissociates into ammonia and HOCN upon heating. These species then react to form urea, see equation 6.1.[282]

$$NH_4(OCN) \longrightarrow NH_3 + HOCN \rightleftharpoons (NH_2)_2CO \qquad (6.1)$$
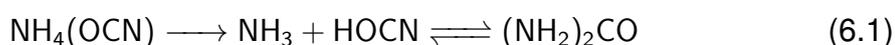
However, concentrations of ammonia cyanate on early Earth are thought to be low. This problem is compounded by the fact that cyanates are not stable in water over geological timescales. This would substantially limit the amount

of cyanate present. As a consequence this formation route for urea would be difficult.[283] An alternative to terrestrial formation would be that urea is formed in space and brought to Earth during cometary impacts. The identification of urea on the Murchison meteorite, along with amino acids and nuclear bases, lends credence to this theory.[113]

While urea has been confirmed *via* its presence on the Murchison meteorite it has yet to be firmly detected in any region of the interstellar medium (ISM).[284] Experiments using model ices as well as chemical networks suggest that a variety of amines and amides should be present though, suggesting urea could be formed readily. Formamide, acetamide and methylamine have all been detected in the ISM. Formamide in particular, is a possible precursor to urea.[285–287] Urea has been synthesised in model ices irradiated to mimic the galactic cosmic rays (GCRs) and high energy UV present in the ISM. However, the exact mechanism has yet to be determined.[114] Two potential mechanisms have been suggested by Förstel *et al.* using formamide, detected along with urea in the model ice experiments, as an intermediate. These proposals are outlined in figure 6.2.
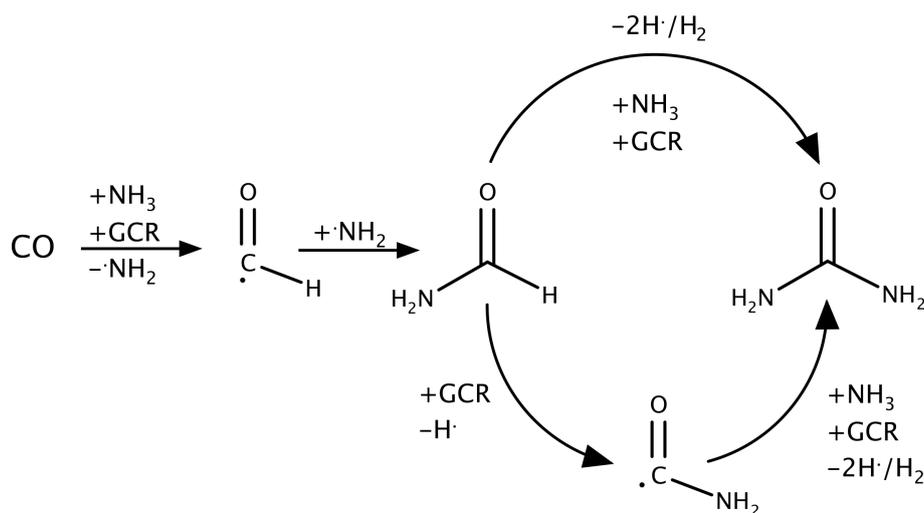
Figure 6.2: The possible mechanisms for the formation of urea from formamide and ammonia proposed by Förstel *et al.*.[114]

The successful detection of formamide in the ISM and the wealth of existing work on the formation thereof means that the way in which it is produced in the ISM is not probed in this work.[288–291] Instead, this work aims to use computational means to test the feasibility of these proposed mechanisms for the formation of urea from formamide and ammonia under astrochemical conditions.

6.2 Methodology

Calculations were performed using the GAUSSIAN 09 suite of electronic structure programmes [*] using DFT at the CAMB3LYP level[148,149,233] using a 6-311G** basis.[234,292] Further calculations were carried out using the MOLPRO suite of *ab initio* programmes [†].

GAUSSIAN 09, using DFT at the CAMB3LYP/6-311G** level, was used to optimise the ground state structures of all the reactants for the formation of urea and these optimised structures were then used to construct z-matrices for scan type calculations.

The scan type calculations carried out in MOLPRO using the MCSCF/RASSCF programmes and the 6-311G(d,p) basis set (see appendix A.1 for an example of a MCSCF scan input) used the same approach as in Chapter 5. The reactant species were placed at a maximum inter-reactant distance of 5 Å and the distance decreased in steps. The geometries, except for any variables required to maintain the angle of approach for the incoming radical and the inter-radical distance, were re-optimised at the Hartree-Fock level at each step. For systems where a restricted active space was required orbitals with occupations higher than 1.98 were closed so that they were doubly occupied and those with occupancies lower than 0.2 were restricted.[234,267–269,271,272,292] The restricted orbitals were limited to a total of no more than two electrons within those orbitals. The remaining orbitals were left active. An indication of the size of the active space and the number of CSFs included is given in Table 6.1.

Table 6.1: The number of CSFs in each RASSCF calculation

Reaction	Number of CSFs
$H_2NCHO + H_2N^{\bullet}$	182182
$H_2NCO^{\bullet} + H_3N$	182182
$H_2NCO^{\bullet} + H_2N^{\bullet}$	89232

DFT scans with and without additional ice species present were also set up. These were carried out such that the reactants were oriented to yield the target product, similarly to the MCSCF scans, at an initial long distance. This distance between the reactants was then decreased in steps of 0.5 Å to a minimum distance of 1 Å. At each successive step the structures were re-optimised at the CAMB3LYP/6-311G* level. For the isocyanic acid calculations, relaxed scans

---

[*]GAUSSIAN 09, revision D.01, M. J. Frisch et al., Gaussian, Inc., Wallingford, CT, 2009

[†]MOLPRO, version 2015.1, a package of ab initio programs, H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schtz, and others , see http://www.molpro.net.

using redundant coordinates were used to scan over the inter-reactant distance in steps of 0.1 Å with the geometries being re-optimised at each interval at the B3LYP/6-311G** level.[148,149,233]

6.3 Comparison of the potential formation reactions

All the routes explored in this work are outlined in figure 6.3. The intermediates and final products as calculated are given for those steps where the structures are known. The subsequent sections of this report will address each type of route; neutral, single and double radical and charged, in turn and assess their feasibility under astrochemical conditions.
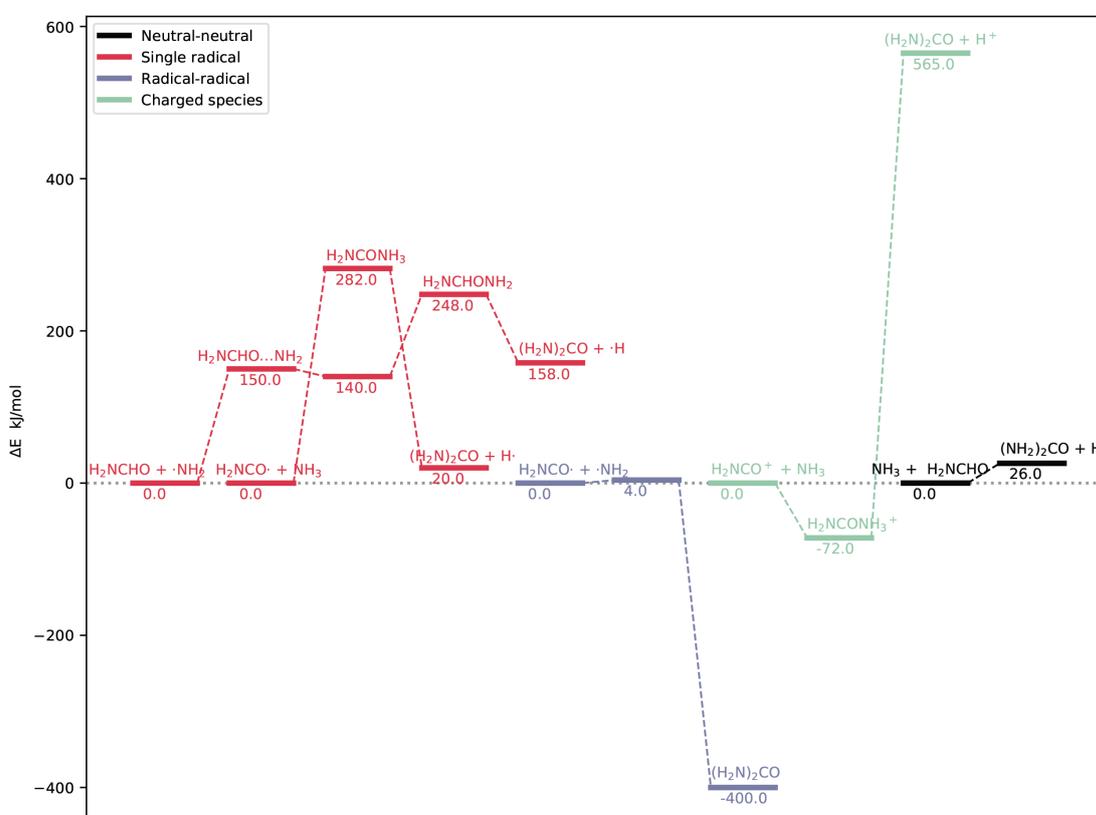
Figure 6.3: The reaction profiles for the urea formation routes investigated in this work. The energies in each route are plotted relative to the energy of the initial reactants. Routes in black are neutral-neutral reactions, red routes involving a single radical, blue routes are between two radicals and green routes use charged species.

6.3.1 Neutral-neutral Pathways

As shown in the proposed formation from Förstel *et al.* one potential route to forming urea is the reaction between formamide ($H_2C(O)H$) and ammonia.[114]

This reaction requires the incoming ammonia to cleave an NH bond and the formamide to break its CH bond allowing dihydrogen to form as a side-product in a concerted reaction. In the first instance, this reaction was followed in the gasphase by scanning along the formamide-ammonia intermolecular distance using DFT. The resulting potential energy curve can be seen in figure 6.4. The potential energy curve shows that the reaction products are less energetically stable than the reactants, being almost 400 kJ mol^{-1} higher in energy. Closer inspection of the optimised structures along the scan show that the NH and CH bonds highlighted previously do not break. No dihydrogen side product is seen and consequently urea does not form in this case. The structure at the point with an intermolecular distance like the CN bond in urea, which one would expect to be at the bottom of the well of the curve, is given in the inset of figure 6.4.
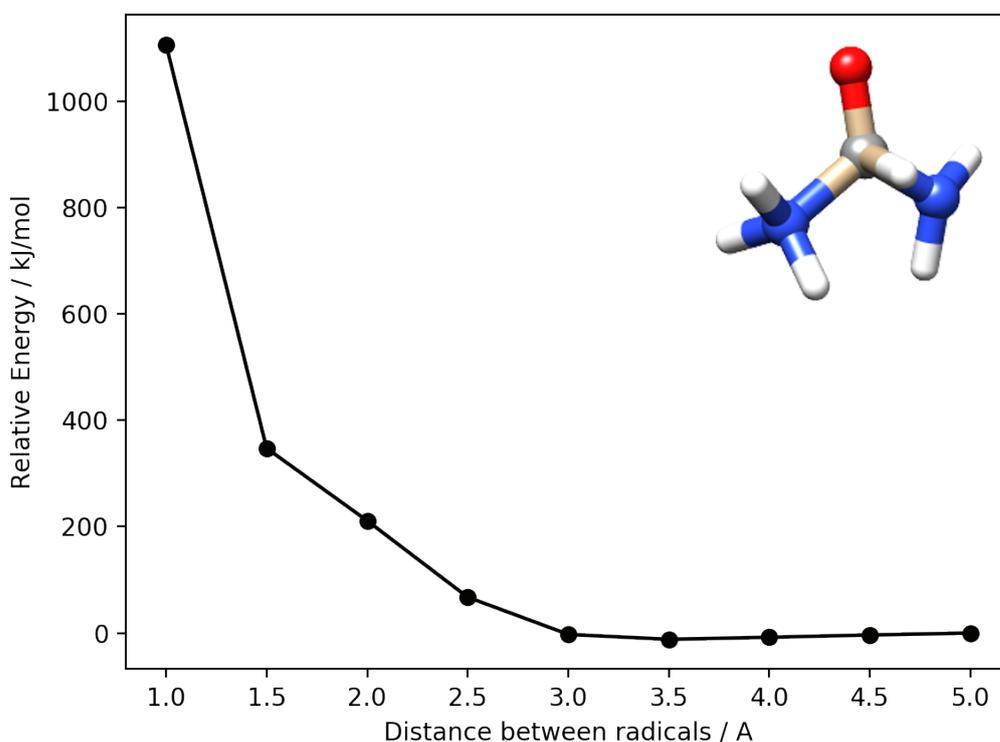
Figure 6.4: The potential energy curve for the addition of ammonia to formamide in the gas-phase to give urea as calculated by DFT where the energy at an intermolecular distance of 5 Å is taken as the zero. Inset: The structure at an intermolecular distance of 1.5 Å for the reaction between formamide and ammonia to make urea, showing the lack of NH and CH bond cleavage.

Literature gives the homolytic cleavage enthalpy of an NH bond in ammonia as +450 kJ mol^{-1} (4.6 eV) which is a significant energy input.[293] In astrochemical settings this energy would have to come from high energy UV radiation. A,,onia

radicals have been observed in the ISM but their abundance is highest towards the edges of dense clouds where radiation can penetrate more easily.[294] Therefore, this route may only be viable in the gas-phase at the boundaries of interstellar clouds.

However, it is not in the gas-phase that we expect this reaction to be occurring. The experimental work done by Förstel *et al.* was carried out by irradiating a simulated ISM ice. This means that calculating the potential energy curve for this reaction in the absence of any surrounding ice species is not a good description of the system of interest. Surrounding ice species could offer stabilisation to the departing hydrogen atoms and facilitate their removal allowing urea to form where it could not in the gas-phase.

In order to more accurately represent the system, the DFT scan calculation for forming urea from formamide and ammonia was repeated with three surrounding water molecules acting as neighbouring ice species. The resulting potential energy curve from this ice containing scan is shown in figure 6.5. As can clearly be seen, the inclusion of ice species does little to alter the shape of the curve from the earlier gas-phase calculation. Again, inspection of the optimised structures generated at each point along the scan show that the water has no interaction with the departing hydrogens. The structure at the point where the intermolecular distance is similar to that of the CN bond in urea (figure 6.5) indicates that the water would preferentially interact with the carbonyl which leaves them poorly placed to offer any stabilising interactions with the dissociating hydrogen atoms.

The scale of the ice included in this scan is obviously limited; the decision to use only three water molecules was to try to elucidate the nature of the stabilising interactions without the calculation becoming too computationally expensive. It is therefore possible that with with a larger scale ice, interactions not present here would be observed. There is also the possibility that the formation of urea is not seen here as the homolytic cleavage of the NH anbd CH bonds is being poorly treated by DFT, a method known to struggle with radical systems.[295]

To try and confirm the possibility for this reaction route, high accuracy CCSD(T) energy calculations were carried out on the products and reactions of this proposed mechanism step. The results of these calculations in the gas phase show that the formation of urea and dihydrogen from formamide and ammonia is, in fact, endothermic by 26 kJ mol^{-1} and so unlikely under ISM conditions (figure 6.6). The inclusion of the zero point energy correction does reduce the endothermicity to 8 kJ mol^{-1}, an amount that could be surmounted in the ISM. However, this does not take into account any potential transition state which was not able to find. It would be expected to be quite substantial since two bonds are broken
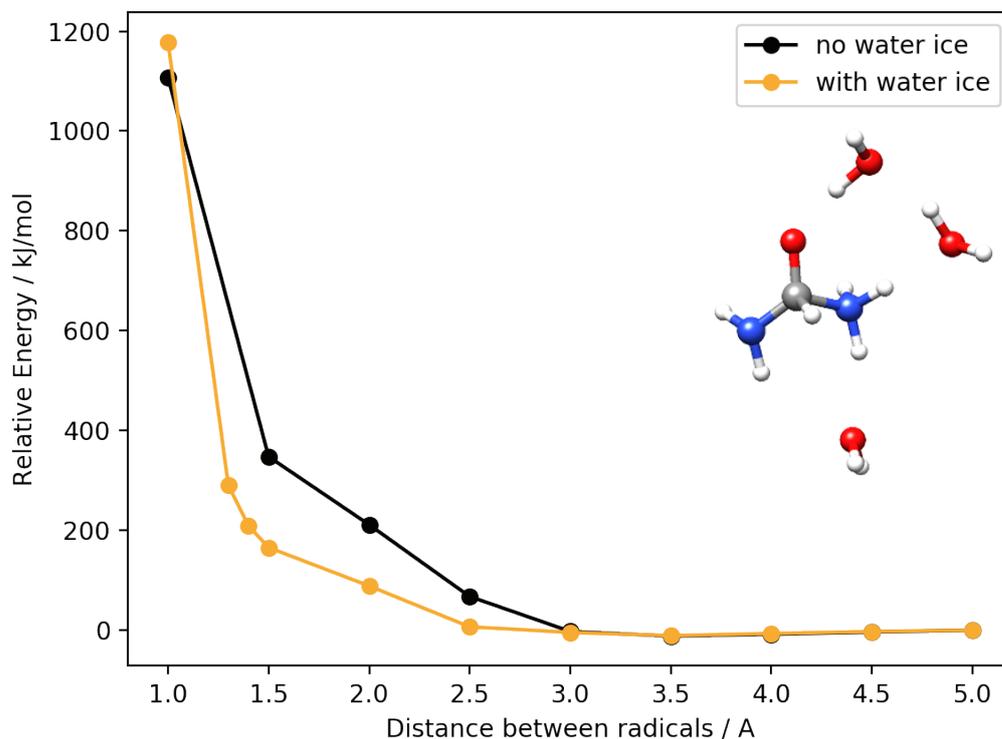
Figure 6.5: The potential energy curve for the addition of ammonia to formamide to give urea as calculated by DFT. Curves with and without three water molecules acting in place of an icy dust grain mantle are shown, where the energy at an intermolecular distance of 5 Å is taken as the zero. Inset: The structure at an intermolecular distance of 1.5 Å for the reaction between formamide and ammonia to make urea, with three water molecules acting as ISM ice species, showing the lack of NH and CH bond cleavage.

in this reaction whereas only one is formed. In light of this, alternative reactions have been investigated.

6.3.2 Single radical pathways

The alternative reaction mechanism for the formation of urea given by Förstel *et al.* uses a formamide radical and ammonia as the reactants. Radical reactions are prevalent in the ISM and contribute to the formation of many COMs making this a promising formation pathway.[296–299] Initially, the dissociation of formamide to yield the radical species was investigated. The same scan type set up as outlined previously was used but this time the reaction coordinate was the formamide CH bond. This process proves to be barrierless and has a dissociation energy of 644 kJ mol^{-1} (6.7 eV) in the gas-phase and 678 kJ mol^{-1} (7.0 eV) when on water ice. The potential energy curve can be seen in figure 6.7. The calculated dissociation
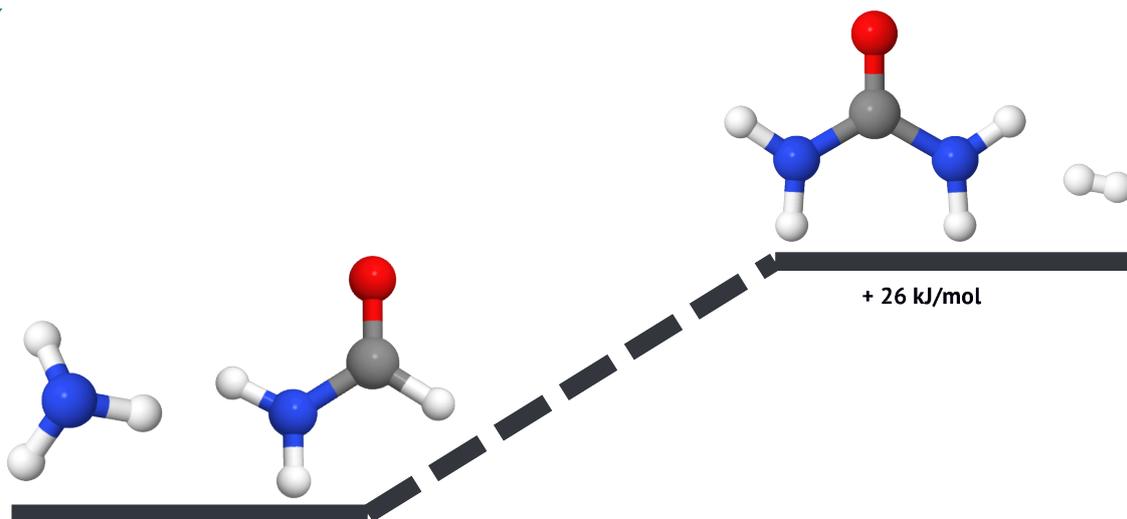
Figure 6.6: The reaction profile for the addition of neutral formamide to neutral ammonia in the gas-phase showing the reaction to be endothermic and thus not feasible under astrochemical conditions
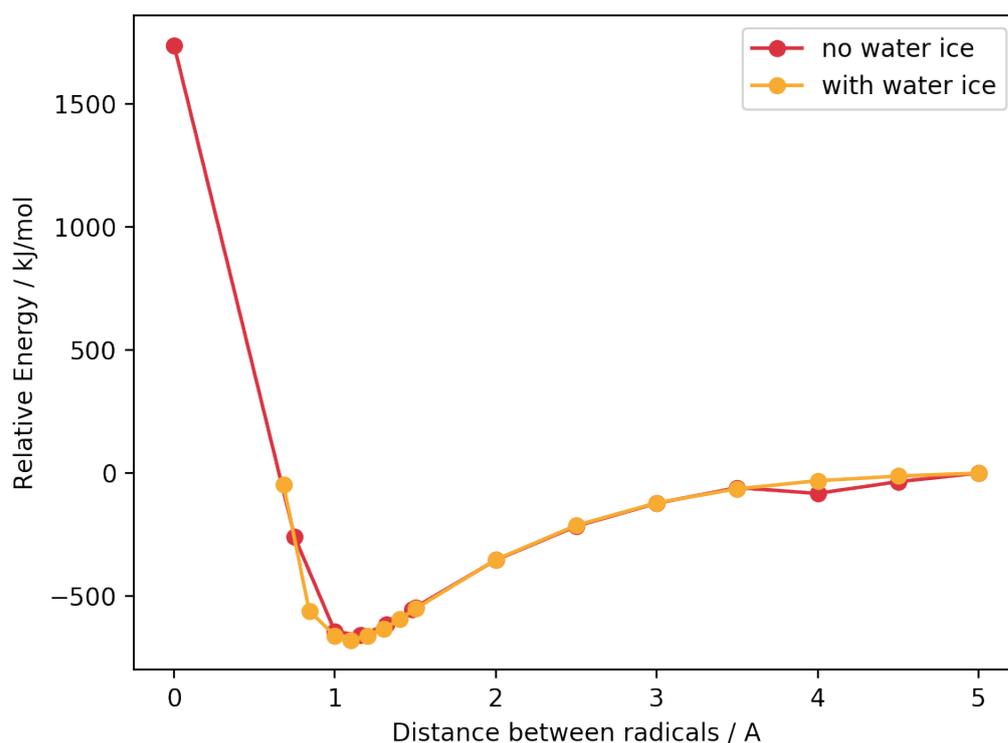
Figure 6.7: The potential energy curve for the dissociation of formamide to give the formamide radical in the gas-phase as calculated by DFT with and without water ice molecules present where the energy at a radical distance of 5 Å is taken as the zero.

energy could sufficiently be overcome by the high energy UV radiation present in

the ISM and so it is highly likely that, so long as the radiation can penetrate far enough into the cloud, the formamide radical will be present.

As before, DFT scan calculations were run to generate the potential energy curve for the reaction between a formamide radical and ammonia with and without water ice present. Both potential energy curves can be seen in figure 6.8. As with the formamide plus ammonia reaction route the curve shows no minimum around the equilibrium CN bond length in urea. However, unlike the previous case, hydrogenated urea is observed as a product. As the scan progresses, a hydrogen radical does dissociate from the NH_3 moiety of the hydrogenated urea but this still does not correspond with a minimum in the potential energy curve. Adding water into the system does not help stabilise the hydrogen radical and thus an energy minimum is still not seen in the ice-containing calculations. As previously stated, DFT does not describe radical reactions well and so these calculations were not expected to yield a complete view of the system but to inform the approach taken with subsequent multiconfigurational calculations.

RASSCF calculations were carried out to follow the reaction between the formamide radical and ammonia. In these calculations, the CN bond was used as the reaction coordinate and the structures were re-optimised at each step. The potential energy curve for this reaction calculated using multiconfigurational methods is given by figure 6.9. Here, urea is observed to form, along with a hydrogen radical. However, the products are higher in energy than the reactants by $20 \, \text{kJ} \, \text{mol}^{-1}$ and there appears to be a substantial barrier to the formation of the products of $282 \, \text{kJ} \, \text{mol}^{-1}$ which would be insurmountable under astrochemical conditions. Inspection of the transition state structure (inset in figure 6.9) shows that the barrier corresponds to the midpoint of NH bond cleavage and CN bond formation. After this point the ammonia nitrogen forms a CN bond with formamide to make urea and the hydrogen radical formed departs. The small discontinuity in the gradient of the potential energy curve at 1.55 Å corresponds to a geometric change in the system. At 1.55 Å the nitrogen switches arrangement from the trigonal ammonia conformation to the planar NH_2 group found in urea.

Clearly, this particular route to urea is not feasible and so alternatives must be sought out and their potential tested. One such possibility would be to use formamide and an amino radical as opposed to the formamide being the radical species. A scan of the CN distance for these reactant species was performed. The potential energy curve, figure 6.10, shows no large barrier however, the reaction is endothermic. The dip in the energy at roughly 3.5 Å is caused by the formation of favourable interactions between the amino radical and the formamide carbonyl and amino group, forming a six-membered hydrogen bonding ring. The disruption of these interactions as the CN intermolecular distance decreases causes

116
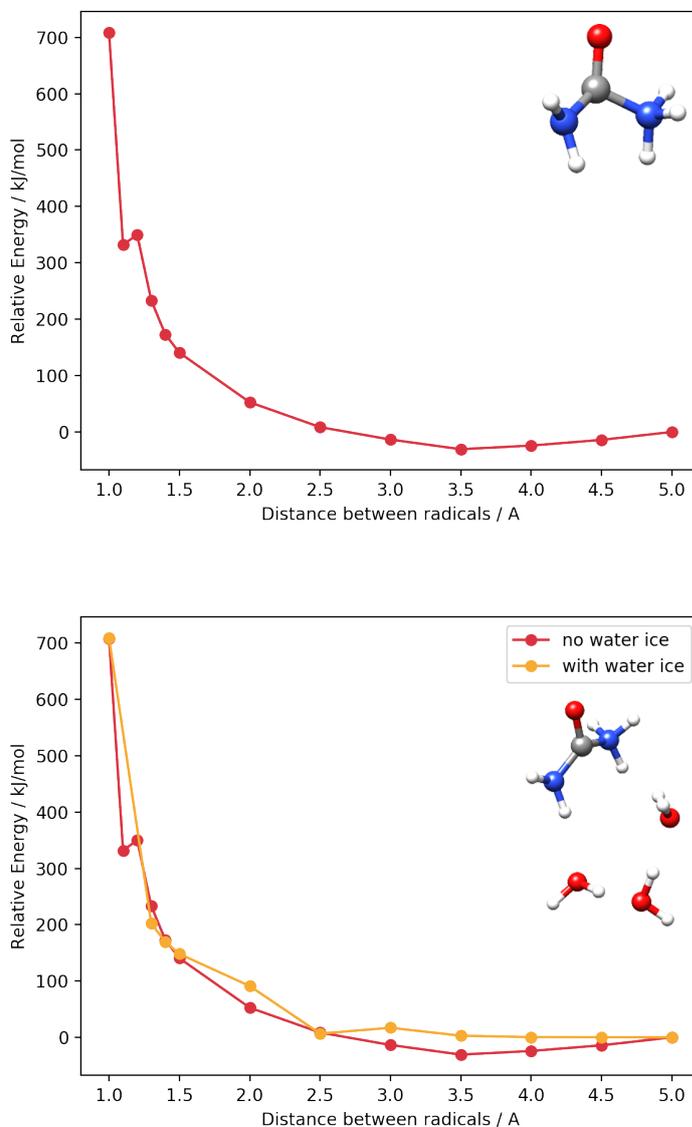
Figure 6.8: Top: The potential energy curve for the addition of ammonia to the formamide radical in the gas-phase as calculated by DFT to give urea. Here, the energy at a radical distance of Å is taken as the zero. Inset: the structure of the species formed at the interspecies distance closest to the length of the CN bond in urea. Bottom: The potential energy curve for the addition of ammonia to the formamide radical, as calculated by DFT, to give urea. Curves with and without three water molecules acting in place of an icy dust grain mantle are shown and the energy at a radical distance of Å is taken as the zero. Inset: the structure of the species formed at the interspecies distance closest to the length of the CN bond in urea.
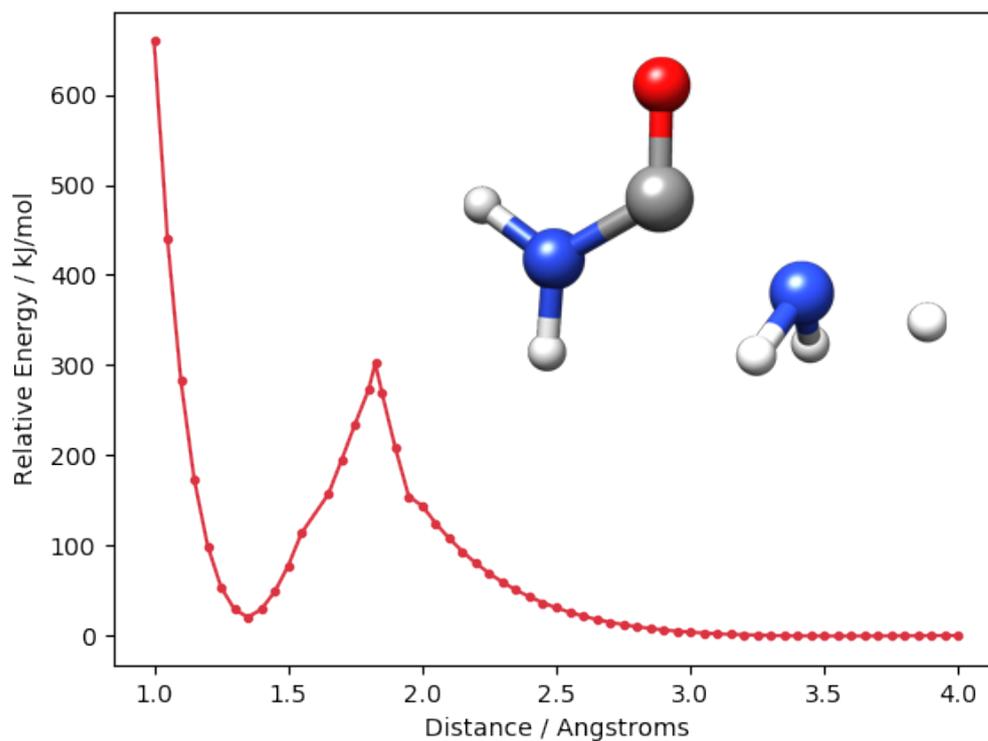
Figure 6.9: The potential energy curve for the addition of ammonia to formamide radical in the gas-phase to give urea as calculated by RASSCF where the energy at a radical distance of 4 Å is taken as the zero. Inset: The optimised structure at an inter-reactant distance of 1.825 Å corresponding to the peak of the barrier in the potential energy curve calculated by RASSCF for formamide radical reacting with ammonia

the small barrier at 2.8 Å. The increase in energy from 2.7 Å onwards in caused by the incoming amino radical distorting the shape of formamide and thus breaking the π delocalisation. The peak of this energy increase lies 150 kJ mol^{-1} above the initial, well-separated species. The second dip in energy at 1.5 Å represents a high energy intermediate and does not correspond to the formation of urea. The optimised structure at this point is shown in the inset in figure 6.10.

As can be seen, the formamide CH bond does not cleave to yield the target product of urea. Scanning the CH bond in this intermediate gave the potential energy curve in figure 6.11. This process displays another barrier of 108 kJ mol^{-1} which corresponds to the point at which the switch is made from tetrahedral intermediate to the planar structure seen in urea. The curve then plateaus after the formation of urea and a hydrogen radical at 2.8 kJ mol^{-1} with the products being 18 kJ mol^{-1} less stable than the tetrahedral intermediate species. Both the barrier and the endothermicity of the reaction make this an unlikely formation route under ISM conditions.

6.3.3 Radical-radical pathways

The variety of substantial barriers to the formation of urea from formamide and an amino radical make this an unlikely astronomical formation pathway. There is a further option for using formamide and ammonia based reactants to yield urea which is to use both a formamide radical and an amino radical. Radical-radical reactions contribute to the presence of many COMs in the ISM, thus it makes sense to consider this candidate mechanism.[300,301] Multiconfigurational methods were used once more to calculate the energy of the system whilst scanning over the CN distance between the radicals.

The singlet state potential energy curve is shown by the blue trace in figure 6.12. The well at 1.4 Å corresponds to the formation of urea via radical-radical recombination and has a depth of roughly 400 kJ mol^{-1}, which is in line with experiment for the formation of a CN bond. The barrier at 2.4 Å is caused by the rearrangement of the formamide radical, which at long inter-radical distances had optimised to a linear isocyanic acid type structure, see inset in figure 6.12.

Given that the unpaired electrons on each radical can take either up or down spins, the radicals may not be in the singlet spin state and so a second triplet state scan was also performed. The results of this scan are denoted by the yellow trace in figure 6.12. The triplet scan does not yield urea like the singlet state scan does, but behaves much better at long range. If the two curves are overlaid using the triplet results at long range and the singlet at short range, then urea can be formed with a minimal barrier of only 4 kJ mol^{-1}. Providing the change
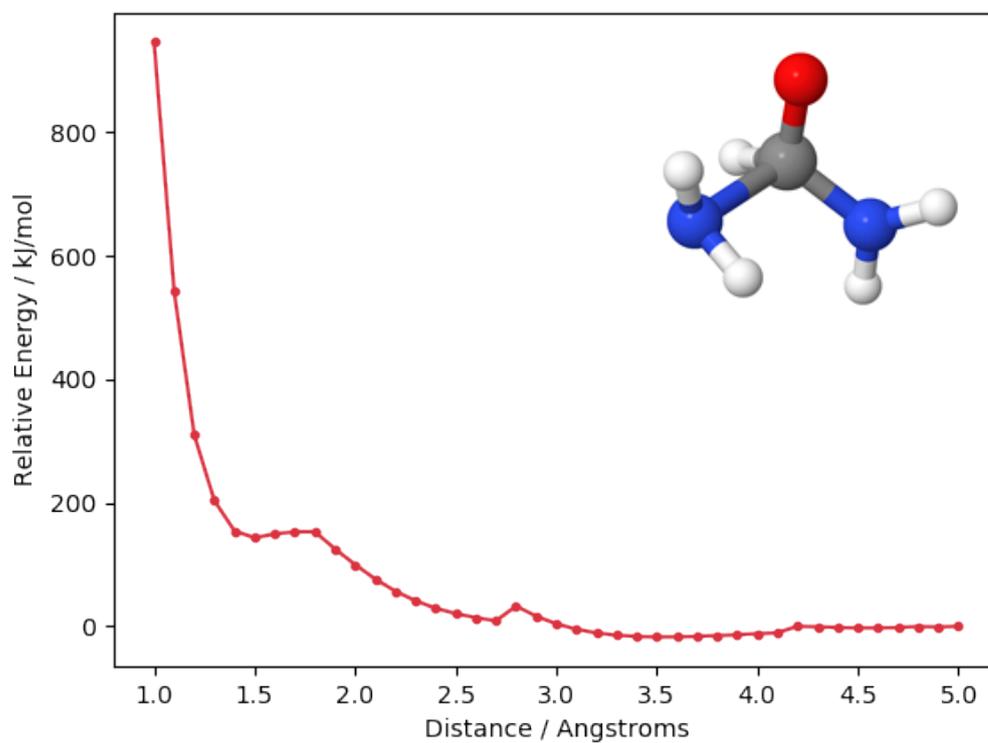
Figure 6.10: The potential energy curve for the addition of an amino radical to formamide in the gas-phase to give urea as calculated by RASSCF where the energy at a radical distance of 5 Å is taken as the zero. Inset: The optimised structure of the intermediate at 1.5 Å in the potential energy curve for the formation of urea from an amino radical and formamide.
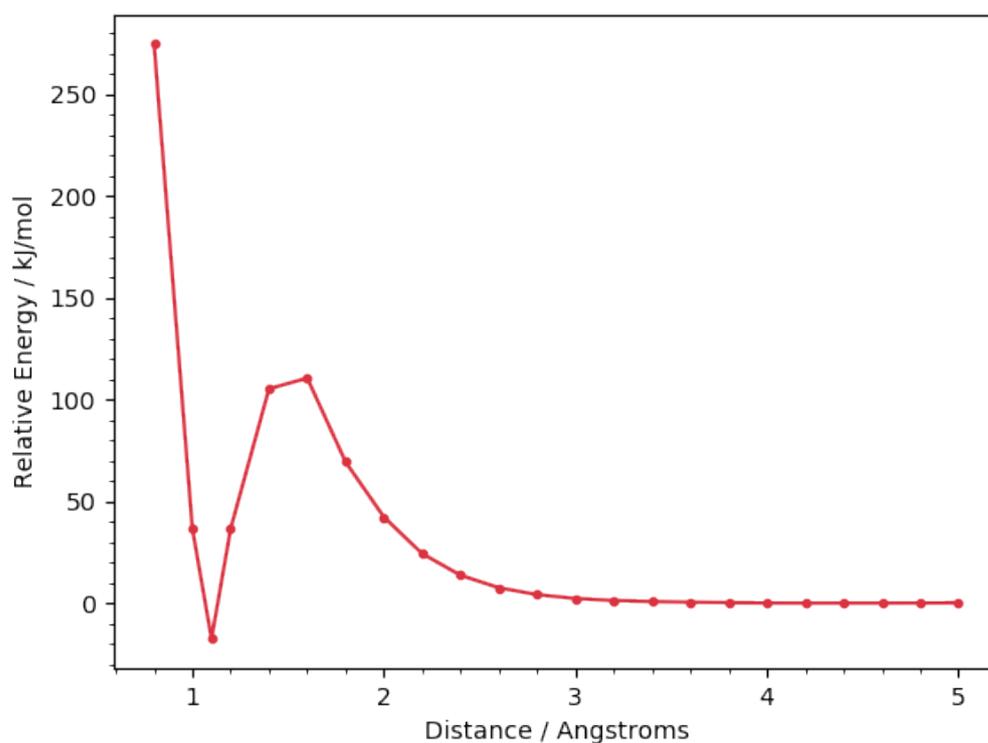
Figure 6.11: The potential energy curve for the dissociation of the CH bond in the high energy intermediate given from the reaction between formamide and an amino radical. Here, the energy of the system at a separation of 5 Å taken as the zero.
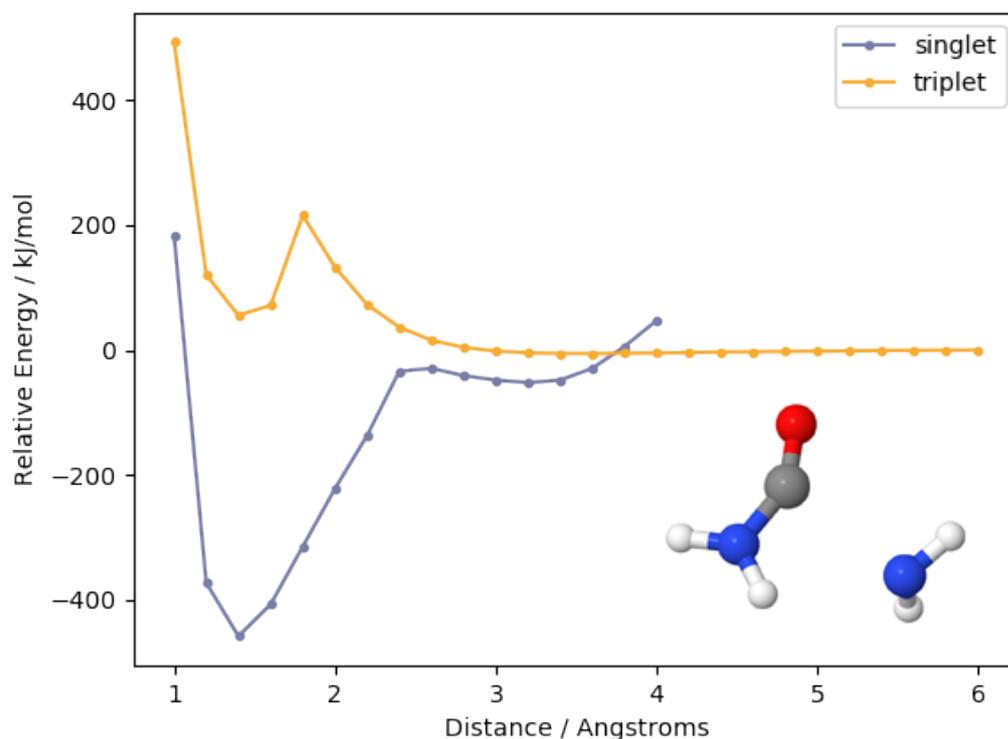
Figure 6.12: The potential energy curve for the addition of an amino radical to a formamide radical to give urea as calculated by RASSCF in the singlet and triplet states with the energy of the triplet system at a 5 Å separation taken as the zero. Inset: The hydrogenated isocyanic acid like structure seen in the system at long separations.

in spin was allowed via some mechanism, this pathway presents the most facile means by which to produce urea in the ISM from radical species generated from formamide and ammonia. There is of course no guarantee that it would be, and even if the magnetic fields of the ISM could achieve this, the likelihood of this pathway being facilitated is low at best. Consequently, further alternatives which do not use radicals were probed.

6.3.4 Charged pathways

As the formamide radical was observed to rearrange to an isocyanic acid type structure, the use of isocyanic acid as a reactant in the formation of urea was also investigated. The long range structure observed in the formamide radical and ammonia calculations was that of hydrogenated isocyanic acid. Additionally, recent work by Brigiano *et al* in which they probed various urea formation routes found the most facile to be a route using an ionic species.[302] Thus, protonated isocyanic acid was used in these calculations. Therefore the first step that was

elucidated using scan type calculations is the *protonation* of isocyanic acid. There are multiple options for where the protonation can occur: on the nitrogen, the carbon or the oxygen atom. The potential energy curves for protonating on each of these atoms is shown in figure 6.13.
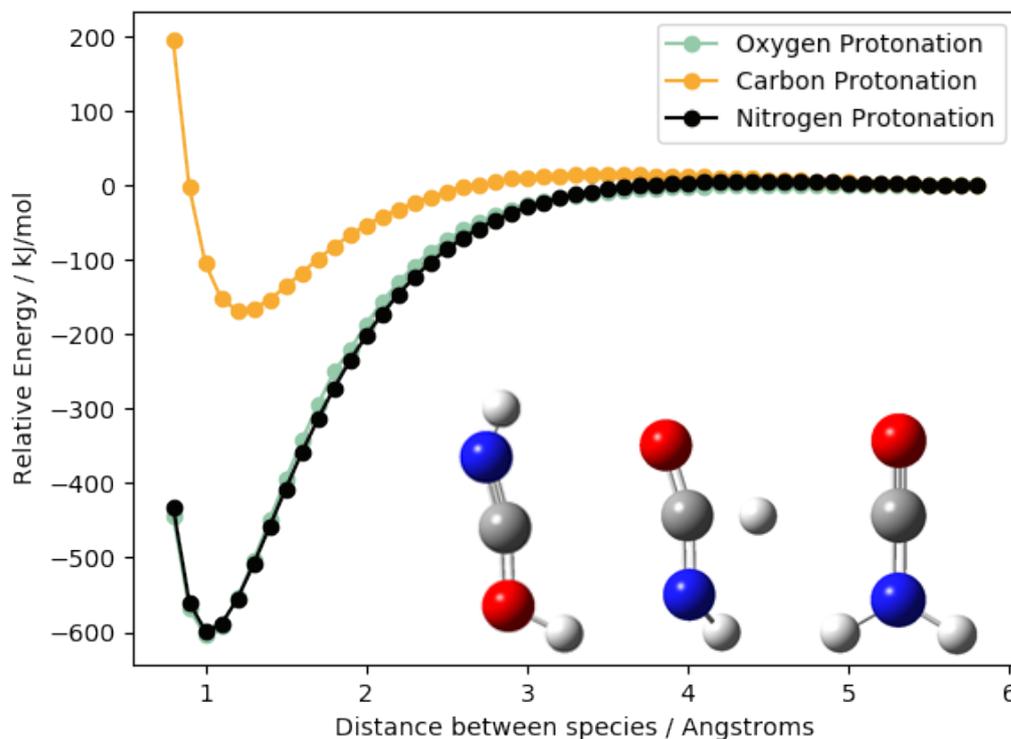
Figure 6.13: The potential energy curves for the addition of a proton to the oxygen, nitrogen and carbon atoms in isocyanic acid with the lowest energy structures for, left to right, the oxygen protonated system, the carbon protonated system and the nitrogen protonated system shown.

For protonation on nitrogen and oxygen the process is barrierless and the well depths are 603.1 and 603.9 kJ mol^{-1} respectively. Protonation on the carbon, however, is not barrierless, but requires an energy input of 14.4 kJ mol^{-1}. It also results in a much less stable species, with a well depth of only 172.8 kJ mol^{-1}, making it the least likely of the three pathways. Of the remaining two protonation reactions nitrogen protonation would appear to be the major route. This is the product that most naturally leads on to the formation of urea as it creates one of the amino groups found in urea. Nonetheless, the similarity in energies for the nitrogen and oxygen protonated species mean the possibility cannot be discounted that the oxygen protonated species can also be formed.

Consequently, the addition of ammonia to each of these species has been probed. Figure 6.14's upper panel shows that, if ammonia reacts with nitrogen

protonated urea, the addition proceeds barrierlessly and has a well depth of 171.9 kJ mol^{-1}. The alternative reaction, where ammonia simply deprotonates the isocyanic acid species is also barrierless but has a much shallower well of 57.4 kJ mol^{-1} making the addition reaction the more energetically favourable of the two. The turquoise trace in figure 6.14 shows the reaction between ammonia and the oxygen protonated variant of protonated isocyanic acid; this reaction is also barrierless, but ammonia appears to bind more weakly. The protonated urea product is only 85.1 kJ mol^{-1} lower in energy than the well-separated reactants (figure 6.14) and the angle between the incoming ammonia and the isocyanic acid remains wider than the equilibrium angle.

It should be noted that the oxygen protonated species would also have to undergo tautomerisation (figure 6.15 upper) after the loss of a proton from the ammonia, to yield the target product urea. If this tautomerisation goes *via* a one step process then there is a barrier of 130 kJ mol^{-1} which is prohibitively high. If the tautomerisation occurs via a series of steps as outlined in figure 6.16 then there are no barriers involved but energy in the form of radiation is required to dissociated the OH bond which has a calculated dissociation energy of 636.7 kJ mol^{-1}, consistent with this type of bond.[303] The conclusion therefore is inevitable that the oxygen protonation route is not available to proceed with.

Just as the oxygen protonated species needs to undergo deprotonation to form urea, so does the nitrogen protonated species. This process, the potential energy of which is given in figure 6.17, is barrierless and has a calculated dissociation energy of 636.7 kJ mol^{-1}. As with the oxygen protonated species this bond could be broken by the high energy radiation present in interstellar clouds but would require the protonated urea species to be in an area of the cloud penetrated by such radiation.

There is also the possibility that including ice species capable of hydrogen bonding in the system will help mediate the deprotonation and so reduce the energy input require. In order to test this an analogous scan calculation was performed but with two water molecules included to act as proton acceptors. As can been seen in figure 6.17 the presence of water vastly reduces the amount of energy required to deprotonate urea to only 7.8 kJmol^{-1}. When the ice is included in the system the lowest energy structure actually becomes neutral urea with the proton shared between the water ice molecules, see figure 6.18. This strongly suggests that the deprotonation would be essentially barrierless under these conditions. If this is the case, then the use of charged species to form urea becomes feasible.
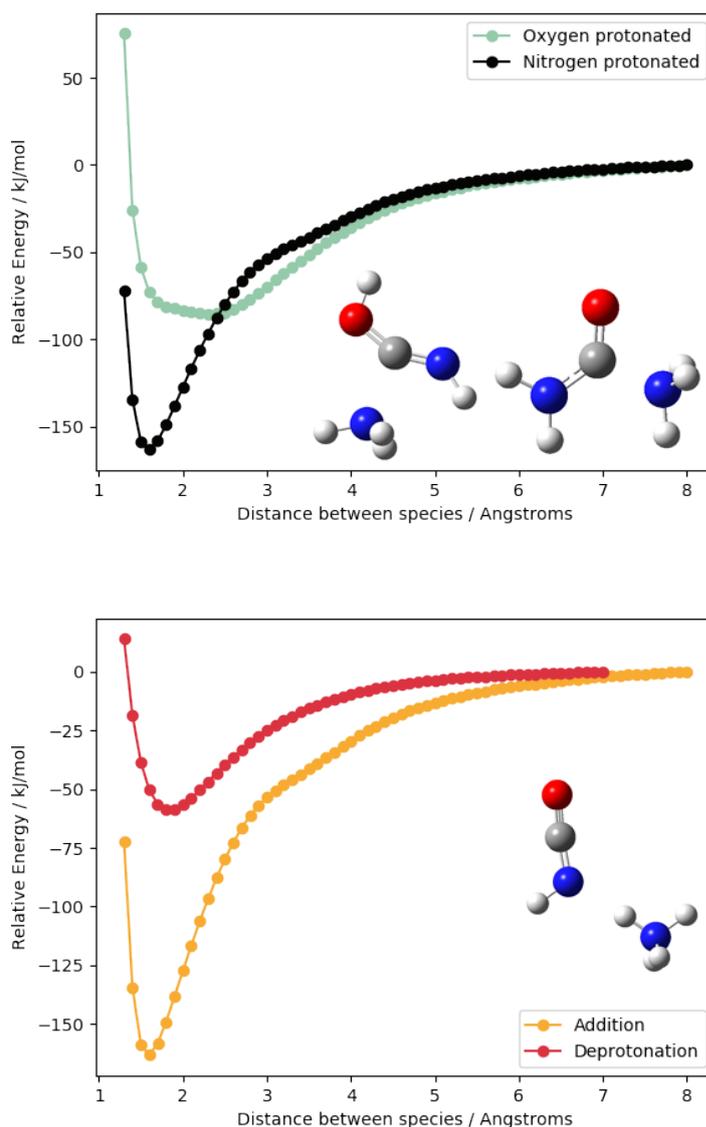
Figure 6.14: Upper: The potential energy curves for the addition of ammonia to the oxygen protonated isocyanic acid and the nitrogen protonated species with the lowest energy structures for each system shown. Lower: The potential energy curves for the addition of ammonia to the nitrogen protonated species and the deprotonation of the nitrogen protonated isocyanic acid by ammonia with the lowest energy structure for the deprotonation shown.

6.4 Conclusions

The calculations that were carried out to probe the feasibility of the mechanism for the formation of urea proposed by Förstel *et al.*, equations 6.2 and 6.3, suggest that these are not an ISM compatible route.

$$NH_2CHO + NH_3 \longrightarrow (NH_2)_2CO + H_2 \qquad (6.2)$$

125
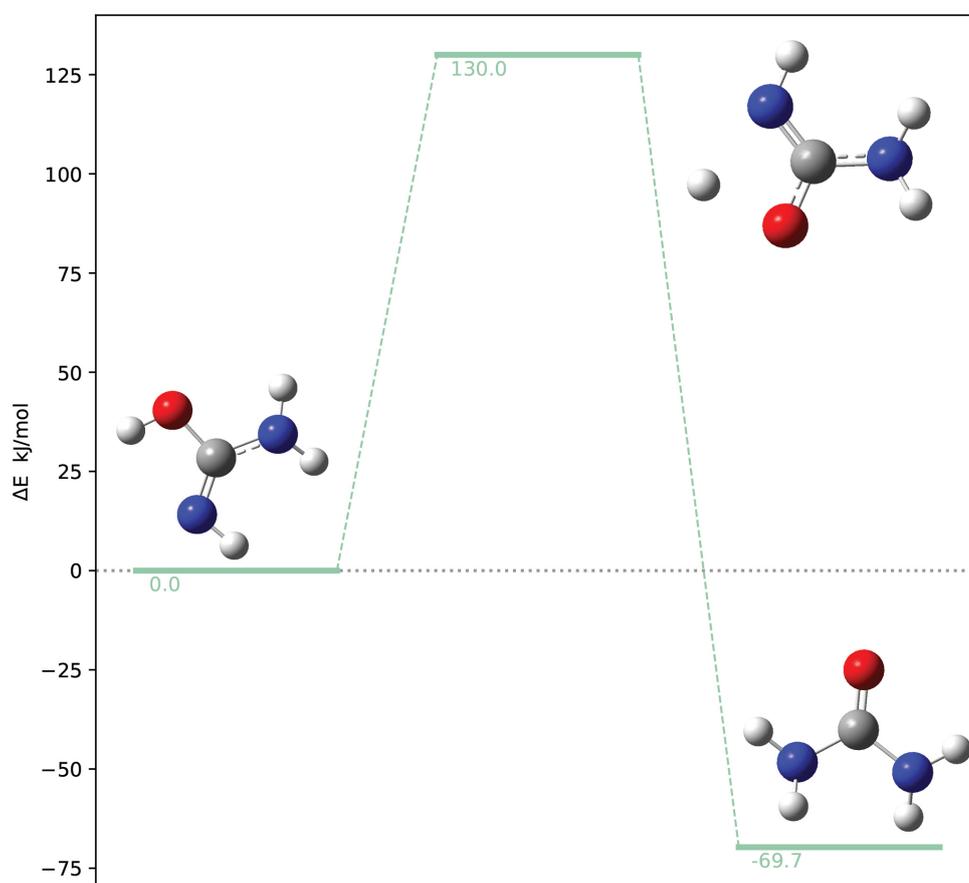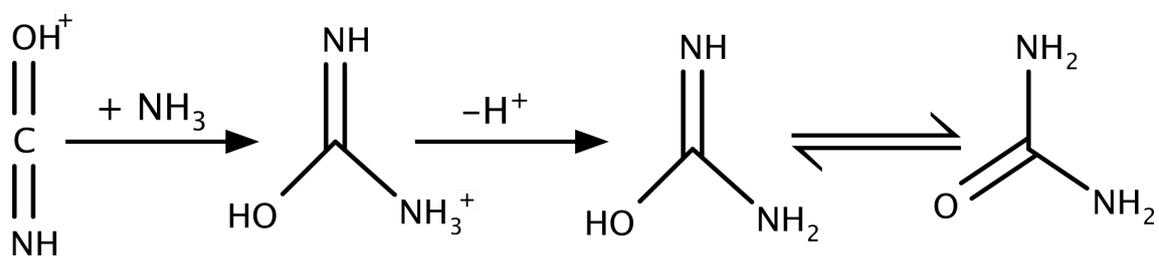
Figure 6.15: Upper: The reaction scheme for the addition of NH_3 to oxygen protonated isocyanic acid and the subsequent tautomerisation of the product to give protonated urea. Lower: The barrier to tautomerisation as calculated using DFT.
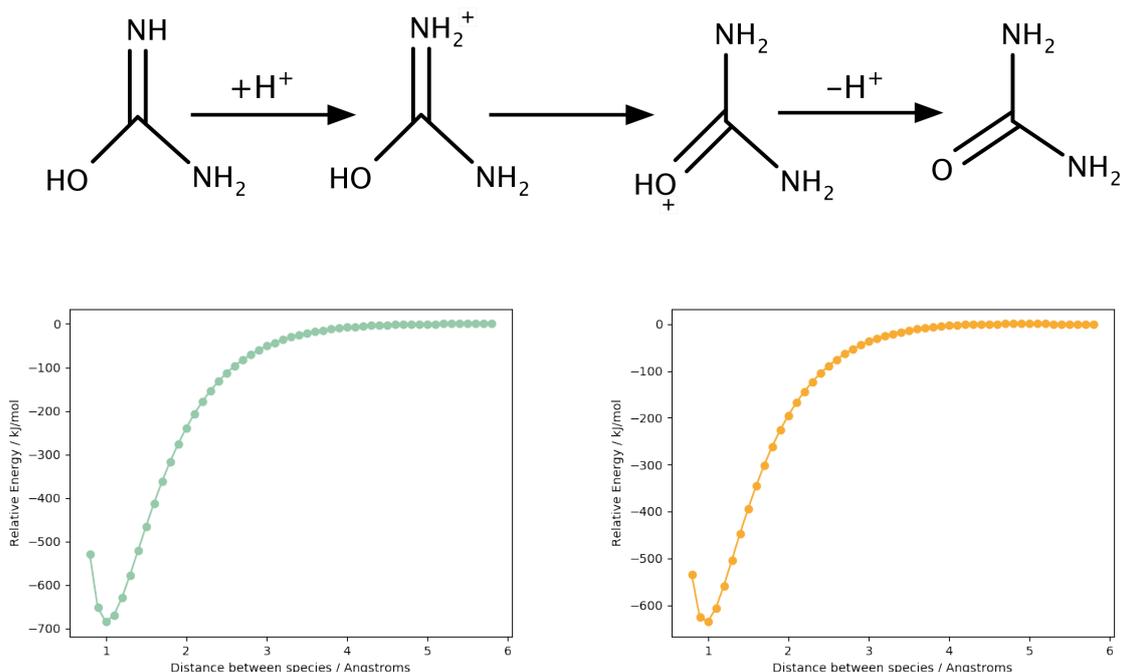
Figure 6.16: Upper: The alternative route for the tautomerisation needed to form protonated urea from oxygen protonated iscyanic acid. Lower: The potential energy curves for (right) the protonation of the imine group and (left) the deprotonation of the carbonyl as calculated using DFT and taking the energy at an inter species distance of 5 Å as the zero.

$$NH_2CO\cdot\ +\ NH_3 \longrightarrow (NH_2)_2CO + H\cdot \tag{6.3}$$

The use of neutral formamide and ammonia does not present itself as a possibility, as the calculations in the gas-phase found this route failed to give a potential energy curve indicating urea was a favourable product. Inspection of the optimised structures showed that this was due to a lack of bond cleavage for the CH bond on formamide and one of the NH bonds in ammonia. In order to create neutral urea, these bonds need to break to give the side product dihydrogen. Including water in the system as an ice species did not facilitate the bond breaking *via* intermolecular interactions. This may have been either due to an insufficient quantity of water in the system or this mechanism simply is not feasible.

When single point CCSD(T) energy calculations were performed for the reactants and products of equation 6.2, the reaction was found to endothermic by 26 kJ mol^{-1}. Zero point energy effects lowered this to 8 kJ mol^{-1}; however, this does not take any potential transition state into account and given the number of bonds breaking during the reaction, this transition state is expected to be substantial. The nature of the ISM therefore makes this particular formation reaction unlikely. The radical route, equation 6.3, presents a much more feasible pathway. The formation of the formamide radical, as calculated by DFT, was found to be

127
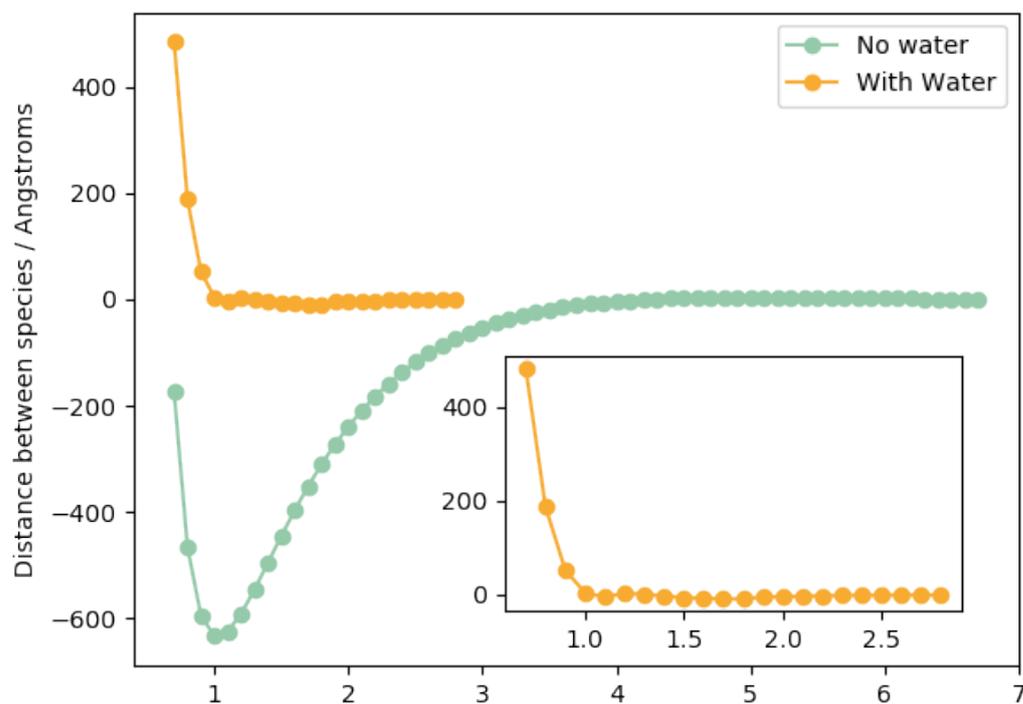
Figure 6.17: The potential energy curve for the deprotonation of protonated urea as calculated using DFT with and without water.
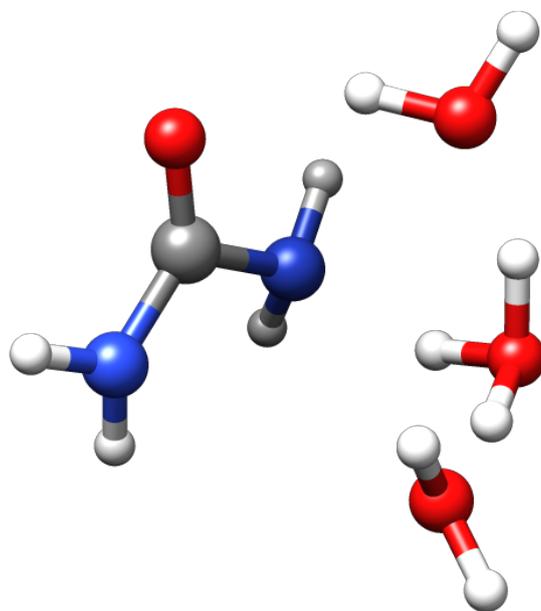
Figure 6.18: The lowest energy structure for protonated urea with water ice included in the system, showing how the water ice helps reduce the energy input required to deprotonate the urea by sharing the removed proton between the water molecules.

barrierless and the dissociation energy could be overcome by high energy UV radiation. Thus, providing the formamide radical could be generated in an area of the interstellar cloud where UV radiation penetrates, the reactants for this step could exist in the ISM. The subsequent reaction between the formamide radical and ammonia is again poorly described by DFT, which does not indicate the reaction is feasible, with the products being higher in energy than the reactants. Multiconfigurational calculations showed protonated urea to be a stable product but found a barrier to the reaction of 282 kJ mol^{-1}. This size of barrier would not be overcome at ISM temperatures.

In light of this, an alternative radical-neutral reaction, equation 6.4, was also tested using multiconfigurational methods.

$$NH_2CHO + NH_2 \cdot \longrightarrow (NH_2)_2CO + H \cdot \qquad (6.4)$$
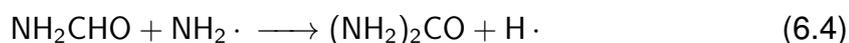
However, this was found to progress through a high energy tetrahedral intermediate giving the reaction a barrier of 150 kJ mol^{-1}. While this barrier is lower than that for the previous radical-neutral reaction it is still too high for this to represent any kind of major formation route for urea in the ISM.

The most promising radical containing reaction for the formation of urea from formamide and ammonia is the radical-radical reaction given in equation 6.5.

$$NH_2CO \cdot + NH_2 \cdot \longrightarrow (NH_2)_2CO \qquad (6.5)$$
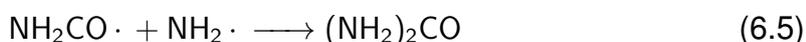
This reaction can progress with a barrier of only 4 kJ mol^{-1}, one low enough that it is not prohibitive under ISM conditions, provided the system can change from triplet to singlet multiplicity as the distance between the two radicals decreases. The likelihood of this being able to occur is contingent on a very specific set of circumstances, particular with regard to the allowing of the spin flip from triplet to singlet, means that it would appear that radical chemistry may not contribute much to the formation of urea in the ISM. Therefore, other, non-radical containing, routes need to be sought. This is why alternatives using charged species were tested for their viability.

The use of protonated isocyanic acid and ammonia to form urea was also investigated and it was found that the most likely pathway used isocyanic acid protonated on the nitrogen. This then reacts barrierlessly with ammonia to form protonated urea. The subsequent deprotonation of this species requires an energy input of 636.7 kJ mol^{-1} to cleave the NH bond. It is possible that photodissociation of this bond could occur. An alternative reaction using oxygen protonated isocyanic acid was found to be unlikely due to the weak binding of the incoming ammonia and the energy required for the tautomerisation to yield urea.

The inclusion of water ice into the system when protonated urea is transformed into the neutral species lowered the energy barrier significantly. In fact, neutral urea with protonated water was found to be lower in energy than the initial protonated urea containing system. The stabilisation afforded by the water ice's ability to hydrogen bond makes this route to urea ISM feasible. Overall, the use of charged species to form urea in the ISM is the most plausible route to urea in an astrochemical setting.

7. The formation of ribose in interstellar ices

Another key element for the origins of life are sugars, specifically ribose which in its deoxygenated, cyclic form makes up part of the genetic information storing RNA and DNA. Chapter 6 has already examined how a precursor to nucleotides is made, another key element of DNA and RNA. Thus, understanding how ribose could form outside of biological processes could shed light on the how life began on early Earth. In this section the formose reaction, which is one possible formation route for ribose and other sugars, is examined.

7.1   Sugars on Earth

One suggestion for the origins of life on early Earth is the RNA world hypothesis, a subject that has already been covered during section 1.1.6 of the Introduction.[304–306] This posits that genetic material in the form of RNA (ribonucleic acid) was formed abiotically and this then formed the basis for the beginnings of the biological processes that led to life. For this to hold true the constituent parts of RNA must have formation routes that could feasibly occur under ISM and/or early Earth conditions. One such constituent piece of RNA is the sugar ribose in its 2-deoxy form, shown in figure 7.1.
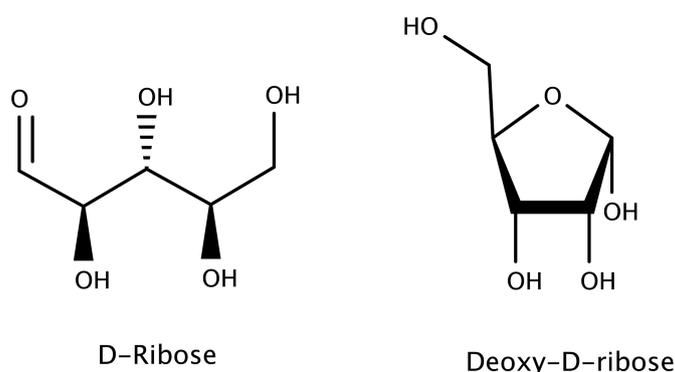
Figure 7.1: The structures of D-ribose (left) and deoxy-ribose (right).

Sugars play an incredibly important role in biology; they are present in genetic code, and make up part of plant cell walls and act as energy storage molecules.[307,308] Sugars and sugar acids, monosaccharides with a carboxylic acid group, with up to six carbon atoms have been detected on the Murchison and Murray meteorites, suggesting that they could have an extraterrestrial origin and have been brought to Earth during cometary impacts.[309,310] This raises the question of what exactly that extraterrestrial formation route is. It has long been known that sugars can be produced abiotically, *via* the formose reaction.[311,312] The formose reaction is a series of aldol reactions starting with formaldehyde and glycolaldehyde, both of which have been confirmed in the ISM.[125,313] A reaction scheme for the formose reaction is given in figure 7.2.
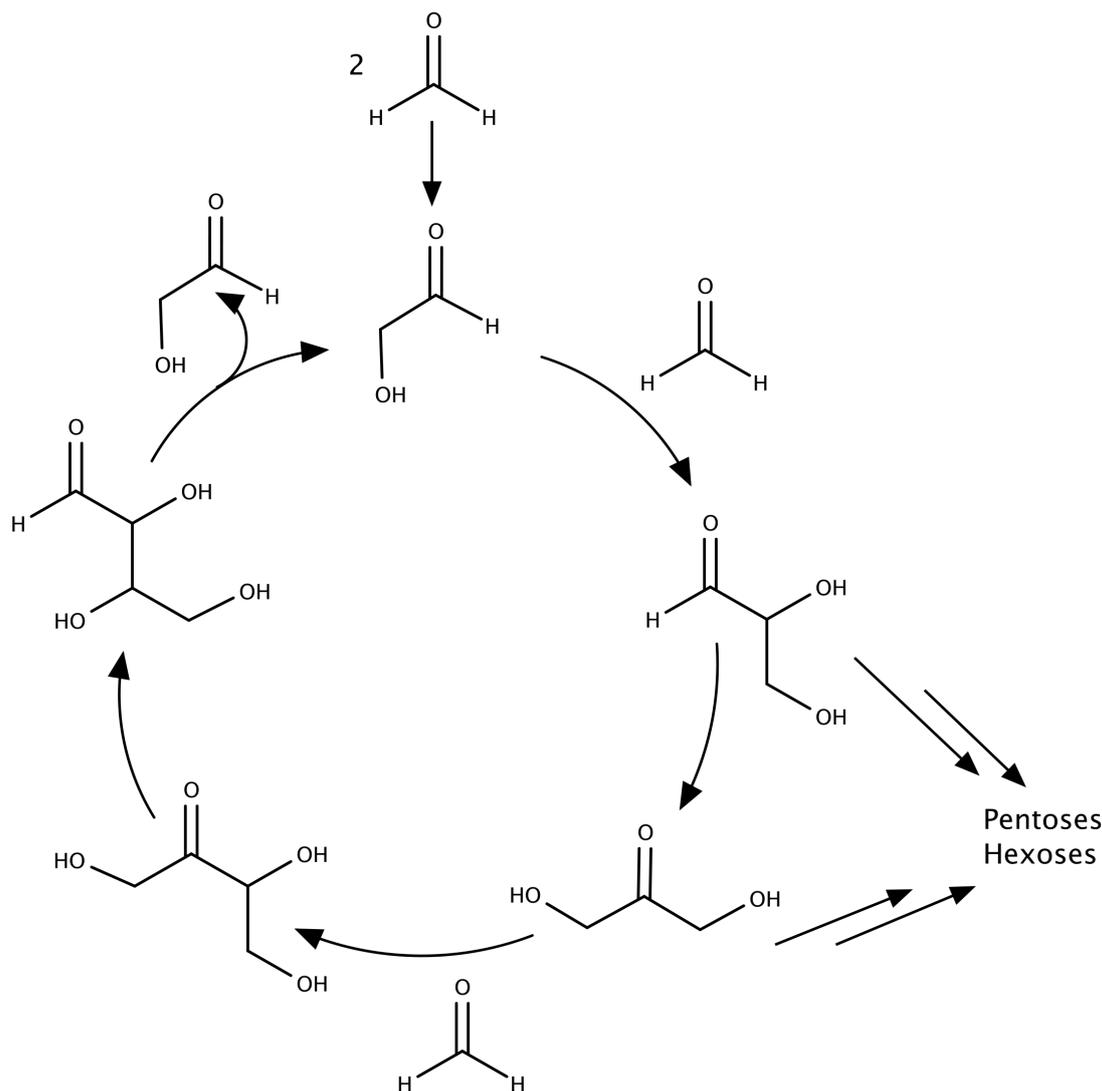
Figure 7.2: The formose reaction scheme as proposed in 1959 by Breslow for the autocatalysis of formaldehyde to produce sugars from trioses up to hexoses.[312]

It has been shown that the formose reaction can be initiated by UV radiation (λ = 308 nm), something that is present within the dust clouds of the ISM.[314] In a cryogenic argon matrix the photoreaction of formaldehyde has been shown to lead to glycolaldehyde.[315–317] Similar behaviour has been observed in model ISM ice systems with a water and methanol mix producing sugars, sugar alcohols and sugar acids when irradiated with UV.[224,318,319] An alternative mix of methanol and carbon monoxide can also lead to sugar formation, with carbenes being suggested as an important intermediate species.[318,320–322] However, it is worth noting that all of these experiments have required energy input, generally from UV photons, in order to drive the reaction forward.

Additionally, a variety of experiments have elucidated possible routes to ribose and other tetroses and pentoses under prebiotic conditions suitable for early

Earth.[323–329] Amino acids have been shown to act both as catalysts for synthesising carbohydrates from simple aldehydes and promoters for aldol reactions to form tetroses.[315,316,330] Amino compounds also potentially hold a route to the kind of enantioselective reactions that lead to only D-sugars being used in DNA and RNA. Amino nitriles can act as promoters that lead to the formation of 2-deoxy-D-ribose from aldol reactions under plausible prebiotic conditions.[331]

The formose reaction has been put forward as the means for forming ribose in an interstellar setting. However, it has not yet been confirmed as the mechanism in operation. In some ice-irradiation experiments the diversity of products yielded is consistent with the formose reaction.[224] The products detected in other ice studies suggest that there must be formation routes at play beyond just formose-type reactions.[319] It has been suggested that photochemistry induced radical reactions could also be responsible for some of the sugars observed, as irradiation of a methanol-ammonia mixed ice led to the formation of glycerol, the simplest sugar molecule.[301,332] Another issue with the formose reaction is that it produces linear ribose and not the cyclic deoxy-ribose found in RNA. Photoredox chemistry has been demonstrated to reduce ribose to deoxyribose, albeit in aqueous solution as opposed to within ISM ices, which suggests that it would be possible to form deoxy-ribose from the products of the formose reaction.[333]

The aim of this work was to make use of computational methods to outline whether or not the formose reaction is feasible under ISM conditions, by examining the energy barriers involved within the successive reaction steps.

7.2 Methodology

All calculations were performed using the GAUSSIAN 09 suite of electronic structure programmes* using DFT at the B3LYP level[148,149,233] using a 6-311G** basis.[234,292] Solvent was added to the systems using the Integral Equation Formalism Polarised Continuum Model (IEFPCM)[334] using water as the bulk solvent.

The geometries for all reactants and intermediates in the formation of ribose from the formose reaction were optimised to give no imaginary frequencies. These optimised geometries were then used to perform rigid and relaxed scans along the forming or breaking bond in each reaction step to locate possible transition state geometries. The scans were run using a step size of 0.05 Å. Once possible transition states were located, QST3 calculations were carried out to optimise the transition state. These structures were optimised until only one imaginary frequency remained that corresponded to the bond breaking or formation in that particular reaction step.[335] Intrinsic Reaction Coordinate (IRC) calculations then

---

*GAUSSIAN 09, revision D.01, M. J. Frisch et al., Gaussian, Inc., Wallingford, CT, 2009

used these transition state geometries to confirm the correct transition state was reached and to find the energy barrier to the reaction.[336,337]

7.3   The formose reaction

Figure 7.3 shows a potential route to form ribose through the formose reaction, starting from methanol and formaldehyde derivatives. The mechanism consists of nine steps, and is largely composed of successive aldol reactions to increase the carbon chain. An additional step, D, not included in this investigation as it is the protonation of formaldehyde and protonated formaldehyde has been detected in the ISM.[338,339] Each step has been labelled with a letter for ease of identification.

The reaction profile for the entire mechanism is given in figure 7.4 and shows that, overall, the reaction is energetically favourable, with all barriers being submerged below the energy of the initial reactants. This is in line with work by Jalbout *et al.* who explored tetrose formation from glycolaldehyde and found their proposed mechanism to be overall energetically favourable.[340] That being said, not all the steps in the mechanism proposed here are barrierless. In some cases, steps B and H, the steps are endothermic. For some steps both a *cis-* and *trans-* orientation of the growing sugar molecule are possible. In these cases both isomers have been investigated. All steps will be discussed in more detail subsequently.

7.3.1   Step A

Step A is the radical-radical recombination of $HOCH_2$ and HCO to form glycolaldehyde. Both reactant species can be made from confirmed ISM molecules, namely methanol and formaldehyde.[196,341] This reaction has been investigated in Chapter 5 and has been shown to proceed barrierlessly and exothermically, giving out on the order of 500 kJ mol^{-1} in energy. This amount of energy being released, should it remain localised around the reactant molecules and not immediately be radiated away by the ice structure, could help overcome any barriers present in early subsequent steps of the mechanism.

7.3.2   Steps B to F

Step B is the protonation of the newly formed glycolaldehyde, which could be in either the *cis-* or *trans-* orientation. *Cis*-glycolaldehyde has been confirmed in the ISM and is the more stable isomer due to intramolecular hydrogen bonding, while *trans-* has only been detected tentatively.[254,342] Both isomers exhibit a
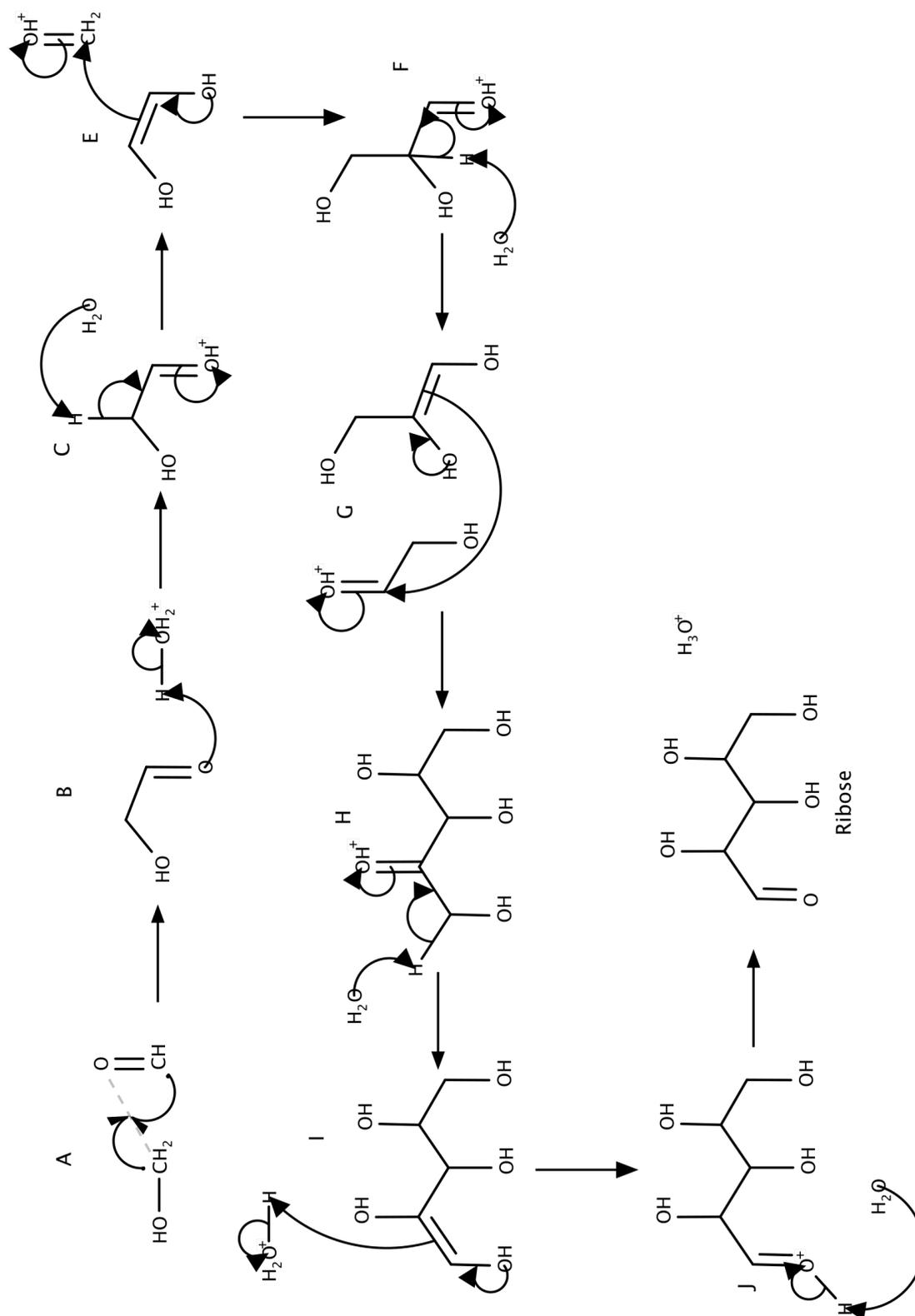
Figure 7.3: The reaction mechanism for a plausible route to the formation of ribose *via* the formose reaction where each step is labeled with a letter, this naming scheme will be maintained throughout. Step D, not shown, is the protonation of formaldehyde.[312]
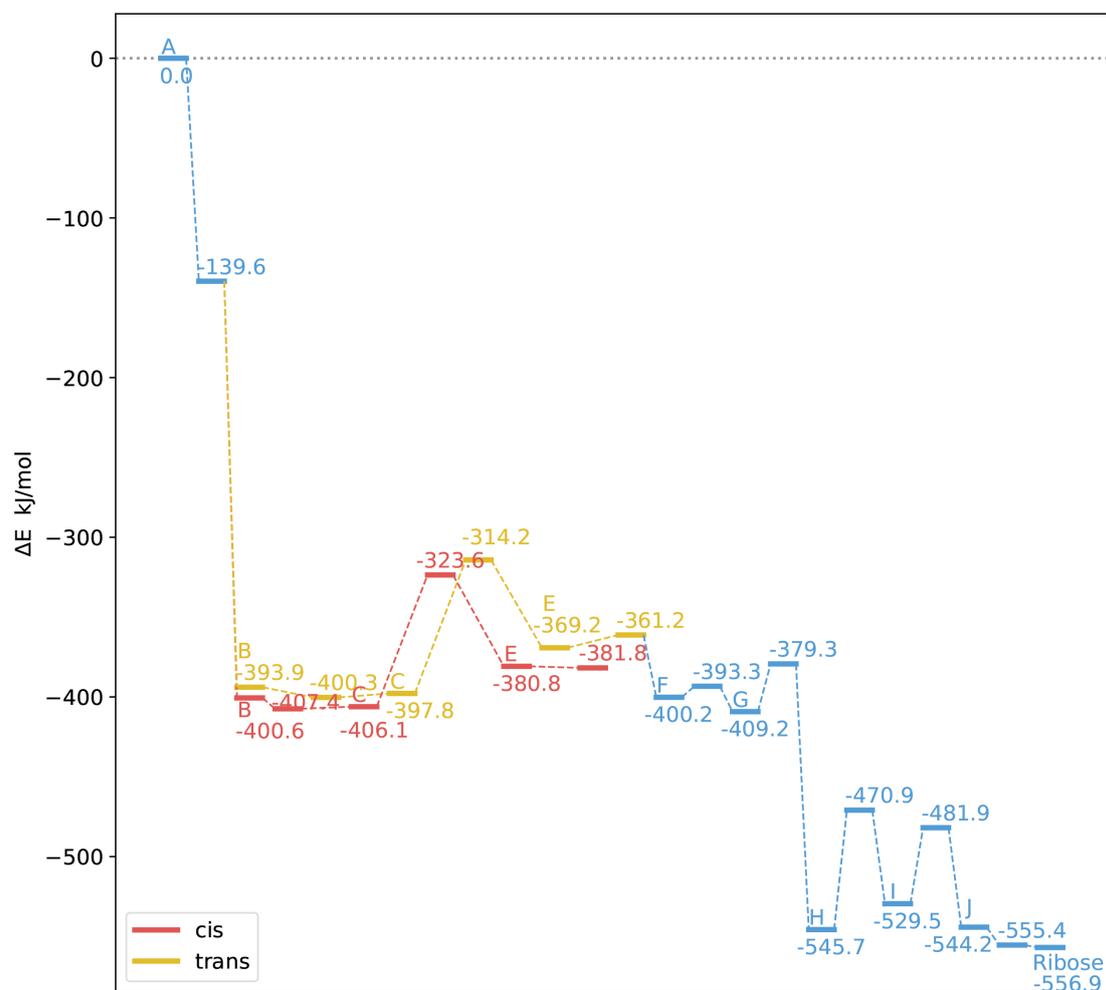
Figure 7.4: The reaction profile for the formation of ribose from the formose reaction as outlined in figure 7.3. All energies are plotted relative to the energy of the reactants of step A (the radical-radical reaction between $HOCH_2$ and HCO) and the energies of all steps include all species present in the entire reaction to maintain consistency. Blue denotes reactions in which no specific isomer was sought whereas the red and yellow denote the specific use of the *cis* and *trans* isomers respectively.

low energy pre-reaction complex which corresponds to the sharing of the proton between the water oxygen and the glycolaldehyde oxygen, see figure 7.5. The fully protonated glycolaldehyde species is lower in energy than the reactants for both isomers (by 5.5 kJ mol^{-1} for *cis-* and 12.2 kJ mol^{-1} for *trans-*) and only marginally higher in energy (1.3 kJ mol^{-1}) than the pre-reaction complex. The energy released by step A would be more than sufficient to overcome this slight endothermicity. Additionally the *cis-* isomer remains lower in energy than the *trans-* throughout steps B to F.

Step C, the enolisation of protonated glycolaldehyde by water, exhibits a barrier for both isomers and is endothermic, as shown in figure 7.6. The barrier for both the *cis-* and *trans-* case corresponds to the switch between an sp^3 and an sp^2 geometry around the carbon being deprotonated. For the *cis-* isomer the barrier height is 82.4 kJ mol^{-1} and for the *trans-* it is 83 kJ mol^{-1}. Depending on whether the system has equilibrated, the energy to overcome these barriers may come from the initial radical-radical recombination or from UV photons. The use of UV radiation to form ribose amongst other sugars in experimental settings implies that this is, in principle, plausible.[224]

An alternative possibility is that the ice mantle could help to stabilise this reaction and thus lower the energy barrier. Given that this step involves the removal of a proton, an extended network of hydrogen bonding capable molecules would allow for greater distribution of the charge, thus helping stabilise the process. To model this pathway *cis-* and *trans-*glycolaldehyde were reoptimised with three surrounding water molecules. The transition state structures and the product, enolised glycolaldehyde, were also reoptimised with three water molecules present. Figure 7.6 shows that the addition of water ice lowers the barrier substantially for both isomers and also stabilises the product to such a degree that the reaction becomes exothermic. This reaction still is not barrierless but a lower barrier increases the likelihood that photons, for example, will be able to supply the energy to overcome it or that the reaction may occur due to tunnelling. Inspection of the optimised geometries shows that the stabilisation does stem from the sharing of the removed proton between the water molecules rather than just one adopting a formal positive charge. The geometries of the water around the transition states can be seen in figure 7.6. From the geometries of the system it is clear to see that a hydrogen bonding chain forms between the glycolaldehyde and the water molecules which distributes the charge between the molecules and stabilises the species present.

Step D is not shown in figure 7.4 but is the protonation of formaldehyde. Protonated formaldehyde has been detected in the ISM and so could be available as a reactant.[338,339] Step E is the aldol reaction between protonated glycolaldehyde
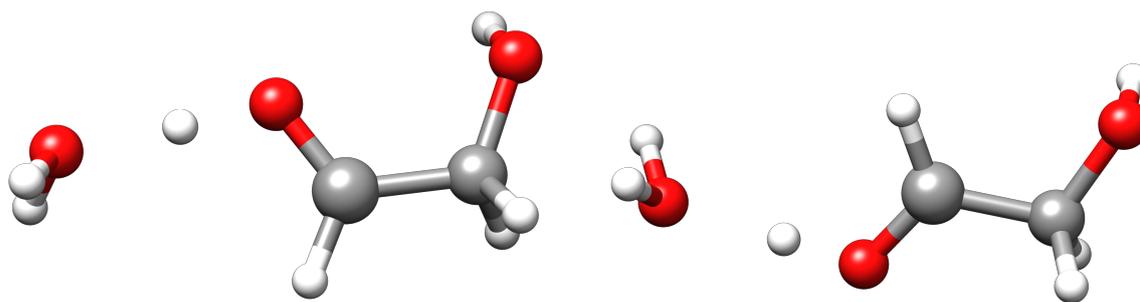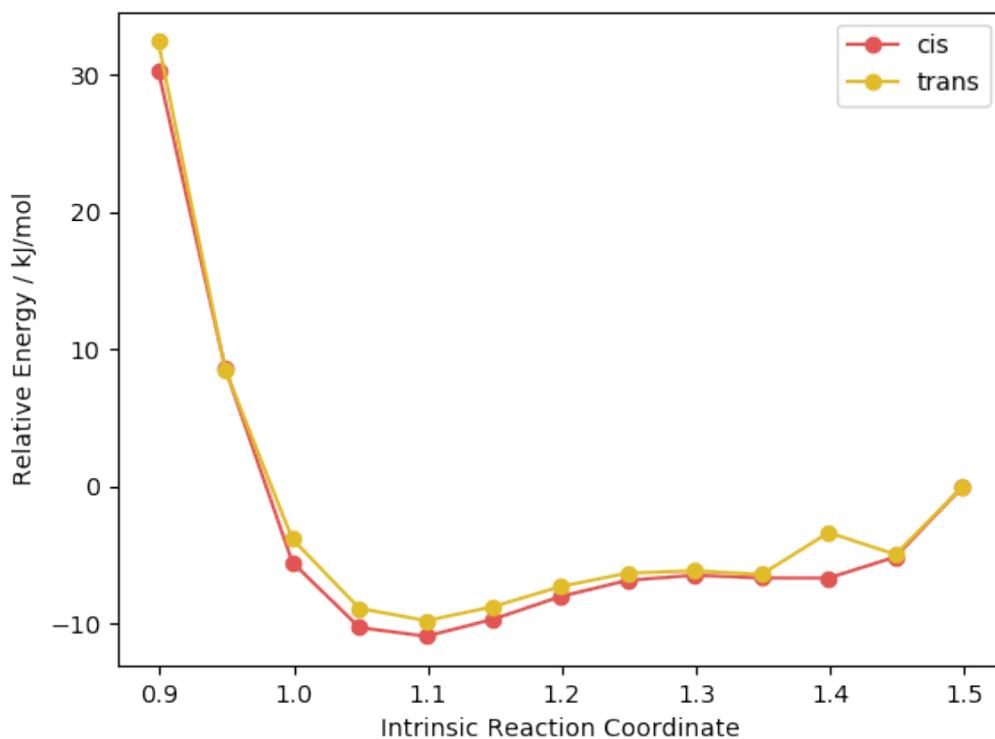
Figure 7.5: Upper: the potential energy curve for the protonation of *cis*- (red) and *trans*-glycolaldhyde (yellow) by H_3O^+. Lower: the geometries of the low energy pre-reaction complexes formed during the protonation process.
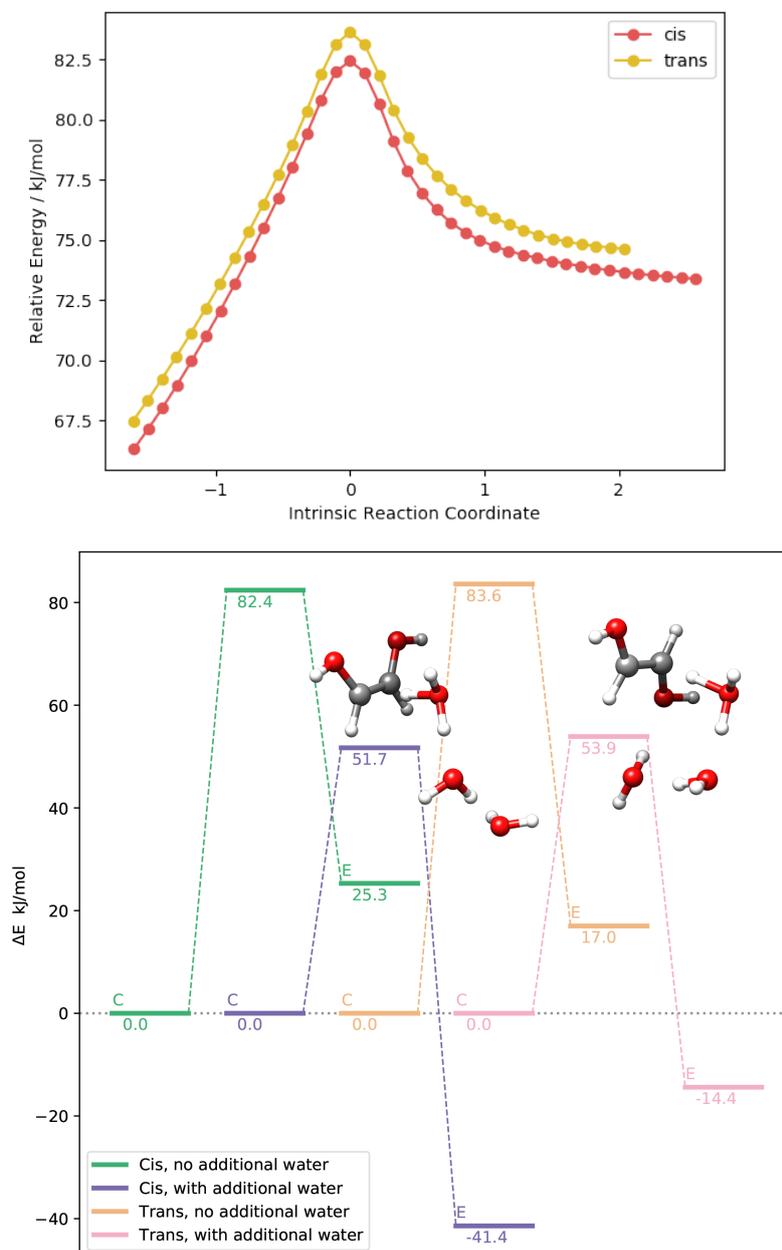
Figure 7.6: Upper: the intrinsic reaction coordinates for the enolisation of glyco-laldehyde (step C) for the *cis-* (red) and *trans-* (yellow) isomers of glycolaldehyde. Lower: the reaction profiles for step C with and without additional, stabilising water species to investigate the effect of water ice on the barrier height with the optimised geometries for the transition states shown.

and protonated formaldehyde to produce protonated glyceraldehyde. This step was successfully modelled using *trans*-glycolaldehyde. However, the transition state for the *cis-* isomer could not be located despite various attempts to do so. The transition state given for the *cis-* isomer containing reaction actually corresponds to the migration of the formaldehyde from one end of the glycolaldehyde to the other (shown diagrammatically in figure 7.7) and not simply the addition of the two species. The *cis-* isomer of protonated glyceraldehyde (left panel in figure 7.8) is lower in energy than the *trans-* isomer by 7 kJ mol^{-1}. However, the most stable structure of protonated glyceraldehyde is that which adopts the staggered kind of conformation expected for linear organic molecules (central structure in figure 7.8). As can be seen, this lowest energy structure more closely resembles the product gained from using *trans*-glycolaldehyde (right panel in figure 7.8). It may, therefore, be that the conformational change needed to move from the *cis*-isomer to the lowest energy geometry for glyceraldehyde makes this route less favourable than using *trans*-glycolaldehyde. Thus, as the correct transition state for the *cis-* isomer reaction in step E could not be found and as the target product bears greater resemblance to the *trans-* reactant, this is the geometry that has been carried forward into the subsequent reaction steps.
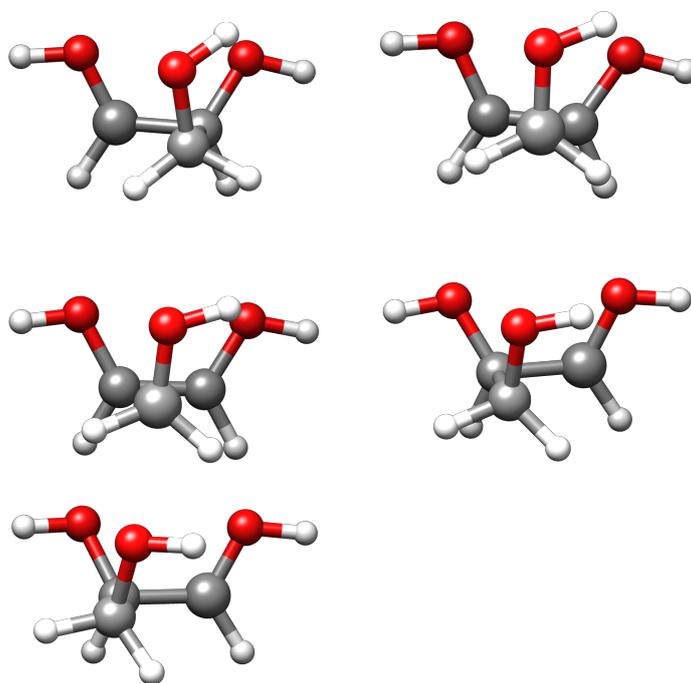
Figure 7.7: The migration of formaldehyde between carbon atoms on protonated glycolaldehyde during the attempt to find the transition state for the formation of protonated glyceraldehyde as part of the formose reaction.

The addition step between *trans*-glycolaldehyde and protonated formaldehyde (step E) is exothermic and has a barrier of 7.9 kJ mol^{-1}, see top left in figure 7.9.
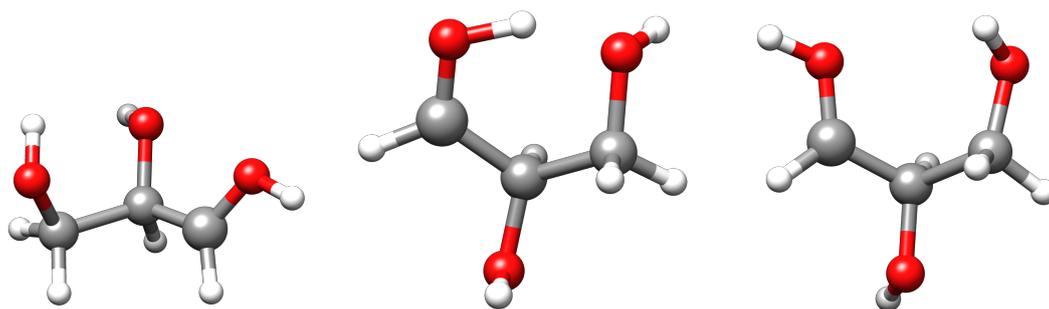
Figure 7.8: The structures of gylceraldehyde obtained using left, a constrained *cis-* glycoladehyde, middle, no constraints and right, a constrained *trans-*glycolaldehyde.

Due to the small size of the barrier and the 30 kJ mol^{-1} energy gain upon the formation of protonated glyceraldehyde there is a chance that the reactants could overcome the barrier, and, having done so would remain as protonated glyceraldehyde. The subsequent enolisation of the protonated glyceraldehyde, step F, is also exothermic and possesses a similarly small barrier of 6.9 kJ mol^{-1}, see top right in figure 7.9. The gain in stability for enol-glyceraldehyde is a product of the delocalisation possible between the double bond and the two oxygen atoms with lone pairs. Additionally, as with step C, both of these steps could be stabilised by the addition of water ice to the system. However, as the barriers are already potentially surmountable without extra ice molecules, exploratory calculations have not been performed in this case.

7.3.3   Steps G to ribose formation

The reaction of this enol with protonated glycolaldehyde to give pentose is then step G (lower in figure 7.9). This reaction is highly exothermic with pentose lying 136 kJ mol^{-1} lower in energy than enolised glyceraldehyde, see figure 7.3. There is, however, a barrier associated with this step. The transition state is 29 kJ mol^{-1} higher in energy than the reactants, which would be unlikely to be crossed at ISM temperatures without an additional energy input. The transition state corresponds to the formation of a tetrahedral intermediate between the enolised glyceraldehyde and the protonated glycolaldehyde which causes a change in the geometry of the glyceraldehyde reactant. The size of the barrier corresponds to 0.3 eV which is far below the amount of energy supplied by UV photons and so they could provide the energy required to drive this reaction step forward.
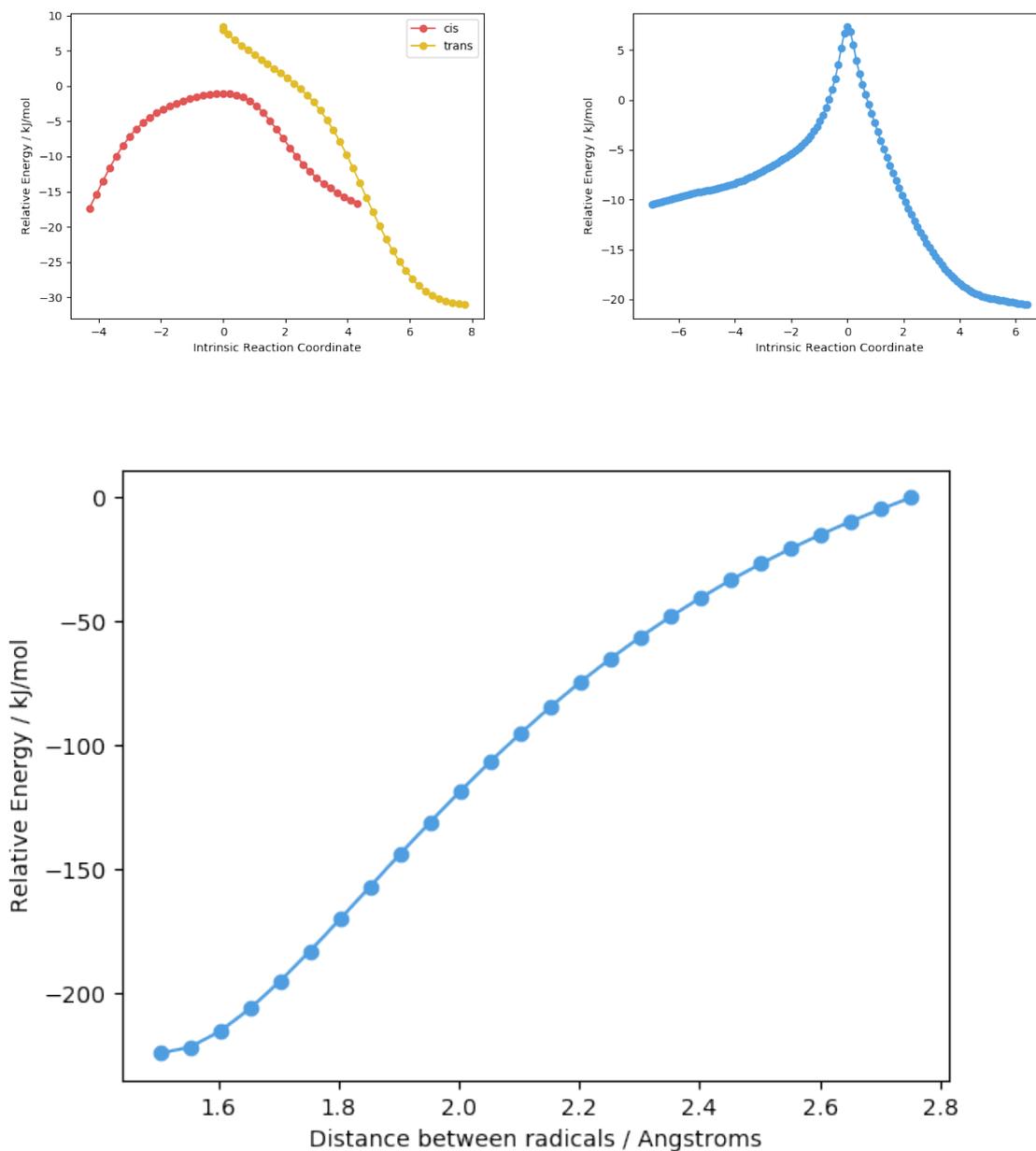
Figure 7.9: Upper: the intrinsic reaction coordinates for left, the reaction between protonated glycolaldehyde and protonated formaldehyde to form glyceraldehyde and right, the enolisation of protonated glyceraldehyde. Lower: the potential energy surface from the transition state to the products for the aldol reaction between protonated glycolaldehyde and the enol of glyceraldehyde to give pentose.

The final steps, H to J, are the rearrangement of the pentose generated in step G to the intended product ribose. This involves an enolisation step, step H, followed by the addition of a hydrogen to the double bond and the reformation of a carbonyl group, step I, and finally the deprotonation to yield ribose, step J, see figures 7.10 and 7.11. All steps are exothermic but the last step is the only barrierless reaction of the three, see figure 7.11. The barriers to reactions H and I are the largest in the reaction so far at 75 and 47 kJ mol^{-1} respectively. The barrier in step H corresponds to the terminal carbon adopting an sp^2 geometry so as to be able to form the carbon-carbon double bond in the enol (see figure 7.10). Conversely, the barrier in step I is the point at which the incoming hydrogen causes a tetrahedral geometry to be adopted, the structure is given in figure 7.10. The barrier heights are the equivalent of 0.77 and 0.49 eV and so could be overcome by the input of energy from a UV photon.

Given the effect of water on the barrier height in step C, additional water was added to step H, as it possesses the largest barrier in the reaction. As with the previous case the additional water allowed for the sharing of the proton between them, however; the barrier height was not lowered compared to the gas-phase case, it actually increases slightly to 80 kJ mol^{-1}. As the barrier corresponds to the change in geometry of the carbon atom this is not unexpected. The energy of the product is lowered significantly to the extent the reaction becomes exothermic by 14 kJ mol^{-1}. This would increase the barrier to step I. As the barrier to step I also relates to a change in geometry of a carbon atom it is thought that it will behave like the barrier in step J and water will be unlikely to reduce it significantly. Thus, assuming the barrier height remains roughly consistent it would increase from 47 to 79 kJ mol^{-1}. However, both of these barriers still correspond to less than 1 eV in energy meaning that UV photons would be able to provide the energy to overcome them. Should the presence of UV allow for the formation of protonated ribose then due to the barrierless nature of the deprotonation step then the formation of neutral ribose is highly likely.

7.4 Conclusions

The work in this chapter serves best as a proof of concept for the viability of the formose reaction in the ISM. It was found that the overall reaction is exothermic and while barriers are present they are all submerged below the energy of the initial reactants. The initial radical-radical recombination gives out 413 kJ/mol of energy and could provide the energy required to overcome barriers early in the reaction, before equilibration has occurred. Later barriers, such as those encountered during the enolisation of glycolaldehyde and the conversion of protonated
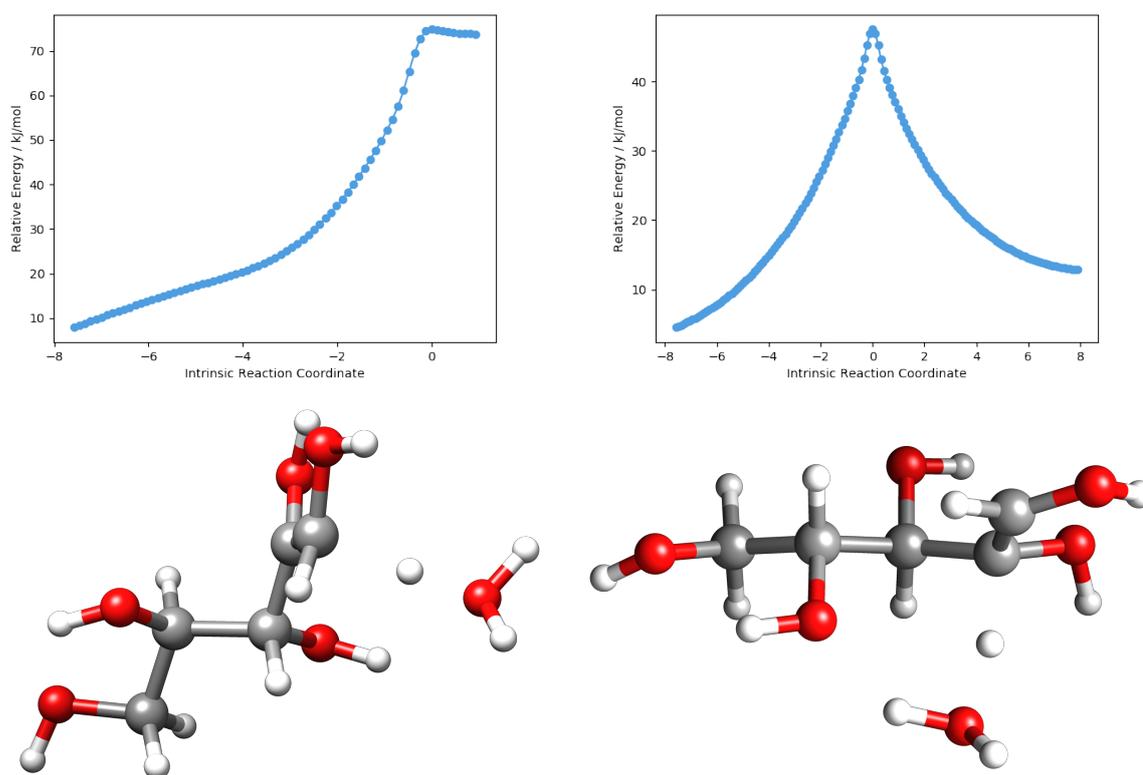
Figure 7.10: Upper: the intrinsic reaction coordinates (IRC) for (left) the enolisation of pentose and (right) the addition of a proton to the pentose enol. Lower: the transition states for left, the enolisation of pentose and right, the addition of a proton to the enol to give protonated ribose.
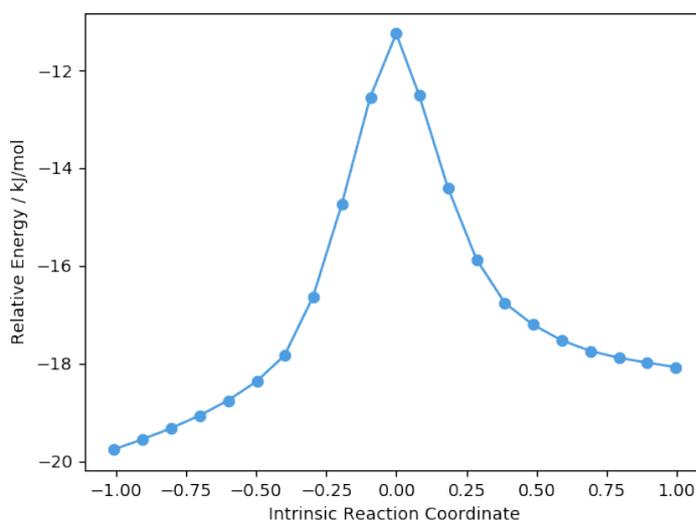
Figure 7.11: The intrinsic reaction coordinate (IRC) for the deprotonation step to produce ribose.

pentose to ribose are too high to be crossed at ISM temperatures with no external energy input. However, these barriers could be overcome if UV photons supply the energy to surmount it. This is consistent with experimental evidence that irradiating laboratory ices with UV can yield sugars.[224] The impact of these barriers could further be reduced by the presence of an ice mantle. The barrier in step C, the enolisation of glycolaldehyde, was lowered by almost 31 kJ/mol when additional water was included in the system. The water stabilised the charged species by means of hydrogen bonding; it also stabilised the product sufficiently that the reaction became exothermic. Therefore, the formation of ribose on the surface of icy dust grains should be possible, particularly in areas of the ISM where UV radiation is present and has penetrated into dust clouds.

What has not been addressed in this chapter are factors such as competing reactions to form other sugars as the products of the formose reaction are simply statistical mixtures of all possible sugars. Chirality has also not been enforced with any particular rigour. It is well known that in biology it is the D stereoisomers that are present and so future work on this subject needs to investigate how a racemic mixture is avoided. Nonetheless, the underlying basis of using the formose reaction to make sugars abiotically is viable and should be given further consideration.

8.   Conclusions and future work

It was the aim in this body of work to investigate the formation of a variety of astrochemical molecules in the ISM. This selection of molecules spans from those with a key biological role, such as ribose and urea, to those like the PAH species that make up the dust grains on which other molecules form. In each case computational methods have been applied to probe the system and processes of interest.

Chapters 3 and 4 do not directly look at the formation routes of COMs in the ISM but seek to computationally elucidate experimental findings. After using experimental and calculated gas-phase spectra to benchmark the solid-state calculation generated spectra, in Chapter 3 is was found that the most likely composition of an irradiated benzene and carbon dioxide ice is a 1:1 mixture of benzoic and isophthalic acid. Single acid systems did not replicate the experimental residue spectra well and addition of the *para* substituted acid, terephthalic acid, did not improve the shape of the simulated spectrum compared to the experimentally obtained one. The apparent preference of formation of isophthalic acid over terephthalic suggests that the reaction mechanism may have an ionic component rather than being purely radical based. This is something that would need further investigation, although at present the form this would take is unkown.

Chapter 4 confirmed that the lowering of PAH ionisation energies in water ices is facilitated by the ability of the water ice to solvate an electron from the PAH. The decrease did have some dependence on the amount of water surrounding the PAH. However, the degree of change in the excitation energy decrease was found to be on the order of 0.1 eV and so the structural dependence is not great. ONIOM calculations were attempted to try and increase the amount of water in the system but these yielded poor UV-Vis spectra and so, the approach to perform structural optimisations using MM and then apply TD-DFT for generating the spectra was employed. Electron density calculations showed electron density on the water ice in various excited states of various PAHs. The calculations also agreed with the previously reported decrease in the extent of the ionisation energy lowering as the size of the PAHs increase, although the predictive model does not hold up well for PAHs that do not have a circular shape.

In Chapter 5 the formation of the isomers methyl formate and glycolaldehyde from radical species was followed. Initial studies into the dissociation of methanol to yield one of the reactant radical species gave gas-phase dissociation energies of 5.6 and 4.5 eV for the CH and OH respectively. This was then found to decrease significantly in the case of the OH bond when water ice was included in the system. The same effect was not observed for the CH bond. Structural examinations suggest that the difference between the dissociation of the two bonds lies in the ability of the water ice to take up the departing hydrogen radical with

148

this occurring only for the OH bond. The outcome of this difference would lead to a greater quantity of OCH_3 being produced which reacts barrierlessly with the HCO radical to produce methyl formate which is the more abundant isomer of the two. More extensive and mixed ice systems that better replicate those found on interstellar dust grains would allow for this formation route to tested under more general circumstances.

Chapter 6 employed a similar approach to Chapter 5 to test a suggested interstellar formation route for urea. Neutral-neutral, radical-neutral and radical-radical as well as ionic pathways were investigated. The neutral-neutral reaction between ammonia and formamide was found to be endothermic making it unfeasible under astrochemical conditions. Both radical-neutral routes ($NH_2 \cdot + H_2NCHO$ and $NH_3 + H_2NCO \cdot$) gave barriers to urea formation that would be insurmountable in the ISM. $NH_2 \cdot + H_2NCHO$ had a barrier of 150 kJ/mol while $NH_3 + H_2NCO \cdot$ had a barrier of 291 kJ/mol. Formation routes involving charged species based around isocyanic acid did provide a barrierless reaction to yield protonated urea but the deprotonation step required a significant energy input. However, addition of water to the system for the deprotonation step lowered the barrier to the extent that the process became effectively barrierless, making this route, with water present, the most feasible. The second most promising reaction was that between the amino radical and the formamide radical ($NH_2 \cdot + H_2NCO \cdot$) which only possessed a barrier of 4 kJ/mol, a barrier which could be overcome in an astrochemical setting. As with the methyl formate work, the inclusion of an ice system would expand the scope of this work to better mimic dust grain surfaces and provide greater insight into the processes taking place.

In was noted in both Chapter 5 and 6 that the use of DFT methods to model radical containing processes gave results that were not representative of the system. This was attributed to the poor behaviour of single determinant methods at long range causing the system to not actually be treated as radicals. Multi-configurational methods performed much better for these systems, however, these vastly limited the scale of the system able to be calculated and the choice of the correct active space became a significant undertaking.

Chapter 7 followed the formation of ribose in the ISM using the formose reaction. Starting from the confirmed ISM species methanol and formaldehyde, successive aldol reactions were used to build larger sugars until the specific tetrose ribose is formed. The initial radical-radical recombination between HCO and $HOCH_2$ was found to be highly exothermic (in excess of 400 kJ/mol) which could provide the energy to cross any subsequent barriers before the system equilibrates. There are barriers along the reaction pathway, all corresponding to the movement of atoms and often due to the formation of tetrahedral intermediates.

These barriers are too substantial to be overcome at ISM temperatures without the input of additional energy either from the first, exothermic step or from UV radiation. Later steps, like the rearrangement of tetrose to give ribose, are too far removed from the initial step for that to be the energy input source and so UV is necessary. It was found that additional water ice in the system, with even small quantities of water making a significant difference, can lower the barrier heights by distributing a proton between the ice molecules. Thus, as the reaction occurs on icy dust grains, the affect of the ice would make the reaction more plausible under ISM conditions.

Overall radical pathways to COM formation present themselves as the most facile way to generate astrochemical species in the ISM. Equally surface reactions in which the ice mantle species can facilitate reactions have shown themselves to open up many more reactions than would be possible using gas-phase chemistry alone. However the computationally intensive nature of multiconfigurational calculations needed to properly describe the radical systems does not allow for work on extensive ice systems. Reconciling these two needs would allow for much more accurate treatments of a vast array of astrochemical reactions. For instance, to date the exact formation route of PAHs in the ISM is not fully understood despite PAHs being widely accepted as a constituent of the carbonaceous dust grains that make up interstellar clouds. One way to achieve this aim would be to make use of combined or embedded methods, like ONIOM or QM:MM. If a method to embed a multiconfigurational calculation within a much larger ice system treated with molecular mechanics was developed then ISM radical chemistry could be accurately modelled in more representative ices. If, ultimately this could lead to the construction of a full-scale, accurate computational model of an ice covered dust grain then the scope of the work carried out in this field would be vastly expanded. This would prove not to be a small task though, and require collaboration between experimental and computational astrochemists as well as method developers to find ways to marry the various necessary techniques together.

The logical expansion of this work would be to use the methods established here to investigate the possibility of forming nucleotides from the urea formed in Chapter 6, see figure 8.1. Both cytosine and uracil are needed to form RNA and so, if they are able to be formed in an astrochemical setting, then it would have significant impacts for the origins of life. Another constituent of RNA is cyclic deoxy-ribose. As Chapter 7 outlines the formation of linear ribose from the formose reaction this could be used as a basis to explore the cyclisation of ribose under ISM conditions. It would also be prudent to explore the alternative reaction products of the formose reaction as it can lead to many sugar products. By follow-
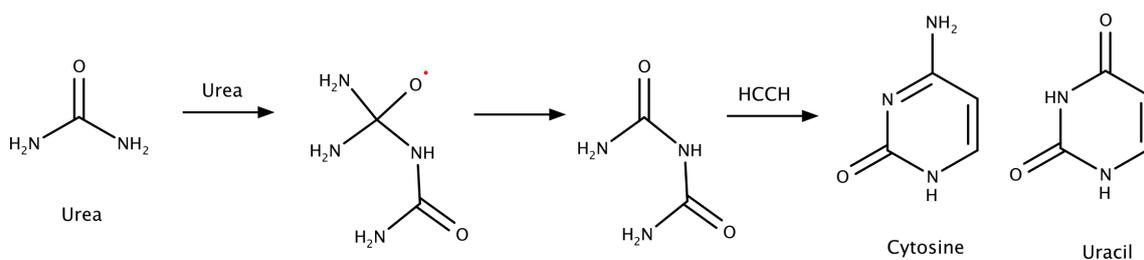
Figure 8.1: A potential mechanism for the formation of nuclear bases from urea

ing the other reaction routes it would provide insight into how much ribose could reasonably be expected to form from this mechanism. Lastly the reaction between deoxy-ribose and the nuclear bases should be explored to ascertain whether the construction of RNA could have taken place on dust grains or comet surfaces. It is hoped that this work would shed a great deal of light into how possible an RNA world origin of life would have been.

Appendices

A. Example inputs

A.1   Molpro input for MCSCF scan calculations

```
***,MeOHCHDissociation
memory,600,m
gprint,orbital
angstrom
geometry={
o1
c2, o1, rOC
h3, o1, rOH, c2, aCOH
h4, c2, r, o1, aOCH1, h3, dHOCH1
h5, c2, rCH, o1, aOCH2, h3, dHCOH2
h6, c2, rCH, o1, aOCH2, h3, dHCOH3
}
rOC=1.42 Ang
rOH=0.96 Ang
r=0.5 Ang
rCH=1.09 Ang
aCOH=107.8 Degree
aOCH1=106.8 Degree
aOCH2=112.5 Degree
dHOCH1=179.9 Degree
dHCOH2=-61.64 Degree
dHCOH3=61.46 Degree
i=0
basis=vdz
do i=1,51
rCH(i)=r
{hf;occ,9;wf,18,1,0}; e_hf(i)=energy
optg;inactive,r
{multi;closed,2;occ,14;wf,18,1,0;canonical,ci}; e_mcscf(i)=energy
r=r+0.05
enddo
{table,rCH,e_hf,e_mcscf
head,r24,EHF,EMCSCF
save,MeOHCHPT2.tab}
```

A.2   VASP inputs for geometry optimisation

A.2.1   INCAR

SYSTE=CHA Si/O Zeolite
ALGO=FAST
ISMEAR=1
SIGMA=0.2
LCHARG=.FALSE.
ISYM=2
LREAL=Auto
ISPIN=2
NSW = 40
IBRION = 2

A.2.2   KPOINTS

Automatic mesh
0 ! number of k-points = 0 -¿automatic generation scheme
Monkhorst-Pack ! select Monkhorst-Pack (first letter is significant)
1 1 1 ! size of mesh (4x4x4 points along b1, b2, b3)
0. 0. 0. ! shift of the k-mesh

A.3   VASP input for IR calculations

A.3.1   INCAR

SYSTE=CHA Si/O Zeolite
ALGO=FAST
ISMEAR=1
SIGMA=0.2
LCHARG=.FALSE.
ISYM=2
LREAL=Auto
ISPIN=2
NSW = 1
IBRION = 7
LEPSILON = .TRUE.
NWRITE = 3

A.4 VASP inputs for TDDFT calculations

A.4.1 Step 1 INCAR

```
SYSTE=CHA Si/O Zeolite
NBANDS=184
ISMEAR=0
SIGMA=0.05
ALGO = D
LHFCALC=.TRUE.
AEXX=0.3
HFSCREEN=0.2
LOPTICS = .TRUE.
LCHARG=.FALSE.
ISYM=3
PRECFOCK=M
LREAL=.FALSE.
ISPIN=2
NSW = 40
IBRION = 2
IVDW=12
```

A.4.2 Step 2 INCAR

```
SYSTE=CHA Si/O Zeolite
NBANDS=184
ISMEAR=0
SIGMA=0.05
ALGO = TDHF
LHFCALC=.TRUE.
AEXX=0.3
HFSCREEN=0.2
LCHARG=.FALSE.
PRECFOCK= N
ISYM=3
LREAL=.FALSE.
ISPIN=2
NSW = 40
IBRION = 2
```

IVDW=12

References

(1) A. M. Shaw, *Astrochemistry: From Astronomy to Astrobiology*, John Wiley & Sons, 2007, p. 352.

(2) A. G. G. M. Tielens, *The Physics and Chemistry of the Interstellar Medium*, Cambridge University Press, 2005.

(3) E. F. van Dishoeck, "Astrochemistry of dust, ice and gas: introduction and overview", *Faraday Discussions*, 2014, **168**, 9.

(4) A. G. G. M. Tielens, "The molecular universe", *Reviews of Modern Physics*, 2013, **85**, 1021–1081.

(5) M.-a. Hashimoto, R. Nakamura, E. P. B. A. Thushari and K. Arai, *Big-bang nucleosynthesis : thermonuclear history in the early universe*, Springer, 2018, p. 84.

(6) A. G. W. A. G. W. Cameron, D. M. Kahl and J. José, *Stellar evolution, nuclear astrophysics, and nucleogenesis*, Courier Corporation, 2013, p. 206.

(7) F. Hoyle, "On Nuclear Reactions Occurring in Very Hot STARS.I. the Synthesis of Elements from Carbon to Nickel.", *The Astrophysical Journal Supplement Series*, 1954, **1**, 121.

(8) E. M. Burbidge, G. R. Burbidge, W. A. Fowler and F. Hoyle, "Synthesis of the Elements in Stars", *Reviews of Modern Physics*, 1957, **29**, 547–650.

(9) M. R. Krumholz, E. Telles, R. Dupke and D. Lazzaro, XV Special Courses at the National Observatory of Rio De Janeiro, AIP Publishing, 2011, vol. 1386, pp. 9–57.

(10) D. A. O. Bradley W. Carroll, *Introduction to Modern Astrophysics, An: Pearson New International Edition*, Pearson Education.

(11) J. H. Jeans, "The Stability of a Spherical Nebula", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 1902, **199**, 1–53.

(12) D. Ward-Thompson and A. P. Whitworth, *An Introduction to Star Formation*, Cambridge University Press, 2011.

(13) A Weiss, "Stellar nucleosynthesis", *Physica Scripta*, 2008, **T133**, 014025.

(14) H. O. U. Fynbo, C. A. Diget, U. C. Bergmann, M. J. G. Borge, J. Cederkäll, P. Dendooven, L. M. Fraile, S. Franchoo, V. N. Fedosseev, B. R. Fulton, W. Huang, J. Huikari, H. B. Jeppesen, A. S. Jokinen, P. Jones, B. Jonson, U. Köster, K. Langanke, M. Meister, T. Nilsson, G. Nyman, Y. Prezado, K. Riisager, S. Rinta-Antila, O. Tengblad, M. Turrion, Y. Wang, L. Weissman, K. Wilhelmsen, J. Äystö and T. I. Collaboration, "Revised rates for the stellar triple-α process from measurement of 12C nuclear resonances", *Nature*, 2005, **433**, 136–139.

(15) W. M. Irvine, in, ed. M. Gargaud, R. Amils, J. C. Quintanilla, H. J. J. Cleaves, W. M. Irvine, D. L. Pinti and M. Viso, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, ch. Diffuse Clouds, pp. 431–431.

(16) H. S. P. Müller, A. Belloche, L.-H. Xu, R. M. Lees, R. T. Garrod, A. Walters, J. van Wijngaarden, F. Lewen, S. Schlemmer and K. M. Menten, "Exploring Molecular Complexity with ALMA (EMoCA): Alkanethiols and alkanols in Sagittarius B2(N2)", *Astronomy & Astrophysics*, 2015, 14.

(17) P. Ehrenfreund and S. B. Charnley, "Organic Molecules in the Interstellar Medium, Comets, and Meteorites: A Voyage from Dark Clouds to the Early Earth", en, *Annual Review of Astronomy and Astrophysics*, 2000, **38**, 427–483.

(18) N. L. J. Cox, P. Pilleri, O. Berné, J. Cernicharo and C. Joblin, "Polycyclic aromatic hydrocarbons and molecular hydrogen in oxygen-rich planetary nebulae: the case of NGC6720", *Monthly Notices of the Royal Astronomical Society: Letters*, 2015, **456**, L89–L93.

(19) T. P. M. Goumans, C. R. A. Catlow, W. A. Brown, J. Kästner and P. Sherwood, "An embedded cluster study of the formation of water on interstellar dust grains.", en, *Physical Chemistry Chemical Physics : PCCP*, 2009, **11**, 5431–6.

(20) T. Millar and D. Williams, *Dust and Chemistry in Astronomy*, CRC Press, 1993, p. 336.

(21) A. P. Jones, N. Ysard, M. Köhler, L. Fanciullo, M. Bocchio, E. Micelotta, L. Verstraete and V. Guillet, "The cycling of carbon into and out of dust", en, *Faraday Discussions*, 2014, **168**, 313.

(22) A. L. Skov, J. D. Thrower and L. Hornekær, "Polycyclic aromatic hydrocarbons catalysts for molecular hydrogen formation", en, *Faraday Discussions*, 2014, **168**, 223.

(23) T. Lamberts, X. de Vries and H. M. Cuppen, "The formation of ice mantles on interstellar grains revisited - the effect of exothermicity", *Faraday Discussions*, 2014, **168**, 327.

(24) D. Williams and S. Viti, "Modelling interstellar physics and chemistry: implications for surface and solid-state processes", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2013, **371**, 20110587.

(25) J. L. Edridge, K. Freimann, D. J. Burke and W. A. Brown, "Surface science investigations of the role of CO_2 in astrophysical ices", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2013, **371**, 20110578.

(26) H. Y. McSween and G. R. Huss, *Cosmochemistry*, Cambridge University Press, 2010.

(27) K. M. Pontoppidan and S. M. Blevins, "The chemistry of planet-forming regions is not interstellar", en, *Faraday Discussions*, 2014, **168**, 49.

(28) P. J. Armitage, in *Handbook of Exoplanets*, Springer International Publishing, Cham, 2018, pp. 2185–2203.

(29) L. A. Hayden and E. B. Watson, "A diffusion mechanism for coremantle interaction", *Nature*, 2007, **450**, 709–711.

(30) S. Marchi, W. F. Bottke, L. T. Elkins-Tanton, M. Bierhaus, K. Wuennemann, A. Morbidelli and D. A. Kring, "Widespread mixing and burial of Earth's Hadean crust by asteroid impacts", *Nature*, 2014, **511**, 578–582.

(31) J. Hartmann, "Investigations on the spectrum and orbit of delta Orionis.", *The Astrophysical Journal*, 1904, **19**, 268.

(32) J. K. Jø rgensen, C. Favre, S. E. Bisschop, T. L. Bourke, E. F. van Dishoeck and M. Schmalzl, "Detection of the Simplest Sugar Glycolaldehyde, in a Solar-type Protostar with ALMA", en, *The Astrophysical Journal*, 2012, **757**, L4.

(33) A. K. Speck, M. Meixner, D. Fong, P. R. McCullough, D. E. Moser and T. Ueta, "Large-Scale Extended Emission around the Helix Nebula: Dust, Molecules, Atoms, and Ions", *The Astronomical Journal*, 2002, **123**, 346–361.

(34) A. Belloche, H. S. P. Müller, K. M. Menten, P. Schilke and C. Comito, "Complex organic molecules in the interstellar medium: IRAM 30 m line survey of Sagittarius B2(N) and (M)", en, *Astronomy & Astrophysics*, 2013, **559**, A47.

(35) M. J. Abplanalp, M. Förstel and R. I. Kaiser, "Exploiting single photon vacuum ultraviolet photoionization to unravel the synthesis of complex organic molecules in interstellar ices", *Chemical Physics Letters*, 2016, **644**, 79–98.

(36) E. Herbst and E. F. van Dishoeck, "Complex Organic Interstellar Molecules", *Annual Review of Astronomy and Astrophysics*, 2009, **47**, 427–480.

(37) H. Kamminga, "Historical perspective: The problem of the origin of life in the context of developments in biology", *Origins of Life and Evolution of the Biosphere*, 1988, **18**, 1–11.

(38) J. W. Schopf, *Earth's earliest biosphere : its origin and evolution*, Princeton University Press, 1983, p. 543.

(39) I. Gilmour and M. A. Sephton, *An introduction to astrobiology*, Open University, 2004, p. 358.

(40) C. F. Chyba and G. D. Mcdonald, *Annu. Rev. Earth Planet. Sci*, 215–264.

(41) L. Remusat, S. Derenne and F. Robert, "New insight on aliphatic linkages in the macromolecular organic fraction of Orgueil and Murchison meteorites through ruthenium tetroxide oxidation", *Geochimica et Cosmochimica Acta*, 2005, **69**, 4377–4386.

(42) N. Fray, A. Bardyn, H. Cottin, K. Altwegg, D. Baklouti, C. Briois, L. Colangeli, C. Engrand, H. Fischer, A. Glasmachers, E. Grün, G. Haerendel, H. Henkel, H. Höfner, K. Hornung, E. K. Jessberger, A. Koch, H. Krüger, Y. Langevin, H. Lehto, K. Lehto, L. Le Roy, S. Merouane, P. Modica, F.-R. Orthous-Daunay, J. Paquette, F. Raulin, J. Rynö, R. Schulz, J. Silén, S. Siljeström, W. Steiger, O. Stenzel, T. Stephan, L. Thirkell, R. Thomas, K. Torkar, K. Varmuza, K.-P. Wanczek, B. Zaprudin, J. Kissel and M. Hilchenbach, "High-molecular-weight organic matter in the particles of comet 67P/Churyumov-Gerasimenko", *Nature*, 2016, **538**, 72–74.

(43) A. Bardyn, D. Baklouti, H. Cottin, N. Fray, C. Briois, J. Paquette, O. Stenzel, C. Engrand, H. Fischer, K. Hornung, R. Isnard, Y. Langevin, H. Lehto, L. Le Roy, N. Ligier, S. Merouane, P. Modica, F.-R. Orthous-Daunay, J. Rynö, R. Schulz, J. Silén, L. Thirkell, K. Varmuza, B. Zaprudin, J. Kissel and M. Hilchenbach, "Carbon-rich dust in comet 67P/Churyumov-Gerasimenko measured by COSIMA/Rosetta", *Monthly Notices of the Royal Astronomical Society*, 2017, **469**, S712–S722.

(44) J. Hartmann, "Investigations on the spectrum and orbit of delta Orionis.", *The Astrophysical Journal*, 1904, **19**, 268.

(45) G Wächtershäuser, "Evolution of the first metabolic cycles.", *Proceedings of the National Academy of Sciences of the United States of America*, 1990, **87**, 200–4.

(46) A. G. Cairns-Smith, A. J. Hall and M. J. Russell, in *Marine Hydrothermal Systems and the Origin of Life*, Springer Netherlands, Dordrecht, 1992, pp. 161–180.

(47) J. Bernal, *The physical basis of life*, Routledge and Paul, 1951.

(48) A. Rich, in *Editors Horizons in Biochemistry*, New York: Academic Press, 1962, pp. 103–126.

(49) K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling and T. R. Cech, "Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena", *Cell*, 1982, **31**, 147–157.

(50) C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace and S. Altman, "The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme", *Cell*, 1983, **35**, 849–857.

(51) L. E. Orgel, "RNA catalysis and the origins of life", *Journal of Theoretical Biology*, 1986, **123**, 127–149.

(52) V. I. Shematovich, "Formation of complex chemical species in astrochemistry (a review)", *Solar System Research*, 2012, **46**, 391–407.

(53) V. I. Shematovich, B. M. Shustov, D. S. Wiebe, Y. N. Pavlyuchenkov and Z.-Y. Li, "Self-Consistent Theoretical Models of Collapsing Pre-Stellar Cores", English, *Proceedings of the International Astronomical Union*, 2006, **1**, 37.

(54) E. Herbst and W. Klemperer, "The Formation and Depletion of Molecules in Dense Interstellar Clouds", *The Astrophysical Journal*, 1973, **185**, 505.

(55) S. Yamamoto, *Introduction to astrochemistry : chemical evolution from interstellar clouds to star and planet formation*, Springer, 2016.

(56) T. E. Cravens and A. Dalgarno, "Ionization, dissociation, and heating efficiencies of cosmic rays in a gas of molecular hydrogen", *The Astrophysical Journal*, 1978, **219**, 750.

(57) R. Gredel, S. Lepp, A. Dalgarno and E. Herbst, "Cosmic-ray-induced photodissociation and photoionization rates of interstellar molecules", *The Astrophysical Journal*, 1989, **347**, 289.

(58) J. Woodall, M. Agúndez, A. Markwick-Kemper and T. Millar, *The UMIST database for astrochemistry*, 2012.

(59) V Wakelam, J.-C Loison, K. M. Hickson and M Ruaud, "A proposed chemical scheme for HCCO formation in cold dense clouds", *MNRAS*, 2015, **453**, 48–52.

(60) L Reboussin, V Wakelam, S Guilloteau and F Hersant, "Grain-surface reactions in molecular clouds: the effect of cosmic rays and quantum tunnelling", *MNRAS*, 2014, **440**, 3557–3567.

(61) Q. Chang, H. M. Cuppen and E. Herbst, "Gas-grain chemistry in cold interstellar cloud cores withamicroscopic Monte Carlo approach to surface chemistry", en, *Astronomy and Astrophysics*, 2007, **469**, 973–983.

(62) M. A. Requena-Torres, J. Martín-Pintado, A. Rodríguez-Franco, S. Martín, N. J. Rodríguez-Fernández and P. de Vicente, "Organic molecules in the Galactic center", en, *Astronomy and Astrophysics*, 2006, **455**, 971–985.

(63) J. L. Neill, A. L. Steber, M. T. Muckle, D. P. Zaleski, V. Lattanzi, S. Spezzano, M. C. McCarthy, A. J. Remijan, D. N. Friedel, S. L. Widicus Weaver and B. H. Pate, "Spatial distributions and interstellar reaction processes.", *The journal of physical chemistry. A*, 2011, **115**, 6472–80.

(64) D. A. Adriaens, T. P. M. Goumans, C. R. A. Catlow and W. A. Brown, "Computational Study of Carbonyl Sulphide Formation on Model Interstellar Dust Grains", *The Journal of Physical Chemistry C*, 2010, **114**, 1892–1900.

(65) Z. Yang and N. Pan, "Computational studies of ion-neutral reactions of astrochemical relevance: Formation of hydrogen peroxide, acetamide, and amino acetonitrile", *International Journal of Mass Spectrometry*, 2015, **378**, 364–368.

(66) H. M. Cuppen, C. Walsh, T. Lamberts, D. Semenov, R. T. Garrod, E. M. Penteado and S. Ioppolo, "Grain Surface Models and Data for Astrochemistry", *Space Science Reviews*, 2017, **212**, 1–58.

(67) S. Morisset, F. Aguillon, M. Sizun and V. Sidis, "Wave-packet study of H_2 formation on a graphite surface through the Langmuir-Hinshelwood mechanism", *The Journal of Chemical Physics*, 2005, **122**, 194702.

(68) R. T. Garrod and E. Herbst, "Formation of methyl formate and other organic species in the warm-up phase of hot molecular cores", en, *Astronomy and Astrophysics*, 2006, **457**, 927–936.

(69) R. T. Garrod, S. L. W. Weaver and E. Herbst, "Complex Chemistry in Star-forming Regions: An Expanded Gas-Grain Warm-up Chemical Model", en, *The Astrophysical Journal*, 2008, **682**, 283–302.

(70) B. J. Irving, A. J. H. M. Meijer and D. Morgan, "On multiple adsorptions of hydrogen atoms on graphene", en, *Physica Scripta*, 2011, **84**, 028108.

(71) T. P. Snow, V Le Page, Y Keheyan and V. M. Bierbaum, "The interstellar chemistry of PAH cations.", *Nature*, 1998, **391**, 259–60.

(72) L. Hornekaer, W. Xu, R. Otero, E. Læ gsgaard and F. Besenbacher, "Long range orientation of meta-stable atomic hydrogen adsorbate clusters on the graphite(0001) surface", *Chemical Physics Letters*, 2007, **446**, 237–242.

(73) A. G. G. M. Tielens and A. Hagen, "Model calculations of the molecular composition of interstellar grain mantles", *Astronomy and Astrophysics*, 1982, **114**, 245–260.

(74) R. Papoular, "On water ice formation in interstellar clouds", *Monthly Notices of the Royal Astronomical Society*, 2005, **362**, 489–497.

(75) H. M. Cuppen and E. Herbst, "Simulation of the Formation and Morphology of Ice Mantles on Interstellar Grains", en, *The Astrophysical Journal*, 2007, **668**, 294–309.

(76) S. Ioppolo, H. M. Cuppen, C. Romanzin, E. F. van Dishoeck and H. Linnartz, "Laboratory Evidence for Efficient Water Formation in Interstellar Ices", en, *The Astrophysical Journal*, 2008, **686**, 1474–1479.

(77) N. Miyauchi, H. Hidaka, T. Chigai, A. Nagaoka, N. Watanabe and A. Kouchi, "Formation of hydrogen peroxide and water from the reaction of cold hydrogen atoms with solid oxygen at 10K", *Chemical Physics Letters*, 2008, **456**, 27–30.

(78) J. Holdship, I. Jimenez-Serra, S. Viti, C. Codella, M. Benedettini, F. Fontani, M. Tafalla, R. Bachiller, C. Ceccarelli and L. Podio, "Sulfur Chemistry in L1157-B1", *The Astrophysical Journal*, 2019, **878**, 64.

(79) M. E. Palumbo, T. R. Geballe and A. G. G. M. Tielens, "Solid Carbonyl Sulfide (OCS) in Dense Molecular Clouds", en, *The Astrophysical Journal*, 1997, **479**, 839–844.

(80) M. A. Leitch-Devlin and D. A. Williams, "The mobility of species adsorbed on the surfaces of interstellar grains: consequences for the formation of interstellar molecules", *Monthly Notices of the Royal Astronomical Society*, 1984, **210**, 577–587.

(81) C. J. Bennett and R. I. Kaiser, "On the Formation of Glycolaldehyde ($HCOCH_2OH$) and Methyl Formate ($HCOOCH_3$) in Interstellar Ice Analogs", en, *The Astrophysical Journal*, 2007, **661**, 899–909.

(82) E. L. Gibb, D. C. B. Whittet, A. C. A. Boogert and A. G. G. M. Tielens, "Interstellar Ice: The Infrared Space Observatory Legacy", en, *The Astrophysical Journal Supplement Series*, 2004, **151**, 35–73.

(83) C. J. Bennett and R. I. Kaiser, "The Formation of Acetic Acid (CH_3COOH) in Interstellar Ice Analogs", *The Astrophysical Journal*, 2007, **660**, 1289–1295.

(84) M. P. Collings, M. A. Anderson, R. Chen, J. W. Dever, S. Viti, D. A. Williams and M. R. S. McCoustra, "A laboratory survey of the thermal desorption of astrophysically relevant molecules", *Monthly Notices of the Royal Astronomical Society*, 2004, **354**, 1133–1140.

(85) D. J. Burke, F. Puletti, W. A. Brown, P. M. Woods, S. Viti and B. Slater, "Glycolaldehyde, methyl formate and acetic acid adsorption and thermal desorption from interstellar ices", *Monthly Notices of the Royal Astronomical Society*, 2015, **447**, 1444–1451.

(86) D. J. Burke, F. Puletti, P. M. Woods, S. Viti, B. Slater and W. A. Brown, "Trapping and desorption of complex organic molecules in water at 20 K.", *The Journal of Chemical Physics*, 2015, **143**, 164704.

(87) S.-. Liu, D. M. Mehringer and L. E. Snyder, "Observations of Formic Acid in Hot Molecular Cores", *The Astrophysical Journal*, 2001, **552**, 654–663.

(88) M. Ikeda, M. Ohishi, A. Nummelin, J. E. Dickens, P. Bergman, A. Hjalmarson and W. M. Irvine, "Survey Observations of C_2H_4O and CH_3CHO toward Massive Star-forming Regions", *The Astrophysical Journal*, 2001, **560**, 792–805.

(89) S.-. Liu, D. M. Mehringer and L. E. Snyder, "Observations of Formic Acid in Hot Molecular Cores", *The Astrophysical Journal*, 2002, DOI: 10.1086/320563.

(90) S. Bottinelli, C. Ceccarelli, B. Lefloch, J. P. Williams, A. Castets, E. Caux, S. Cazaux, S. Maret, B. Parise and A. G. G. M. Tielens, "Complex Molecules in the Hot Core of the Low-Mass Protostar NGC 1333 IRAS 4A", en, *The Astrophysical Journal*, 2004, **615**, 354–358.

(91) D Bockelée-Morvan, D. C. Lis, J. E. Wink, D Despois, J Crovisier, R Bachiller, D. J. Benford, N Biver, P Colom, J. K. Davies, E Gérard, B Germain, M Houde, D Mehringer, R Moreno, G Paubert, T. G. Phillips and H Rauer, *New molecules found in comet C/1995 O1 (Hale-Bopp) Investigating the link between cometary and interstellar material*, tech. rep., 2000, pp. 1101–1114.

165

(92) N. Watanabe, O. Mouri, A. Nagaoka, T. Chigai, A. Kouchi and V. Pirronello, "Laboratory Simulation of Competition between Hydrogenation and Photolysis in the Chemical Evolution of H_2OCO Ice Mixtures", *The Astrophysical Journal*, 2007, **668**, 1001–1011.

(93) R. Hudson and M. Moore, "IR Spectra of Irradiated Cometary Ice Analogues Containing Methanol: A New Assignment, a Reassignment, and a Nonassignment", *Icarus*, 2000, **145**, 661–663.

(94) T. P. M. Goumans, M. A. Uppal and W. A. Brown, "Formation of CO_2 on a carbonaceous surface: a quantum chemical study", *Monthly Notices of the Royal Astronomical Society*, 2008, **384**, 1158–1164.

(95) C. J. Bennett, T. Hama, Y. S. Kim, M. Kawasaki and R. I. Kaiser, "Laboratory studies on the formation of formic acid (HCOOH) in interstellar and cometary ices", *The Astrophysical Journal*, 2011, **727**, 27.

(96) S. Ioppolo, H. M. Cuppen, E. F. van Dishoeck and H. Linnartz, "Surface formation of HCOOH at low temperature", *Monthly Notices of the Royal Astronomical Society*, 2011, **410**, 1089–1095.

(97) C. M. Leung, E. Herbst and W. F. Huebner, "Synthesis of complex molecules in dense interstellar clouds via gas-phase chemistry - A pseudo time-dependent calculation", *The Astrophysical Journal Supplement Series*, **56**, 231–256.

(98) W. M. Irvine, P. Friberg, N. Kaifu, H. E. Matthews, Y. C. Minh, M. Ohishi and S. Ishikawa, "Detection of formic acid in the cold, dark cloud L134N", *Astronomy and Astrophysics*, 1990, **229**, L9–L12.

(99) E. Vigren, M. Hamberg, V. Zhaunerchyk, M. Kaminska, J. Semaniak, M. Larsson, R. D. Thomas, M. af Ugglas, I. Kashperka, T. J. Millar, C. Walsh, H. Roberts and W. D. Geppert, "Dissociative Recombination Of Protonated Formic Acid: Implications For Molecular Cloud and Cometary Chemistry", *The Astrophysical Journal*, 2010, **709**, 1429–1434.

(100) J. Cernicharo, A. M. Heras, A. G. G. M. Tielens, J. R. Pardo, F. Herpin, M. Guélin and L. B. F. M. Waters, "Infrared Space Observatory's Discovery of C_4H_2, C_6H_2, and Benzene in CRL 618", *The Astrophysical Journal*, 2001, **546**, L123–L126.

(101) J. Clairemidi, P. Bréchignac, G. Moreels and D. Pautet, "Tentative identification of pyrene as a polycyclic aromatic molecule in UV spectra of comet P/Halley: An emission from 368 to 384 nm", *Planetary and Space Science*, 2004.

(102) J. Clairemidi, G. Moreels, O. Mousis and P. Bréchignac, "Identification of anthracene in Comet 1P/Halley", *Astronomy & Astrophysics*, 2008, **492**, 245–250.

(103) B. A. McGuire, A. M. Burkhardt, S. Kalenskii, C. N. Shingledecker, A. J. Remijan, E. Herbst and M. C. McCarthy, "Detection of the aromatic molecule benzonitrile (c- C_6H_5CN) in the interstellar medium.", *Science (New York, N.Y.)*, 2018, **359**, 202–205.

(104) L. Kolesniková, A. Daly, J. Alonso, B. Tercero and J. Cernicharo, "The millimeter wave tunneling-rotational spectrum of phenol", *Journal of Molecular Spectroscopy*, 2013, **289**, 13–20.

(105) L. Zhou, W. Zheng, R. I. Kaiser, A. Landera, A. M. Mebel, M.-C. Liang and Y. L. Yung, "Cosmic-ray-mediated Formation of Benzene on the Surface of Saturn's Moon Titan", *The Astrophysical Journal*, 2010, **718**, 1243–1251.

(106) J. H. Lacy, I. Evans, Neal J., J. M. Achtermann, D. E. Bruce, J. F. Arens and J. S. Carr, "Discovery of interstellar acetylene", *The Astrophysical Journal*, 1989, **342**, L43.

(107) E. H. Wilson, S. K. Atreya and A. Coustenis, "Mechanisms for the formation of benzene in the atmosphere of Titan", *Journal of Geophysical Research: Planets*, 2003, **108**, 5014.

(108) B. Sivaraman, R. Mukherjee, K. P. Subramanian and S. B. Banerjee, "Benzene Formation On Interstellar Icy Mantles Containing Propargyl Alcohol", *The Astrophysical Journal*, 2014, **798**, 72.

(109) J. M. Hollis, P. R. Jewell, F. J. Lovas, A. Remijan and H. Møllendal, "Green Bank Telescope Detection of New Interstellar Aldehydes: Propenal and Propanal", *The Astrophysical Journal*, 2004, **610**, L21–L24.

(110) P. Gorai, A. Das, L. Majumdar, S. K. Chakrabarti, B. Sivaraman and E. Herbst, "The Possibility of Forming Propargyl Alcohol in the Interstellar Medium", *Molecular Astrophysics*, 2017, **6**, 36–46.

(111) S. L. Miller and H. C. Urey, "Organic Compound Synthesis on the Primitive Earth: Several questions about the origin of life have been answered, but much remains to be studied", en, *Science*, 1959, **130**, 245–251.

(112) E. T. Parker, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, M. Callahan, A. Aubrey, A. Lazcano and J. L. Bada, "Primordial synthesis of amines and amino acids in a 1958 Miller H_2S-rich spark discharge experiment.", *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**, 5526–31.

(113) R. Hayatsu, M. H. Studier, L. P. Moore and E. Anders, "Purines and tri-azines in the Murchison meteorite", *Geochimica et Cosmochimica Acta*, 1975, **39**, 471–488.

(114) M Förstel, P Maksyutenko, B. M. Jones, B.-J. Sun, A. H. H. Chang and R. I. Kaiser, "Synthesis of urea in cometary model ices and implications for Comet 67P/Churyumov-Gerasimenko.", *Chemical communications*, 2015, **52**, 741–4.

(115) J. E. Elsila, D. P. Glavin and J. P. Dworkin, "Cometary glycine detected in samples returned by Stardust", *Meteoritics & Planetary Science*, 2009, **44**, 1323–1330.

(116) K. Altwegg, H. Balsiger, A. Bar-Nun, J.-J. Berthelier, A. Bieler, P. Bochsler, C. Briois, U. Calmonte, M. R. Combi, H. Cottin, J. De Keyser, F. Dhooghe, B. Fiethe, S. A. Fuselier, S. Gasc, T. I. Gombosi, K. C. Hansen, M. Haessig, A. Jäckel, E. Kopp, A. Korth, L. Le Roy, U. Mall, B. Marty, O. Mousis, T. Owen, H. Rème, M. Rubin, T. Sémon, C.-Y. Tzou, J. Hunter Waite and P. Wurz, "Prebiotic chemicals-amino acid and phosphorus in the coma of comet 67P/Churyumov-Gerasimenko", *Science Advances*, 2016, **2**, e1600285.

(117) M. P. Bernstein, J. E. Elsila, J. P. Dworkin, S. A. Sandford, L. J. Allamandola and R. N. Zare, "Side Group Addition to the Polycyclic Aromatic Hydrocarbon Coronene by Ultraviolet Photolysis in Cosmic Ice Analogs", *The Astrophysical Journal*, 2002, **576**, 1115–1120.

(118) G. M. Muñoz Caro, U. J. Meierhenrich, W. A. Schutte, B Barbier, A Arcones Segovia, H Rosenbauer, W. H.-P. Thiemann, A Brack and J. M. Greenberg, "Amino acids from ultraviolet irradiation of interstellar ice analogues.", *Nature*, 2002, **416**, 403–6.

(119) W. H. Sorrell, "Origin of Amino Acids and Organic Sugars in Interstellar Clouds", en, *The Astrophysical Journal*, 2001, **555**, L129–L132.

(120) V. Blagojevic, S. Petrie and D. K. Bohme, "Gas-phase syntheses for interstellar carboxylic and amino acids", *Monthly Notices of the Royal Astronomical Society*, 2003, **339**, L7–L11.

(121) D. E. Woon, "Pathways to Glycine and Other Amino Acids in Ultraviolet-irradiated Astrophysical Ices Determined via Quantum Chemical Modeling", en, *The Astrophysical Journal*, 2002, **571**, L177–L180.

(122) R. T. Garrod and S. L. W. Weaver, "Simulations of hot-core chemistry.", EN, *Chemical reviews*, 2013, **113**, 8939–60.

(123) A. Horn, H. Mollendal, O. Sekiguchi, E. Uggerud, H. Roberts, E. Herbst, A. A. Viggiano and T. D. Fridgen, "The Gas-Phase Formation of Methyl Formate in Hot Molecular Cores", en, *The Astrophysical Journal*, 2004, **611**, 605–614.

(124) N. Lavado, M. Ávalos, R. Babiano, P. Cintas, M. E. Light, J. L. Jiménez and J. C. Palacios, "On the Plausibility of Pseudosugar Formation in Cometary Ices and Oxygen-rich Tholins.", *Origins of life and evolution of the biosphere : the journal of the International Society for the Study of the Origin of Life*, 2016, **46**, 31–49.

(125) J. M. Hollis, F. J. Lovas and P. R. Jewell, "Interstellar Glycolaldehyde: The First Sugar", *The Astrophysical Journal*, 2000, **540**, L107–L110.

(126) A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, 1996.

(127) E. Schrödinger, "Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinem", *Annalen der Physik*, 1926, **384**, 734–756.

(128) E. Schrödinger, "Quantisierung als Eigenwertproblem", *Annalen der Physik*, 1926, **384**, 489–527.

(129) E. Schrödinger, "Quantisierung als Eigenwertproblem", *Annalen der Physik*, 1926, **386**, 109–139.

(130) E. Schrödinger, "Quantisierung als Eigenwertproblem", *Annalen der Physik*, 1926, **385**, 437–490.

(131) P. W. Atkins and J. De Paula, *Atkins' Physical chemistry*, Oxford University Press, 2010, p. 972.

(132) M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln", *Annalen der Physik*, 1927, **389**, 457–484.

(133) J. Binney and D. D. B. Skinner, *The physics of quantum mechanics*, p. 392.

(134) P. Echenique and J. L. Alonso, "A mathematical and computational review of Hartree-Fock SCF methods in Quantum Chemistry", *Molecular Physics*, 2007, 3057–3098.

(135) J. C. Slater, "The Theory of Complex Spectra", *Physical Review*, 1929, **34**, 1293–1322.

(136) E. U. Condon, "The Theory of Complex Spectra", *Physical Review*, 1930, **36**, 1121–1133.

(137) P.-O. Löwdin, "Quantum Theory of Many-Particle Systems. I. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction", *Physical Review*, 1955, **97**, 1474–1489.

(138) C. C. J. Roothaan, "New Developments in Molecular Orbital Theory", *Reviews of Modern Physics*, 1951, **23**, 69–89.

(139) J. Cioslowski, "Density-driven self-consistent-field method: Density-constrained correlation energies in the helium series", *Physical Review A*, 1991, **43**, 1223–1228.

(140) E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar and C. F. Fischer, "Ground-state correlation energies for two- to ten-electron atomic ions", *Physical Review A*, 1991, **44**, 7071–7083.

(141) E. Valderrama, E. V. Ludeña and J. Hinze, "Analysis of dynamical and nondynamical components of electron correlation energy by means of local-scaling density-functional theory", *The Journal of Chemical Physics*, 1997, **106**, 9227–9235.

(142) E. Ramos-Cordoba, P. Salvador and E. Matito, "Separation of dynamic and nondynamic correlation", *Physical Chemistry Chemical Physics*, 2016, **18**, 24015–24023.

(143) P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", *Physical Review*, 1964, **136**, B864–B871.

(144) W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", *Physical Review*, 1965, **140**, A1133–A1138.

(145) J. A. Pople, P. M. W. Gill, B. G. Johnson and P. M. W. Gill, "Kohn-Sham density-functional theory within a finite basis set", *Chemical Physics Letters*, 1992, **199**, 557.

(146) J. P. Perdew and W. Yue, "Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation", *Physical Review B*, 1986, **33**, 8800–8802.

(147) A. J. Cohen, P. Mori-Sánchez and W. Yang, "Challenges for Density Functional Theory", *Chemical Reviews*, 2012, **112**, 289–320.

(148) A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange", *The Journal of Chemical Physics*, 1993, **98**, 5648.

(149) C. Lee, W. Yang and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Physical Review B*, 1988, **37**, 785–789.

(150) C. Latouche, F. Palazzetti, D. Skouteris and V. Barone, "High-Accuracy Vibrational Computations for Transition-Metal Complexes Including Anharmonic Corrections: Ferrocene, Ruthenocene, and Osmocene as Test Cases", *Journal of Chemical Theory and Computation*, 2014, **10**, 4565–4573.

(151) M. Steinmetz and S. Grimme, "Benchmark study of the performance of density functional theory for bond activations with (Ni,Pd)-based transition-metal catalysts.", *ChemistryOpen*, 2013, **2**, 115–24.

(152) T. Weymuth, E. P. A. Couzijn, P. Chen and M. Reiher, "New Benchmark Set of Transition-Metal Coordination Reactions for the Assessment of Density Functionals", *Journal of Chemical Theory and Computation*, 2014, **10**, 3092–3103.

(153) N. Mardirossian and M. Head-Gordon, "Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V", *The Journal of Chemical Physics*, 2015, **142**, 074111.

(154) E. Runge and E. K. U. Gross, "Density-Functional Theory for Time-Dependent Systems", *Physical Review Letters*, 1984, **52**, 997–1000.

(155) E. Gross and W. Kohn, "Time-Dependent Density-Functional Theory", *Advances in Quantum Chemistry*, 1990, **21**, 255–291.

(156) M. A. Mosquera, "Action formalism of time-dependent density-functional theory", *Physical Review A*, 2013, **88**, 022515.

(157) M. E. Casida, in *Recent Advances in Density Functional Methods, vol 1*, ed. D. P. Chong, World Scientific, 1995, pp. 155–192.

(158) A. Dreuw and M. Head-Gordon, "Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin-Bacteriochlorin and Bacteriochlorophyll-Spheroidene Complexes", *J. Am. Chem. Soc.*, 2004, **126**, 4007–4016.

(159) C. Jamorski, M. E. Casida and D. R. Salahub, "Dynamic polarizabilities and excitation spectra from a molecular implementation of time-dependent density-functional response theory: N2 as a case study", *The Journal of Chemical Physics*, 1996, **104**, 5134–5147.

(160) S. Hirata and M. Head-Gordon, "Time-dependent density functional theory within the Tamm-Dancoff approximation", *Chemical Physics Letters*, 1999, **314**, 291–299.

(161) A. Chantzis, A. D. Laurent, C. Adamo and D. Jacquemin, "Is the Tamm-Dancoff Approximation Reliable for the Calculation of Absorption and Fluorescence Band Shapes?", *Journal of Chemical Theory and Computation*, 2013, **9**, 4517–4525.

(162) S. Miertuš, E. Scrocco and J. Tomasi, "Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects", *Chemical Physics*, 1981, **55**, 117–129.

(163) C. Kittel and P. McEuen, *Introduction to solid state physics*, Wiley, 1996.

(164) J. J. Kohanoff, *Electronic structure calculations for solids and molecules : theory and computational methods*, Cambridge University Press, 2006, p. 348.

(165) F. Bloch, "Über die Quantenmechanik der Elektronen in Kristallgittern", *Zeitschrift fr Physik*, 1929, **52**, 555–600.

(166) N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976.

(167) O. de Souza and R. Ornstein, "Effect of periodic box size on aqueous molecular dynamics simulation of a DNA dodecamer with particle-mesh Ewald method", *Biophysical Journal*, 1997, **72**, 2395–2397.

(168) A. R. Leach, *Molecular modelling : principles and applications*, Prentice Hall, 2001, p. 744.

(169) G Kresse and J Hafner, "Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements", *Journal of Physics: Condensed Matter*, 1994, **6**, 8245–8257.

(170) P. E. Blöchl, C. J. Först and J. Schimpl, "Projector augmented wave method:ab initio molecular dynamics with full wave functions", *Bulletin of Materials Science*, 2003, **26**, 33–41.

(171) A. C. Wahl and G. Das, in *Methods of Electronic Structure Theory*, Springer US, Boston, MA, 1977, pp. 51–78.

(172) B. O. Roos, P. R. Taylor and P. E. Sigbahn, "A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach", *Chemical Physics*, 1980, **48**, 157–173.

(173) P. E. M. Siegbahn, in *Lecture Notes in Quantum Chemistry: European Summer School in Quantum Chemistry*, ed. B. O. Roos, Springer Berlin Heidelberg, Berlin, Heidelberg, 1992, pp. 255–293.

(174) P. A. Malmqvist, A. Rendell and B. O. Roos, "The restricted active space self-consistent-field method, implemented with a split graph unitary group approach", *The Journal of Physical Chemistry*, 1990, **94**, 5477–5482.

(175)   M. W. Schmidt and M. S. Gordon, "The Construction and Interpretation of MCSCF Wavefunctions", *Annual Review of Physical Chemistry*, 1998, **49**, 233–266.

(176)   K. Andersson, P. Malmqvist and B. O. Roos, "Second-order perturbation theory with a complete active space self-consistent field reference function", *The Journal of Chemical Physics*, 1992, **96**, 1218–1226.

(177)   K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, "Second-order perturbation theory with a CASSCF reference function", *The Journal of Physical Chemistry*, 1990, **94**, 5483–5488.

(178)   U. Burkert and N. L. Allinger, *Molecular mechanics*, American Chemical Society, 1982, p. 339.

(179)   N. Allinger, "Calculation of Molecular Structure and Energy by Force-Field Methods", *Advances in Physical Organic Chemistry*, 1976, **13**, 1–82.

(180)   S. Dapprich, I. Komáromi, K. Byun, K. Morokuma and M. J. Frisch, "A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives", *Journal of Molecular Structure: Theochem*, 1999, **461-462**, 1–21.

(181)   T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery, K. Morokuma and M. J. Frisch, "Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM", *Journal of Chemical Theory and Computation*, 2006, **2**, 815–826.

(182)   P. Ehrenfreund, M. P. Bernstein, J. P. Dworkin, S. A. Sandford and L. J. Allamandola, "The Photostability of Amino Acids in Space", *The Astrophysical Journal*, 2001, **550**, L95–L99.

(183)   G. M. Muñoz Caro, U. J. Meierhenrich, W. A. Schutte, B. Barbier, A. Arcones Segovia, H. Rosenbauer, W. H.-P. Thiemann, A. Brack and J. M. Greenberg, "Amino acids from ultraviolet irradiation of interstellar ice analogues", *Nature*, 2002, **416**, 403–406.

(184)   J. E. Elsila, J. P. Dworkin, M. P. Bernstein, M. P. Martin and S. A. Sandford, "Mechanisms of Amino Acid Formation in Interstellar Ice Analogs", *The Astrophysical Journal*, 2007, **660**, 911–918.

(185)   D. M. Koch, C. Toubin, G. H. Peslherbe and J. T. Hynes, "A Theoretical Study of the Formation of the Aminoacetonitrile Precursor of Glycine on Icy Grain Mantles in the Interstellar Medium", *Journal of Physical Chemistry*, 2008, **112**, 2972–2980.

(186) C. D. Georgiou and D. W. Deamer, "Lipids as Universal Biomarkers of Extraterrestrial Life", *Astrobiology*, 2014, **14**, 541–549.

(187) K. E. Smith, P. A. Gerakines and M. P. Callahan, "Metabolic precursors in astrophysical ice analogs: implications for meteorites and comets", *Chemical Communications*, 2015, **51**, 11787–11790.

(188) B. Zuckerman, J. A. Ball and C. A. Gottlieb, "Microwave Detection of Interstellar Formic Acid", *The Astrophysical Journal*, 1971, **163**, L41.

(189) A. Remijan, L. E. Snyder, S.-. Liu, D. Mehringer and Y.-J. Kuan, "Acetic Acid in the Hot Cores of Sagitarrius B2(N) and W51", *The Astrophysical Journal*, 2002, **576**, 264–273.

(190) Y.-S. J. Shiao, L. W. Looney, A. J. Remijan, L. E. Snyder and D. N. Friedel, "First acetic acid survey with carma in hot molecular cores", *The Astrophysical Journal*, 2010, **716**, 286–298.

(191) A. Palau, C. Walsh, Á. Sánchez-Monge, J. M. Girart, R. Cesaroni, I. Jiménez-Serra, A. Fuente, L. A. Zapata and R. Neri, "Complex Organic Molecules tracing shocks along the outflow cavity in the high-mass protostar IRAS20126+4104", *Monthly Notices of the Royal Astronomical Society*, 2017, **467**, stx004.

(192) C. J. Bennett, T. Hama, Y. S. Kim, M. Kawasaki and R. I. Kaiser, "Laboratory Studies on the Formation of Formic Acid (HCOOH) in Interstellar and Cometary Ices", *The Astrophysical Journal*, 2011, **727**, 27.

(193) Y. S. Kim and R. I. Kaiser, "Abiotic formation of carboxylic acids (RCOOH) in interstellar and solar model ices", *The Astrophysical Journal*, 2010, **725**, 1002–1010.

(194) C. Zhu, A. M. Turner, M. J. Abplanalp and R. I. Kaiser, "Formation and High-order Carboxylic Acids (RCOOH) in Interstellar Analogous Ices of Carbon Dioxide (CO_2) and Methane(CH_4) ", *The Astrophysical Journal Supplement Series*, 2018, **234**, 15.

(195) Z. Martins, J. S. Watson, M. A. Sephton, O. Botta, P. Ehrenfreund and I. Gilmour, "Free dicarboxylic and aromatic acids in the carbonaceous chondrites Murchison and Orgueil", *Meteoritics & Planetary Science*, 2006, **41**, 1073–1080.

(196) J. A. Ball, C. A. Gottlieb, A. E. Lilley and H. E. Radford, "Detection of Methyl Alcohol in Sagittarius", *The Astrophysical Journal*, 1970, **162**, L203.

(197) B. M. McMurtry, S. E. J. Saito, A. M. Turner, H. K. Chakravarty and R. I. Kaiser, "On the formation of benzoic acid and higher-order benzene carboxylic acids in the interstellar model ice grains", *The Astrophysical Journal*, 2016, **831**, 174.

(198) R. L. James, N. C. Jones, S. V. Hoffmann and A. Dawes, "VUV spectroscopy of an electron irradiated benzene:carbon dioxide interstellar ice analogue", *RSC Advances*, 2019, **9**, 5453–5459.

(199) Y. Zhao and D. G. Truhlar, "The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function", *Theoretical Chemistry Accounts*, 2008, **120**, 215–241.

(200) F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy", *Physical Chemistry Chemical Physics*, 2005, **7**, 3297.

(201) F. Weigend, "Accurate Coulomb-fitting basis sets for H to Rn", *Physical Chemistry Chemical Physics*, 2006, **8**, 1057.

(202) G. Kresse and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set", *Computational Materials Science*, 1996, **6**, 15–50.

(203) J. Paier, R. Hirschl, M. Marsman and G. Kresse, "The Perdew-Burke-Ernzerhof exchange-correlation functional applied to the G2-1 test set using a plane-wave basis set", *The Journal of Chemical Physics*, 2005, **122**, 234102.

(204) G. A. Sim, J. M. Robertson and T. H. Goodwin, "The crystal and molecular structure of benzoic acid", *Acta Crystallographica*, 1955, **8**, 157–164.

(205) T. G. D. van Schalkwyk, "Phthalic acid at low temperature", *Acta Crystallographica*, 1954, **7**, 775–775.

(206) R. Alcala and S. Martinez-Carrera, "The crystal structure of isophthalic acid", *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 1972, **B28**, 1671.

(207) M. Bailey and C. J. Brown, "The crystal structure of terephthalic acid", *Acta Crystallographica*, 1967, **22**, 387–391.

(208) M. E. Casida, C. Jamorski, K. C. Casida and D. R. Salahub, "Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold", *Journal of Chemical Physics*, 1998, 4439.

(209) R. Bauernschmitt and R. Ahlrichs, "Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory", *Chemical Physics Letters*, 1996, **256**, 454–464.

(210) G. Bruno, L. Randaccio and IUCr, "A refinement of the benzoic acid structure at room temperature", *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 1980, **36**, 1711–1712.

(211) A. L. Rheingold, "CCDC 8844328", *CSD Communication*, 2011, DOI: `10.5517/ccxbld6`.

(212) J. L. Derissen, "The crystal structure of isophthalic acid", *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 1974, **30**, 2764–2765.

(213) M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller and F. Bechstedt, "Linear optical properties in the projector-augmented wave methodology", *Physical Review B*, 2006, **73**, 045112.

(214) H. Ratajczak, "Charge-transfer properties of the hydrogen bond. I. Theory of the enhancement of dipole moment of hydrogen-bonded systems", *The Journal of Physical Chemistry*, 1972, **76**, 3000–3004.

(215) B. Cullity and S. Stock, *Elements of X-ray Diffraction, Third Edition*, 2001.

(216) D. Jacquemin, V. Wathelet, E. A. Perpète and C. Adamo, "Extensive TD-DFT Benchmark: Singlet-Excited States of Organic Molecules", *Journal of Chemical Theory and Computation*, 2009, **5**, 2420–2435.

(217) D. A. Williams and E. Herbst, "It's a dusty Universe: surface science in space", *Surface Science*, 2002, **500**, 823–837.

(218) B. Draine, "Interstellar Dust Grains", *Annual Review of Astronomy and Astrophysics*, 2003, **41**, 241–289.

(219) L. J. Allamandola, G. G. M. Tielens and J. R. Barker, "Interstellar polycyclic aromatic hydrocarbons - The infrared emission bands, the excitation/emission mechanism, and the astrophysical implications", *The Astrophysical Journal Supplement Series*, 1989, **71**, 733.

(220) L. J. Allamandola, D. M. Hudgins and S. A. Sandford, "Modeling the Unidentified Infrared Emission with Combinations of Polycyclic Aromatic Hydrocarbons", *The Astrophysical Journal*, 1999, **511**, L115–L119.

(221) T. Q. Zhao, Q. Li, B. S. Liu, R. K. E. Gover, P. J. Sarre and A. S.-C. Cheung, "Laboratory astrochemistry: catalytic conversion of acetylene to polycyclic aromatic hydrocarbons over SiC grains", *Phys. Chem. Chem. Phys.*, 2016, **18**, 3489–3496.

(222) T. Stein, B. Bandyopadhyay, T. P. Troy, Y. Fang, O. Kostko, M. Ahmed and M. Head-Gordon, "Ab initio dynamics and photoionization mass spectrometry reveal ion-molecule pathways from ionized acetylene clusters to benzene cation", *Proceedings of the National Academy of Sciences*, 2017, **114**, E4125–E4133.

(223) K. I. Öberg, R. T. Garrod, E. F. van Dishoeck and H. Linnartz, "Formation rates of complex organics in UV irradiated CH_3OH -rich ices", en, *Astronomy and Astrophysics*, 2009, **504**, 891–913.

(224) C. Meinert, I. Myrgorodska, P. de Marcellus, T. Buhse, L. Nahon, S. V. Hoffmann, L. L. S. DHendecourt and U. J. Meierhenrich, "Ribose and related sugars from ultraviolet irradiation of interstellar ice analogs", en, *Science*, 2016, **352**, 208–212.

(225) M. S. Gudipati and L. J. Allamandola, "Facile Generation and Storage of Polycyclic Aromatic Hydrocarbon Ions in Astrophysical Ices", *The Astrophysical Journal*, 2003, **596**, L195–L198.

(226) J. Bouwman, D. M. Paardekooper, H. M. Cuppen, H. Linnartz and L. J. Allamandola, "Real-time optical spectroscopy of vacuum ultraviolet irradiated pyrene: H_2O interstellar ice", *The Astrophysical Journal*, 2009, **700**, 56–62.

(227) M. P. Bernstein, S. A. Sandford, A. L. Mattioda and L. J. Allamandola, "Near and Mid-Infrared Laboratory Spectra of PAH Cations in Solid $H2O$ ", *The Astrophysical Journal*, 2007, **664**, 1264–1272.

(228) A. M. Cook, A. Ricca, A. L. Mattioda, J. Bouwman, J. Roser, H. Linnartz, J. Bregman and L. J. Allamandola, "Photochemistry of polycyclic aromatic hydrocarbons in cosmic water ice: the role of PAH ionization and concentration", *The Astrophysical Journal*, 2015, **799**.

(229) J. Y. Park and D. E. Woon, "Theoretical investigation of OCN- charge-transfer complexes in condensed-phase media: Spectroscopic properties in amorphous ice", *Journal of Physical Chemistry A*, 2004, **108**, 6589–6598.

(230) M. S. Gudipati and L. J. Allamandola, "Unusual Stability of Polycyclic Aromatic Hydrocarbon Radical Cations in Amorphous Water Ices up to 120 K: Astronomical Implications", *The Astrophysical Journal*, 2006, **638**, 286–292.

(231) J. W. Boyle, J. A. Ghormley, C. J. Hochanadel and J. F. Riley, "Production of hydrated electrons by flash photolysis of liquid water with light in the first continuum", *Journal of Physical Chemistry*, 1969, **73**, 2886–2890.

(232) M. Faubel, B. Steiner and J. P. Toennies, "Photoelectron spectroscopy of liquid water, some alcohols, and pure nonane in free micro jets", *The Journal of Chemical Physics*, 1997, **106**, 9013–9031.

(233) T. Yanai, D. P. Tew and N. C. Handy, "A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP)", *Chemical Physics Letters*, 2004, **393**, 51–57.

(234) A. D. McLean and G. S. Chandler, "Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z = 11 - 18", *The Journal of Chemical Physics*, 1980, **72**, 5639–5648.

(235) S. Grimme, S. Ehrlich and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory", *Journal of Computational Chemistry*, 2011, **32**, 1456–1465.

(236) J. W. Ponder and D. A. Case, "Force fields for protein simulations.", *Advances in protein chemistry*, 2003, **66**, 27–85.

(237) W. L. Jorgensen and J. Tirado-Rives, "The OPLS (optimized potentials for liquid simulations) potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin", *Journal of the American Chemical Society*, 1988, **110**, 1657–1666.

(238) W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids", 1996.

(239) G. H. Diercksen, B. O. Roos and A. J. Sadlej, "Legitimate calculation of first-order molecular properties in the case of limited CI functions. Dipole moments", *Chemical Physics*, 1981, **59**, 29–39.

(240) G. H. F. Diercksen and A. J. Sadlej, "Perturbation theory of the electron correlation effects for atomic and molecular properties. Second- and third-order correlation corrections to molecular dipole moments and polarizabilities", *The Journal of Chemical Physics*, 1981, **75**, 1253–1266.

(241) N. C. Handy and H. F. Schaefer, "On the evaluation of analytic energy derivatives for correlated wave functions", *The Journal of Chemical Physics*, 1984, **81**, 5031–5033.

(242) K. B. Wiberg, C. M. Hadad, T. J. LePage, C. M. Breneman and M. J. Frisch, "Analysis of the effect of electron correlation on charge density distributions", *The Journal of Physical Chemistry*, 1992, **96**, 671–679.

(243) F. Trani, G. Scalmani, G. Zheng, I. Carnimeo, M. J. Frisch and V. Barone, "Time-Dependent Density Functional Tight Binding: New Formulation and Benchmark of Excited States", *Journal of Chemical Theory and Computation*, 2011, **7**, 3304–3313.

(244) M. P. Bernstein, S. A. Sandford and L. J. Allamandola, "The Mid-Infrared Absorption Spectra of Neutral Polycyclic Aromatic Hydrocarbons in Conditions Relevant to Dense Interstellar Clouds", *The Astrophysical Journal Supplement Series*, 2005, **161**, 53–64.

(245) G. Nemeth, H. Selzle and E. Schlag, "Magnetic ZEKE experiments with mass analysis", *Chemical Physics Letters*, 1993, **215**, 151–155.

(246) M. S. Deleuze, L. Claes, E. S. Kryachko and J.-P. Fran?ois, "Benchmark theoretical study of the ionization threshold of benzene and oligoacenes", *The Journal of Chemical Physics*, 2003, **119**, 3106–3119.

(247) M. Parac and S. Grimme, "A TDDFT study of the lowest excitation energies of polycyclic aromatic hydrocarbons", *Chemical Physics*, 2003, **292**, 11–21.

(248) M. C. R. Cockett, H. Ozeki, K. Okuyama and K. Kimura, "Vibronic coupling in the ground cationic state of naphthalene: A laser threshold photoelectron [zero kinetic energy (ZEKE) photoelectron] spectroscopic study", *The Journal of Chemical Physics*, 1993, **98**, 7763–7772.

(249) J. W. Hager and S. C. Wallace, "Two-Laser Photoionization Supersonic Jet Mass Spectrometry of Aromatic Molecules", *Analytical Chemistry*, 1988, **60**, 5–10.

(250) M. I. Shchuka, A. L. Motyka and M. R. Topp, "Two-photon threshold ionization spectroscopy of perylene and van der Waals complexes", *Chemical Physics Letters*, 1989, **164**, 87–95.

(251) E. Clar, J. M. Robertson, R. Schloegl and W. Schmidt, "Photoelectron spectra of polynuclear aromatics. 6. Applications to structural elucidation: "circumanthracene"", *Journal of the American Chemical Society*, 1981, **103**, 1320–1328.

(252) F. R. Clemente, T. Vreven and M. J. Frisch, in *Quantum Biochemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010, pp. 61–83.

(253) M. S. Gudipati and L. J. Allamandola, "Polycyclic Aromatic Hydrocarbon Ionization Energy Lowering in Water Ices", *The Astrophysical Journal*, 2004, **615**, L177–L180.

(254) R. D. Brown, J. G. Crofts, P. D. Godfrey, F. F. Gardner, B. J. Robinson and J. B. Whiteoak, "Discovery of interstellar methyl formate", *The Astrophysical Journal*, 1975, **197**, L29.

(255) D. M. Mehringer, L. E. Snyder, Y. Miao and F. J. Lovas, "Detection and Confirmation of Interstellar Acetic Acid", *The Astrophysical Journal*, 1997, **480**, L71–L74.

(256) J. M. Hollis, S. N. Vogel, L. E. Snyder, P. R. Jewell and F. J. Lovas, "The Spatial Scale of Glycolaldehyde in the Galactic Center", *The Astrophysical Journal*, 2001, **554**, L81–L85.

(257) J. B. Bergner, K. I. Oberg and M. Rajappan, "Methanol Formation via Oxygen Insertion Chemistry in Ices", *The Astrophysical Journal*, 2017, **845**, 29.

(258) S. B. Charnley, M. E. Kress, A. G. G. M. Tielens and T. J. Millar, "Interstellar Alcohols", *The Astrophysical Journal*, 1995, **448**, 232.

(259) J. C. Laas, R. T. Garrod, E. Herbst and S. L. Widicus Weaver, "Contributions From Grain Surface and Gas Phase Chemistry to the Formation of Methyl Formate and its Structural Isomers", en, *The Astrophysical Journal*, 2011, **728**, 71.

(260) J. L. Neill, M. T. Muckle, D. P. Zaleski, A. L. Steber, B. H. Pate, V. Lattanzi, S. Spezzano, M. C. McCarthy and A. J. Remijan, "Laboratory and Tentative Interstellar Detection of Trans-methyl Formate Using the Publicly Available Green Bank Telescope PRIMOS Survey", en, *The Astrophysical Journal*, 2012, **755**, 153.

(261) P. Modica and M. E. Palumbo, "Formation of methyl formate after cosmic ion irradiation of icy grain mantles", en, *Astronomy and Astrophysics*, 2010, **519**, A22.

(262) A. Occhiogrosso, S. Viti, P. Modica and M. E. Palumbo, "A study of methyl formate in astrochemical environments", *Monthly Notices of the Royal Astronomical Society*, 2011, **418**, 1923–1927.

(263) K.-J. Chuang, G. Fedoseev, S. Ioppolo, E. F. van Dishoeck and H. Linnartz, "H-atom addition and abstraction reactions in mixed CO, H_2CO and CH_3OH ices - an extended view on complex organic molecule formation", *Monthly Notices of the Royal Astronomical Society*, 2015, **455**, 1702–1712.

(264) G. Fedoseev, H. M. Cuppen, S. Ioppolo, T. Lamberts and H. Linnartz, "Experimental evidence for glycolaldehyde and ethylene glycol formation by surface hydrogenation of CO molecules under dense molecular cloud conditions", *Monthly Notices of the Royal Astronomical Society*, 2015, **448**, 1288–1297.

(265) M. Bertin, C. Romanzin, M. Doronin, L. Philippe, P. Jeseck, N. Litgerink, H. Linnartz, X. Michaut and J.-H. Fillion, "UV photodesorption of methanol in pure and CO-rich ices: desorption rates of the intact molecule and of the photofragments", *The Astrophysical Journal Letters*, 2016, **817**, L12.

(266) T. L. Cottrell, *The Strengths of Chemical Bonds*, Academic Press, 1954.

(267) H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz, "Molpro: a general-purpose quantum chemistry program package", *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2012, **2**, 242–253.

(268) H. Werner and P. J. Knowles, "A second order multiconfiguration SCF procedure with optimum convergence", *The Journal of Chemical Physics*, 1985, **82**, 5053–5063.

(269) P. J. Knowles and H.-J. Werner, "An efficient second-order MC SCF method for long configuration expansions", *Chemical Physics Letters*, 1985, **115**, 259–267.

(270) T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen", *The Journal of Chemical Physics*, 1989, **90**, 1007–1023.

(271) M. J. Frisch, M. Head-Gordon and J. A. Pople, "Semi-direct algorithms for the MP2 energy and gradient", *Chemical Physics Letters*, 1990, **166**, 281–289.

(272) M. J. Frisch, M. Head-Gordon and J. A. Pople, "A direct MP2 gradient method", *Chemical Physics Letters*, 1990, **166**, 275–280.

(273) C. W. Bauschlicher, S. R. Langhoff and S. P. Walch, "Theoretical study of the bond dissociation energies of methanol", *The Journal of Chemical Physics*, 1992, **96**, 450–454.

(274) J. Cernicharo, N. Marcelino, E. Roueff, M. Gerin, A. Jiménez-Escobar and G. M. Muñoz Caro, "Discovery of the Methoxy Radical, CH_3O, Toward B1: Dust Grain and Gas-phase Chemistry in Cold Dark Clouds", *The Astrophysical Journal*, 2012, **759**, L43.

(275) P. M. Woods, G. Kelly, S. Viti, B. Slater, W. A. Brown, F. Puletti, D. J. Burke and Z. Raza, "On The Formation of Glycolaldehyde in Dense Molecular Cores", *The Astrophysical Journal*, 2012, **750**, 19.

(276) P. M. Woods, B. Slater, Z. Raza, S. Viti, W. A. Brown and D. J. Burke, "Glycolaldehyde Formation via the Dimerization of the Formyl Radical", *The Astrophysical Journal*, 2013, **777**, 90.

(277) P. R. Schreiner, J. P. Wagner, H. P. Reisenauer, D. Gerbig, D. Ley, J. Sarka, A. G. Császár, A. Vaughn and W. D. Allen, "Domino Tunneling", *Journal of the American Chemical Society*, 2015, **137**, 7828–7834.

(278) F. Wöhler, "Ueber künstliche Bildung des Harnstoffs", *Annalen der Physik und Chemie*, 1828, **88**, 253–256.

(279) M. P. Robertson and S. L. Miller, "An efficient prebiotic synthesis of cytosine and uracil", *Nature*, 1995, **375**, 772–774.

(280) Y. A. Jeilani, C. Fearce and M. T. Nguyen, "Acetylene as an essential building block for prebiotic formation of pyrimidine bases on Titan", *Physical Chemistry Chemical Physics*, 2015, **17**, 24294–24303.

(281) B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, "The RNA World and the Origins of Life", 2002.

(282) E Kinne-Saffran and R. K. Kinne, "Vitalism and synthesis of urea. From Friedrich Wöhler to Hans A. Krebs.", *American journal of nephrology*, 1999, **19**, 290–4.

(283) S. L. Miller and L. E. Orgel, *The origins of life on the earth*, Prentice-Hall, 1974, p. 229.

(284) A. J. Remijan, L. E. Snyder, B. A. McGuire, H.-L. Kuo, L. W. Looney, D. N. Friedel, G. Y. Golubiatnikov, F. J. Lovas, V. V. Ilyushin, E. A. Alekseev, S. F. Dyubko, B. J. McCall and J. M. Hollis, "Observational results of a multi-telescope campaign in search of interstellar urea $(NH_2)_2CO$ ", *The Astrophysical Journal*, 2014, **783**, 77.

(285) J. M. Hollis, F. J. Lovas, A. J. Remijan, P. R. Jewell, V. V. Ilyushin and I. Kleiner, "Detection of Acetamide (CH_3CONH_2): The Largest Interstellar Molecule with a Peptide Bond", *The Astrophysical Journal*, 2006, **643**, L25–L28.

(286) R. H. Rubin, J. Swenson, G. W., R. C. Benson, H. L. Tigelaar and W. H. Flygare, "Microwave Detection of Interstellar Formamide", *The Astrophysical Journal*, 1971, **169**, L39.

(287) N. Fourikis, K. Takagi and M. Morimoto, "Detection of Interstellar Methylamine by its $2_{02} \rightarrow 1_{10}A_a$-State Transition", *The Astrophysical Journal*, 1974, **191**, L139.

(288) R. Spezia, Y. Jeanvoine, W. L. Hase, K. Song and A. Largo, "Sythesis of formamide and related organic species in the interstellar medium via chemical dynamics simulations", *The Astrophysical Journal*, 2016, **826**, 107.

(289) Z. Kaňuchová, R. G. Urso, G. A. Baratta, J. R. Brucato, M. E. Palumbo and G. Strazzulla, "Synthesis of formamide and isocyanic acid after ion irradiation of frozen gas mixtures", *Astronomy & Astrophysics*, 2016, **585**, A155.

(290) L. Song and J. Kästner, "Formation of the prebiotic molecule NH_2CHO on astronomical amorphous solid water surfaces: accurate tunneling rate calculations", *Phys. Chem. Chem. Phys.*, 2016, **18**, 29278–29285.

(291) P. Redondo, C. Barrientos and A. Largo, "Some insights into formamide formation through gas-phase reactions in the interstellar medium", *The Astrophysical Journal*, 2013, **780**, 181.

(292) R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, "Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions", *The Journal of Chemical Physics*, 1980, **72**, 650–654.

(293) M. Szwarc, "The Dissociation Energy of the First NH Bond in Ammonia", *The Journal of Chemical Physics*, 1949, **17**, 505–507.

(294) E. F. van Dishoeck, D. J. Jansen, P. Schilke and T. G. Phillips, "Detection of the Interstellar NH_2 Radical", *The Astrophysical Journal*, 1993, **416**, L83.

(295) E. I. Izgorodina, D. R. B. Brittain, J. L. Hodgson, K. E. H, Y. L. Ching, M. Namazain and M. L. Coote, "Should Contemporary Density Functional Theory Methods Be Used to Study the Thermodynamics of Radical Reactions?", *Journal of Physical Chemistry A*, 2007, **111**, 10754–10768.

(296) L. Zhao, M. B. Prendergast, R. I. Kaiser, B. Xu, W. Lu, U. Ablikim, M. Ahmed, A. D. Oleinikov, V. N. Azyazov, A. M. Mebel, A. H. Howlader and S. F. Wnuk, "Reactivity of the Indenyl Radical C_9H_7 with Acetylene C_2H_2 and Vinylacetylene C_4H_4", *ChemPhysChem*, 2019, **20**, 1437–1447.

(297) M. N. R. Ashfold, R. A. Ingle, T. N. V. Karsili and J. Zhang, "Photoinduced C-H bond fission in prototypical organic molecules and radicals", *Physical Chemistry Chemical Physics*, 2019, **21**, 13880–13901.

(298) M. K. Sharma, "LVG analysis of amidogen radical (NH_2) found in interstellar medium and in cometary material", *Molecular Astrophysics*, 2019, **15**, 1–7.

(299) T Butscher, F Duvernay, G Danger, R Torro, G Lucas, Y Carissan, D Hagebaum-Reignier and T Chiavassa, "Radical-assisted polymerization in interstellar ice analogues: formyl radical and polyoxymethylene", *Monthly Notices of the Royal Astronomical Society*, 2019, **486**, 1953–1963.

(300) T. Butscher, F. Duvernay, A. Rimola, M. Segado-Centellas and T. Chiavassa, "Radical recombination in interstellar ices, a not so simple mechanism", *Physical Chemistry Chemical Physics*, 2017, **19**, 2857.

(301) K. I. Öberg, "Photochemistry and Astrochemistry: Photochemical Pathways to Interstellar Complex Organic Molecules", *Chemical Reviews*, 2016, **116**, 9631–9663.

(302) F. S. Brigiano, Y. Jeanvoine, A. Largo and R. Spezia, "The formation of urea in space I. Ion-molecule, neutral-neutral, and radical gas-phase reactions", 2017.

(303) S. J. Blanksby and G. B. Ellison, "Bond Dissociation Energies of Organic Molecules", *Acc. Chem. Res.*, 2003, **36**, 255–263.

(304) D. P. Bartel and P. J. Unrau, "Constructing an RNA world", *Trends in Cell Biology*, 1999, **9**, M9–M13.

(305) L. E. Orgel, "Some Consequences of the RNA World Hypothesis", *Origins of Life and Evolution of the Biosphere*, 2003, **33**, 211–218.

(306) L. E. Orgel, "Prebiotic Chemistry and the Origin of the RNA World", *Critical Reviews in Biochemistry and Molecular Biology*, 2004, **39**, 99–123.

(307) R. Stern and M. J. Jedrzejas, "Carbohydrate Polymers at the Center of Life's Origins: The Importance of Molecular Processivity", *Chemical Reviews*, 2008, **108**, 5061–5085.

(308) D. L. D. L. Nelson, M. M. Cox and A. L. Lehninger, *Lehninger principles of biochemistry*, W.H. Freeman and Company, 2013.

(309) G. Cooper, N. Kimmich, W. Belisle, J. Sarinana, K. Brabham and L. Garrel, "Carbonaceous meteorites as a source of sugar-related organic compounds for the early Earth", *Nature*, 2001, **414**, 879–883.

(310) G. Cooper and A. C. Rios, "Enantiomer excesses of rare and common sugar derivatives in carbonaceous meteorites.", *Proceedings of the National Academy of Sciences of the United States of America*, 2016, **113**, E3322–31.

(311) A. Butlerow, "Bildung einer zuckerartigen Substanz durch Synthese", *Annalen der Chemie und Pharmacie*, 1861, **120**, 295–298.

(312) R. Breslow, "On the mechanism of the formose reaction", *Tetrahedron Letters*, 1959, **1**, 22–26.

(313) L. E. Snyder, D. Buhl, B. Zuckerman and P. Palmer, "Microwave Detection of Interstellar Formaldehyde", *Physical Review Letters*, 1969, **22**, 679–681.

(314) O. A. Snytnikova, A. N. Simonov, O. P. Pestunova, V. N. Parmon and Y. P. Tsentalovich, "Study of the photoinduced formose reaction by flash and stationary photolysis", *Mendeleev Communications*, 2006, **16**, 9–11.

(315) S. Pizzarello and A. L. Weber, "Prebiotic amino acids as asymmetric catalysts.", *Science (New York, N.Y.)*, 2004, **303**, 1151.

(316) A. Córdova, I. Ibrahem, J. Casas, H. Sundén, M. Engqvist and E. Reyes, "Amino Acid Catalyzed Neogenesis of Carbohydrates: A Plausible Ancient Transformation", *Chemistry - A European Journal*, 2005, **11**, 4772–4784.

(317) C. Anastasi, M. A. Crowe, M. W. Powner and J. D. Sutherland, "Direct Assembly of Nucleoside Precursors from Two- and Three-Carbon Units", *Angewandte Chemie International Edition*, 2006, **45**, 6176–6179.

(318) P. de Marcellus, C. Meinert, I. Myrgorodska, L. Nahon, T. Buhse, L. L. S. D'Hendecourt and U. J. Meierhenrich, "Self-assembling amphiphilic molecules: Synthesis in simulated interstellar/precometary ices", *PNAS*, 2015, **98**, 815–819.

(319) M. Nuevo, G. Cooper and S. A. Sandford, "Deoxyribose and deoxysugar derivatives from photoprocessed astrophysical ice analogues and comparison to meteorites", *Nature Communications*, 2018, **9**, 5276.

(320) K. Kvenvolden, J. Lawless, K. Pering, E. Peterson, J. Flores, C. Ponnamperuma, I. R. Kaplan and C. Moore, "Evidence for Extraterrestrial Amino-acids and Hydrocarbons in the Murchison Meteorite", *Nature*, 1970, **228**, 923–926.

(321) M. H. Engel and S. A. Macko, "Isotopic evidence for extraterrestrial non-racemic amino acids in the Murchison meteorite", *Nature*, 1997, **389**, 265–268.

(322) B. H. Patel, C. Percivalle, D. J. Ritson, C. D. Duffy and J. D. Sutherland, "Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism", *Nature Chemistry*, 2015, **7**, 301–307.

(323) L. Burroughs, M. E. Vale, J. A. R. Gilks, H. Forintos, C. J. Hayes and P. A. Clarke, "Efficient asymmetric organocatalytic formation of erythrose and threose under aqueous conditions", *Chemical Communications*, 2010, **46**, 4776.

(324) L. Burroughs, P. A. Clarke, H. Forintos, J. A. R. Gilks, C. J. Hayes, M. E. Vale, W. Wade and M. Zbytniewski, "Asymmetric organocatalytic formation of protected and unprotected tetroses under potentially prebiotic conditions", *Organic & Biomolecular Chemistry*, 2012, **10**, 1565.

(325) H.-J. Kim, A. Ricardo, H. I. Illangkoon, M. J. Kim, M. A. Carrigan, F. Frye and S. A. Benner, "Synthesis of Carbohydrates in Mineral-Guided Prebiotic Cycles", *Journal of the American Chemical Society*, 2011, **133**, 9457–9468.

(326) S. A. Benner, H.-J. Kim and M. A. Carrigan, "Asphalt, Water, and the Prebiotic Synthesis of Ribose, Ribonucleosides, and RNA", *Accounts of Chemical Research*, 2012, **45**, 2025–2034.

(327) J. Kofoed, J.-L. Reymond and T. Darbre, "Prebiotic carbohydrate synthesis: zinc-proline catalyzes direct aqueous aldol reactions of α-hydroxy aldehydes and ketones", *Organic & Biomolecular Chemistry*, 2005, **3**, 1850.

(328) S. Pizzarello and A. L. Weber, "Stereoselective Syntheses of Pentose Sugars Under Realistic Prebiotic Conditions", *Origins of Life and Evolution of Biospheres*, 2010, **40**, 3–10.

(329) R. Breslow and C. Appayee, "Transketolase reaction under credible prebiotic conditions.", *Proceedings of the National Academy of Sciences of the United States of America*, 2013, **110**, 4184–7.

(330) M. Tena-Solsona, J. Nanda, S. Díaz-Oltra, A. Chotera, G. Ashkenasy and B. Escuder, "Emergent Catalytic Behavior of Self-Assembled Low Molecular Weight Peptide-Based Aggregates and Hydrogels", *Chemistry - A European Journal*, 2016, **22**, 6687–6694.

(331) A. M. Steer, N. Bia, D. K. Smith and P. A. Clarke, "Prebiotic synthesis of 2-deoxy- d -ribose from interstellar building blocks promoted by amino esters or amino nitriles", *Chemical Communications*, 2017, **53**, 10362–10365.

(332) C. R. Arumainayagam, R. T. Garrod, M. C. Boyer, A. K. Hay, S. T. Bao, J. S. Campbell, J. Wang, C. M. Nowak, M. R. Arumainayagam and P. J. Hodge, "Extraterrestrial prebiotic molecules: photochemistry vs. radiation chemistry of interstellar ices", *Chemical Society Reviews*, 2019, **48**, 2293–2314.

(333) D. J. Ritson and J. D. Sutherland, "Conversion of Biosynthetic Precursors of RNA to Those of DNA by Photoredox Chemistry", *Journal of Molecular Evolution*, 2014, **78**, 245–250.

(334) G. Scalmani and M. J. Frisch, "Continuous surface charge polarizable continuum models of solvation. I. General formalism", *The Journal of Chemical Physics*, 2010, **132**, 114110.

(335) C. Peng and H. Bernhard Schlegel, "Combining Synchronous Transit and Quasi-Newton Methods to Find Transition States", *Israel Journal of Chemistry*, 1993, **33**, 449–454.

(336) K. Fukui, "The path of chemical reactions - the IRC approach", *Accounts of Chemical Research*, 1981, **14**, 363–368.

(337) H. P. Hratchian and H. B. Schlegel, "Accurate reaction paths using a Hessian based predictor-corrector integrator", *The Journal of Chemical Physics*, 2004, **120**, 9918–9924.

(338) M. Ohishi, S.-i. Ishikawa, T. Amano, H. Oka, W. M. Irvine, J. E. Dickens, L. M. Ziurys and A. J. Apponi, "Detection of A New Interstellar Molecular Ion, H_2COH^+ (Protonated Formaldehyde)", *The Astrophysical Journal*, 1996, **471**, L61–L64.

(339) A. Bacmann, E. García-García and A. Faure, "Detection of protonated formaldehyde in the prestellar core L1689B", *Astronomy & Astrophysics*, 2016, **588**, L8.

(340) A. F. Jalbout, "Prebiotic Synthesis of Simple Sugars by an Interstellar Formose Reaction", *Origins of Life and Evolution of Biospheres*, 2008, **38**, 489–497.

(341) L. E. Snyder, D. Buhl, B. Zuckerman and P. Palmer, "Microwave Detection of Interstellar Formaldehyde", *Physical Review Letters*, 1969, **22**, 679–681.

(342) J. L. Neill, M. T. Muckle, D. P. Zaleski, A. L. Steber, B. H. Pate, V. Lattanzi, S. Spezzano, M. C. McCarthy and A. J. Remijan, "Laboratory and tentative interstellar detection of trans-methyl formate using the publicly available Green Bank Telescope Primos survey", en, *The Astrophysical Journal*, 2012, **755**, 153.