Formation of prebiotic Hydrogen Cyanide dimers in an astrophysical environment

PL Els

orcid.org 0000-0003-2141-7705

Dissertation submitted in partial fulfilment of the requirements for the degree *Masters of Science in Space Physics* at the North-West University

Supervisor: Prof DJ van der Walt

Co-supervisor: Prof F Petruccione

Assistant Supervisor: Prof CGCE van Sittert

Graduation May 2019

22934537

Abstract

The past half-century has seen much advancement in the fields of astronomy and astrochemistry, but an emerging and highly interdisciplinary field, known as astrobiology, now seeks to answer one of the oldest and most fundamental questions in science: What is the origin of life? Prebiotic molecules — those that are proposed to be part of the processes leading to the origin of life — and the formation of nucleobases in different astrophysical environments have thus far been the primary area of focus. This study was spurred on by the detection of an HCN dimer ($H_2C_2N_2$) by the Green Bank Telescope, along with possible formation routes leading to adenine ($H_5C_5N_5$), one of the nucleobases. The relatively low temperatures ($\sim$ 10 K) and densities of the relevant environments naturally leads one to suspect the involvement of non-trivial quantum mechanical effects in the chemical processes. The possibility of using an open quantum systems approach to the problem of HCN dimer formation or even of spontaneous dimerisation on the surfaces of interstellar ice grains was investigated. This required gaining an understanding of current methods of investigation regarding surface reactions, seeing if there are any gaps in the theory that can be filled by quantum mechanics, attempting to remedy this if they are, indeed, found. We expected to be able to model the ice-surface and accompanying molecules as a two-level system (or, at the very least, not something incredibly complex) in which the important quantum-mechanical effects are accounted for, and to subsequently model the reaction process. The ice-surface had to first be modelled using current, conventional methods. The complexity and thoroughness involved in the modelling by means of computational quantum chemistry (CQC) was unclear on the outset of the study, which has, admittedly, turned into more of a literature review regarding the problem as described in the title. This project thus attempts to provide the reader with background on the problem, current and previous methods with which prebiotic chemical problems have been investigated, a familiarity with many of the different concepts, and finally how such a problem can be solved given a larger time-investment. To this end we look specifically at the problem of HCN dimerisation both on the surface of interstellar ice-grains, as well as in the gas-phase. A chemical reaction pathway for gas-phase dimerisation of HCN and a model for the ice-grain surface are both developed.

Keywords:

astrobiology, open quantum systems
astrochemistry, density functional theory

Contents

List of Figures

List of Tables

Aims

The aim of this work is to provide a review of the results, methods, and avenues of investigation pertaining to the problem of non-terrestrial prebiotic molecule formation — especially of the dimer of hydrogen cyanide, and the possibility of astrophysical ice-grains acting as a catalyst for its formation. We wish to detail the current state of affairs and provide clarity on how problems such as this, and those of a similar nature, can be investigated, i.e., what Astrobiology is, and what research in the field entails. Furthermore, one of the initial outcomes of this work was to look at the possible incorporation of Open Quantum Systems with the methods one would usually use to solve such chemical problems; specifically, if the method of Open Quantum Systems uniquely adds to, or simplifies, the investigation by means of Density Functional Theory in any way. As such, the underlying mathematics and assumptions of Density Functional Theory has to be thoroughly investigated.

Objectives

1. Investigate the claim that the formation of the HCN dimer via ($HCN + HCN \longrightarrow H_2C_2N_2$), in the gas-phase, is chemically unfavourable given the environmental conditions in which the dimer was detected.

2. Scrutinise the results of Vazart et al. (2015), which would suggest an alternative formation route for the HCN dimer in the gas-phase.

3. Model the HCN dimer formation via ($HCN + HCN \longrightarrow H_2C_2N_2$) on the surface of an astrophysical ice-grain, using Density Functional Theory (DFT), and compare the chemical favourability of this process with the two aforementioned points.

4. Review the quantitative background of DFT in order to then use the previous point as a basis from which to generalise the dimerisation process using Open Quantum Systems, should it be feasible to do so.

Chapter 1

Introduction

Science has long since tried to answer some of the most important and fundamental questions in the universe. Although significant advances have been made in many fields, greatly increasing our collective knowledge, the question with possibly the most philosophical impact has yet to be answered: What are the origins of life? This has been a question long on the minds of some in the fields of Astronomy, Astrochemistry, and, more recently, Astrobiology.

The investigation of life beyond our own poses its own unique problems: 1) No life beyond that known to us on Earth has been found and, without experimental evidence, we have to restrict ourselves to what is most likely, instead of what is empirically verifiable. 2) Although current computational power and numerical methods allow us to simulate complex chemical systems, it is still far from being a trivial process to come up with, and then subsequently test, a new idea involving a realistic chemical system. Thus an attempt at a chemical investigation of the emergence of life is subject to what can be solved computationally in a realistic amount of time.

We choose to first give a general background of problems relating to the modelling of prebiotic molecule formation, as well as how the methods used to model quantum and chemical system have developed over the past century. There are a great many different fields involved in these problems and by giving a quick summary of the involvement of each, as well as related literature, we hope to simplify the task of familiarising the uninitiated reader with the different concepts. At the start of writing this dissertation it was, to the author's knowledge, the first attempt of incorporating computational quantum chemistry and open quantum systems (OQS). This turned out to be wrong, and the related literature and discussions can be found in Section 1.3. We give an overview of our methodology and the paths our research led us down at the end of this chapter. This project was thus written with the purpose of being used as a foundation on which to base future research into the origins of life and modelling prebiotic molecule formations.

1.1 Historical background

Not only did the birth of quantum mechanics in the mid-1920s give insight into the fundamental nature of atoms, but it also provided us with a tool for mathematically solving simple atoms

and molecules (see, e.g., Heitler & London 1927). Within a couple of years, these ideas were refined to the point where Dirac (1929) stated that:

> *"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."*

Pioneering work in tackling the difficulty that Dirac spoke of was done by Hartree (1928) in starting development of a method, called self-consistent field theory, to handle quantum many-body problems *ab initio* — specifically as a method of approximation of the wave function and energy of the system. The method that had emerged by the 1930s, known as the Hartree-Fock method (see, e.g., Slater 1930, Brillouin 1934, Hartree & Hartree 1935), did in no way prove Dirac's statement wrong, as proposed improvements and additional terms to increase the accuracy of the approximations are still commonly found nearly a century later (Sousa et al. 2007, Riley et al. 2007, Korth & Grimme 2009, Cramer & Truhlar 2009, Goerigk & Grimme 2011). Although there was an improvement in the theoretical treatment of chemical systems, such as valence bond theory and molecular orbital theory (Chapter 3 on page 27), it is the development of automatic computing devices in the period after the second world war which is the next milestone in the pre-history of Astrobiology. Increase in the level of complexity of what is being modelled has necessitated a decrease in the desired computational cost, since *ab initio* methods yield mathematical problems that would take too long to solve in any reasonable amount of time.

A great leap forward came in the 1960s (Hohenberg & Kohn 1964, Kohn & Sham 1965) where the ground-work was laid for the development of *density functional theory* (DFT). This provided a method of determining the properties of a molecule based on solving for the electron density of the molecule. More specifically, it solves the energy of the molecule which, in turn, is a function of the electron density (hence functional). The importance of DFT cannot be overstated. Unlike the solution to the wavefunction of the system — which contains all possible information of the system — in *ab initio* methods, which scale as N^4 in terms of computational cost with the size of the system, solving for the electron density scales as N^3, where N is the number of *basis functions*[1] under consideration. DFT calculations are thus faster in reaching the desired accuracy than the *ab initio* methods (Kohn et al. 1996, Burke 2012). Many of the most sophisticated methods that are currently in use make a compromise: They try and incorporate the more useful aspects of *ab initio* methods (mostly from Hartree-Fock methods) while trying to maintain the improvements provided by DFT. These methods are known as hybrid methods and some of the most common of them are discussed in more detail in Section 3.3. The partial inclusion of the Hartree-Fock method, even in the most sophisticated methods, means that it remains invaluable to research based in computational quantum chemistry.

The advent of the computer was not the only necessity: methods for the detection and identification of molecules and their astrophysical environments were needed before any scientific attempt at an answer to where life came from could be made. A total of nearly 200 molecules have been detected in the interstellar medium (ISM) and circumstellar shells thus far (reported

[1] See Section 3.4

in Table 2.1 on page 8), and investigation of the formation of prebiotic molecules is an active field of research in which models can and *must* be tested against observation.

1.2 Early Astrochemistry and Astrobiology

The simplest, in terms of structural and mathematical treatment, and most abundant molecule in the universe is molecular hydrogen H_2. Although its theoretical treatment was done mostly around the 1930s (see, e.g., Heitler & London 1927, Wang 1928, Kemble & Zener 1929, Rosen 1931, James & Coolidge 1933), it wasn't until Gould & Salpeter (1963) that its formation processes were thoroughly investigated. They proposed that association on the surface of interstellar grains were responsible for the high molecular abundance. This process was reinvestigated and significantly expanded upon by Knaap et al. (1966) and later Hollenbach & Salpeter (1970), and theirs is still the approach taken by recent papers, albeit with improvements on, for example, the modelling of the interstellar grain surface (Biham et al. 1998, Roser et al. 2002, Cazaux & Tielens 2002; 2004, Cuppen & Herbst 2005, Navarro-Ruiz et al. 2014). This includes, for example, modelling the chemical surface as amorphous solid water (ASW), ice containing a variety of impurities, as having realistic topologies (such as including ridges and cavities that can act as chemical repositories), or modelling the grain surface as graphite (Fromherz et al. 1993, Parneix & Bréchignac 1998, Farebrother et al. 2000, Hornekær et al. 2003). Laboratory experiments have followed a similar path, testing for the formation of H_2 on a variety of grain analogues. Specifically, carbonaceous surfaces (Pirronello et al. 1997a;b; 1999, Vidali et al. 1998, Zecho et al. 2002, Perry & Price 2003, Güttler et al. 2004), silicates (Pirronello et al. 1997a;b, Vidali et al. 1998), and amorphous ices (Manico et al. 2001, Roser et al. 2001; 2002, Hornekær et al. 2003; 2005, Perets et al. 2005). These results provided temperature ranges in which H_2 formation would efficiently take place on the given surfaces, described in further detail in Section 2.2.2 on page 13.

The building blocks of deoxyribonucleic- and ribonucleic acid (DNA and RNA) are the so-called nucleobases. The nucleobases are cytosine, guanine, adenine, thymine, and uracil, with their chemical configuration and formulae given in Figure 2.5 on page 22. Consideration of the formation of some of these nucleobases, along with the relevant environmental parameters, began as early as the 1960s (Miller & Urey 1959, Ponnamperuma et al. 1963). The relatively simple molecular structure of one of these components, namely adenine, along with the recent identification of one of the *dimers* [2] of hydrogen cyanide by Zaleski et al. (2013) awards it the prime candidacy for further investigation. Indeed, hydrogen cyanide (HCN), its dimer ($H_2C_2N_2$), and adenine ($H_5C_5N_5$) have been the focus of many research projects in recent years (Chakrabarti & Chakrabarti 2000, Smith et al. 2001, Woon 2002, Matthews & Minard 2006, Wang et al. 2007, Roy et al. 2007, Gupta et al. 2011). These included many proposed reactions paths for adenine formation accounting for many different factors such as limitation to neutral-neutral collisions, density and temperature of the environments, as well as the different surface properties mentioned in the previous paragraph. A recent paper by Vazart et al. (2015), motivated by the detection of the dimer, showed that its formation can be accounted for by reactions, subject

[2] HCN is considered a *monomer*, meaning it can react with identical molecules to form more complex structures consisting of several monomers. If the process can carry on *ad infinitum*, the molecule is called a *polymer*, if it only consists of a few monomers, it is referred to as an *oligomer*. A dimer will consist of two monomers, a trimer of three monomers, *etc.*

to characteristic conditions of interstellar clouds, between the cyano radical (CN) and methanimine (CH_2NH). If one does not want to specifically look at chemical reactions on the surface of ice-grains, then the results of Vazart et al. (2015) are worth further investigation.

1.3 Quantum approach

The relevant environments (see Section 2.1 on page 7) for these types of calculations are of such low density and temperature, of the order of $\sim$30 K (Li & Draine 2001, McElroy et al. 2013), that it is reasonable to expect quantum mechanical effects to be involved in the processes. The abundance of molecular and atomic hydrogen would mean that these would also contribute to the surface reactions (Cuppen & Herbst 2005), likely undergoing quantum-mechanical tunnelling between binding sites on the surface due to the combination of low temperature and their small mass (Chang et al. 2005).

The system should thus contain non-trivial quantum properties, something not taken into account by the previously mentioned computational quantum chemistry methods. Furthermore, more recent quantum mechanical methods, specifically the method of *Open quantum systems*, allows for the treatment of complex systems to be simplified to one where the quantum system is separated (not isolated) from its environment. We thus end up having two systems: The quantum mechanical system A, which we wish to solve, and the environmental system B surrounding it. System A thus interacts with the environment to some extent and no matter how small the coupling of the systems are, system A will undergo nonunitary evolution (Auletta et al. 2009). At this point an important assumption is made: the environment B (for example a radiation bath) is considered to be sufficiently large such that system A does not influence its evolution, but it, in turn, will directly influence the evolution of the significantly smaller system A.

Thus if we are, for instance, looking at the dimerisation of a molecule on the surface of an interstellar ice-grain, then it might be possible to fully model the situation using open quantum systems. The system will then contain the two molecules of interest, the grain-surface, and the environment. The mathematical basis for the treatment of open quantum systems, which the above-mentioned system could possibly be modelled as, is well established (see, e.g., Auletta et al. 2009, Le Bellac 2011, Brasil et al. 2013), and the possibility of this *would* have been investigated in this dissertation. The naivety of assuming open quantum systems could partially or wholly replace the established methods of the field was unclear to the author until working through the fundamental derivations and approximations found in DFT. Those that the author find essential to developing an understanding of both the mathematical structure of the field, as well as understanding the potential connection between it (DFT) and open quantum systems are presented in Chapter 3. The usefulness of the idea to include methods from OQS in DFT was, however, validated upon discovery of a series of articles where the formal incorporation of open quantum systems and density functional theory is being attempted (see, e.g., Di Ventra & D'Agosta 2007, D'Agosta & Di Ventra 2008; 2013). D'Agosta & Di Ventra (2013) describes the theory as "*still in its infancy*", but the potential contribution of the theory warrants mention of it in this text.

1.4 Project development and outline

This study was spurred on by the work of Smith et al. (2001) that addresses statements made by Chakrabarti & Chakrabarti (2000). Chakrabarti & Chakrabarti (2000) describe a five-step chemical reaction that leads from HCN (hydrogen cyanide) to $H_5C_5N_5$ (adenine), with each step comprised of the simple addition of an additional HCN molecule. Smith et al. (2001) look at the feasibility of the very first step (HCN + HCN $\longrightarrow$ $H_2C_2N_2$) using arguments based on quantum chemistry as well as calculations regarding chemical reactions in interstellar clouds. It was the detection of the HCN dimer ($H_2C_2N_2$) towards the Sagittarius B2(N) region by Zaleski et al. (2013) that piqued interest in the previous articles.

The aforementioned work was the author's first encounter with computational quantum chemistry (CQC), meaning a familiarity with the subject's nomenclature and quantitative inner-workings had to be acquired before any research relating to it could claim scientific value. To those unfamiliar with the field, commonly used expressions such as "B3LYP/6-311+G(2df,pd)" and "MP4SDTQ/6-311+G(2df,pd)" may appear esoteric. These are, however, important statements that will be found in the vast majority of articles in the fields of Astrochemistry and Astrobiology and give us the necessary information regarding the numerical method of investigation of the chemical reaction, as well as the degree of accuracy to which it was modelled. If one becomes familiarised with the nomenclature and use of the quantum chemical software, the risk of still using the different numerical methods as *black-box* solutions and assuming they are empirical of nature remains. It is for this reason that we have dedicated a considerable portion of this project to look at — and work through — the fundamental assumptions that go into the methods of computational quantum chemistry from both a physical and mathematical point of view. Although a review of some of the procedures of computational quantum chemistry is not strictly encapsulated by the title of this dissertation, it was found that the proposed project requires a detailed interdisciplinary knowledge of CQC before one can even start to consider how open quantum system can be incorporated into the problem. Indeed, it was during the review of density functional theory that articles that pertain directly to the incorporation of open quantum system and DFT were found. Beyond a knowledge of DFT and open quantum systems, a knowledge of *surface science*, a field dealing specifically with chemical and physical phenomena occurring on surfaces and interfaces, is extremely useful. It is thus the author's belief that to be able to solve the problem, as stated in the title of this dissertation, would require more time than can reasonably be allocated to this project.

The amount of different fields that are relevant and the depth of each of these field thus means this dissertation can be regarded as an introduction to *the problem* of investigating prebiotic molecule formation in an astrophysical environment and the place which open quantum systems can take in the research thereof.

Chapter 2

Introduction to prebiotic research

This chapter attempts to provide the necessary background information on the topics mentioned in Chapter 1. The different environments, reaction processes, chemical and physical models, and applications relevant to Astrochemistry are discussed in Sections 2.1 and 2.2. The importance of ice-grain chemistry and its role in astrobiological research, as well as other noteworthy prebiotic-related research questions will be made clear in Sections 2.2.3 and 2.3. The chapter is concluded in Section 2.4

2.1 Astrophysical environments

The giant molecular cloud near the centre of the Milky way, known as Sagittarius B2 (Sgr B2), is one of the most active regions of star formation in the Galaxy. It is, therefore, of utmost importance in the study of non-terrestrial molecular formation. Indeed, a large number of complex molecules have been detected towards Sgr B2 by comparing the spectral line survey of the region with models of the emission of known molecules (see, e.g., Nummelin et al. 1998, Friedel et al. 2004, Belloche et al. 2009). The goal of Astrochemistry and, by extension, Astrobiology, is to understand and explain the chemical processes at work in these astrophysical environments. In the case of the latter, the focus is on those molecules that are involved in the processes that lead to the building blocks of DNA and RNA — the nucleobases (see Figure 2.5). Although the eager researcher might want to propose a complete chemical formation route leading all the way to the nucleobases, we must limit ourselves first to explaining those molecules that have been detected, since theorising too far beyond this point is simply conjecture. It is thus the author's opinion that studying the formation of experimentally-detected prebiotic *precursors* should be given particular attention (see, e.g., Basiuk & Bogillo 2002, Lattelais et al. 2007, Belloche et al. 2008, Gupta et al. 2013, Zaleski et al. 2013, Vazart et al. 2015).

The number of classifications for different types of astrophysical environments is so large (and each being its own respective field of research) that we won't describe each of them in this text — we will, however, discuss those conducive to large-scale chemical evolution, and mention how the environments we deem less relevant might play a role as well. Molecular clouds, specifically, are of interest to us. This is because the modelling of their environments have historically included the presence of dust, and thus the synergy between surface chemistry

2 atoms		3 atoms		4 atoms	5 atoms	6 atoms	7 atoms	8 atoms	9 atoms	10 atoms
H_2	NO^+ ?	$C_3{}^*$	C_2N	c-C_3H	$C_5{}^*$	C_5H	C_6H	CH_3C_3N	CH_3C_4H	CH_3C_5N
AlF	ArH^+	C_2H	Si_2C	l-C_3H	C_4H	l-H_2C_4	CH_2CHCN	$HC(O)OCH_3$	CH_3CH_2CN	$(CH_3)_2CO$
AlCl	TiO	C_2O	SiNC	C_3N	C_4Si	$C_2H_4{}^*$	CH_3C_2H	CH_3COOH	$(CH_3)_2O$	$(CH_2OH)_2$
$C_2{}^{**}$	HCl^+	C_2S	HCP	C_3O	l-C_3H_2	CH_3CN	HC_5N	C_7H	HC_7N	CH_3CH_2CHO
CH	FeO ?	CH_2	CCP	C_3S	c-C_3H_2	CH_3NC	CH_3C_2CHO	C_6H_2	C_8H	CH_3CHCH_2O
CH^+	O_2	HCN	AlOH	$C_2H_2{}^*$	H_2CCN	CH_3OH	CH_3NH_2	CH_2OHCHO	$CH_3C(O)NH_2$	CH_3OCH_2OH
CN	CF^+	HCO	H_2O^+	NH_3	$CH_4{}^*$	CH_3SH	c-C_2H_4O	l-HC_6H^*	C_8H^-	
CO	SiH ?	HCO^+	H_2Cl^+	HCCN	HC_3N	HC_3NH^+	H_2CCHOH	CH_2CHCHO (?)	C_3H_6	**11 atoms**
CO^+	PO	HCS^+	KCN	$HCNH^+$	HC_2NC	HC_2CHO	C_6H^-	CH_2CCHCN	CH_3CH_2SH (?)	HC_9N
CP	AlO	HOC^+	FeCN	HNCO	HCOOH	NH_2CHO	CH_3NCO	H_2NCH_2CN	CH_3NHCHO ?	CH_3C_6H
SiC	OH^+	H_2O	HO_2	HNCS	H_2CNH	C_5N	HC_5O	CH_3CHNH	HC_7O	C_2H_5OCHO
HCl	CN^-	H_2S	TiO_2	$HOCO^+$	H_2C_2O	l-HC_4H^*		CH_3SiH_3		$CH_3OC(O)CH_3$
KCl	SH	HNC	HCS	H_2CO	H_2NCN	l-HC_4N				
NH	N_2	HNO	HSC	H_2CN	HNC_3	c-H_2C_3O				**12 atoms**
NO	NS^+	MgCN	NCO	H_2CS	$SiH_4{}^*$	H_2CCNH (?)				c-$C_6H_6{}^*$
NS		MgNC		H_3O^+	H_2COH^+	C_5N^-				n-C_3H_7CN
NaCl		N_2H^+		c-SiC_3	C_4H^-	HNCHCN				i-C_3H_7CN
OH		N_2O		$CH_3{}^*$	HC(O)CN	SiH_3CN				$C_2H_5OCH_3$?
PN		NaCN		C_3N^-	HNCNH	C_5S (?)				
SO		OCS		PH_3	CH_3O					**12+ atoms**
SO^+		SO_2		HCNO	$NH_4{}^+$					$HC_{11}N$?
SiN		c-SiC_2		HOCN	H_2NCO^+					$C_{60}{}^*$
SiO		HS_2		HSCN	$NCCNH^+$					$C_{60}{}^{+*}$
SiS		NH_2		H_2O_2	CH_3Cl					$C_{70}{}^*$
CS		$H_3{}^{+(*)}$		C_3H^+						c-C_6H_5CN
HF		SiCN		HMgNC						
HD		AlNC		HCCO						
SH^+		$CO_2{}^*$		CNCN						

Table 2.1: Molecules detected in the interstellar medium or circumstellar shells. Table adapted from Müller et al. (2001; 2005). https://www.astro.uni-koeln.de/cdms/molecules. Activated complexes are indicated with a "*". Tentative detections that have a reasonable chance of being correct are marked with "?", and secured detections where line overlap cannot be ruled out are indicated with "(?)".

and the gas-phase is well investigated[1](see, e.g., Watson & Salpeter 1972, Tielens & Hagen 1982, d'Hendecourt et al. 1985, Brown 1990, Hasegawa et al. 1992). We are, however, more concerned with those recent models that make use of expanded chemical networks, more accurate descriptions of ice-surfaces of the dust grains, more physical dynamics of the molecular clouds, and, most importantly, benefit from having their physical parameters based on observations (see, e.g., Du et al. 2012, Maret et al. 2013, Lippok et al. 2013, Awad et al. 2014). Of those discussed in the above references, we will make a distinction between two types of molecular clouds: Stable starless cores (also referred to as quiescent-, or cold-cores) and "hot" molecular cores (regions surrounding young stellar objects), the former of which, due to the fact that temperatures don't reach the point of being able to destroy icy dust-grains, has been the primary focus of astrochemical research.

Throughout this study we will be referring to these hot and cold molecular clouds and sometimes simply to molecular clouds in general. We wish to quickly give some typical values of their physical parameters such that there is a reference point when we talk about "dense" regions or "high" temperatures, in the context of molecular clouds, later on. The model of Kalvāns (2015) uses observational data from known starless core regions to investigate their chemical evolution. The data shows that the densities (specifically hydrogen densities n_H) are in the range of 8.7×10^4–$4.0 \times 10^5\,\mathrm{cm}^{-3}$, with the author calculating temperature values in the range of 7.9–11.1 K. In terms of the time-scale for such an environment, he reports that "*... the model indicates dark core lifetimes of* $<1\,\mathrm{Myr}$.", which, although relatively short in the cosmological sense, is plenty of time for chemical evolution to take place[2]. Two further interesting points from this study are that Kalvāns (2015) finds the formation of *complex organic molecules* (details in Section 2.3.1) requires the temperature to spike up to 20 K, and that, given even the limited number of data-points, there is a inverse correlation between the temperatures and the observed hydrogen densities. For hot molecular clouds we have similar time-scales (Andre et al. 1993), but the key difference lies in that these are regions that have either begun, or have completed, gravitational collapse. As such, they are hotter (several 100–1000 K) and more dense (10^6–$10^8\,\mathrm{cm}^{-3}$) than their cold-core counterpart (see, e.g., Ceccarelli et al. 1998, Osorio et al. 1999, Nomura & Millar 2004, Awad et al. 2014). Authors will often refer to cold molecular clouds as also being dense, and we wish to point at that this is within the context of comparing the density with that of the interstellar medium (ISM), which has a density of the order of $n_H \sim 1\,\mathrm{cm}^{-3}$.

The environmental parameters we have mentioned above greatly limit the formation of astromolecules. Given the relatively low temperatures and densities of molecular clouds (and thus also the ISM), the collision rate of atoms and molecules will be very low. That is not to say that gas-phase chemistry does not take place — the timescales applicable to molecular clouds are sufficiently large as to allow for significant chemical evolution to take place (Capriotti & Kozminski 2001, Freyer et al. 2003). Should, however, a reaction require a three-body collision, it would be safe to assume that the particular reaction will be utterly negligible to the chemical composition of its environment[3]. In the case of some molecules, the argument can

[1]More on this in Section 2.2.3 on page 14.
[2]See Section 2.2.4.
[3]Note we are referring to three-body collisions in the gas-phase and thus not considering interstellar dust to

be made that its formation must proceed through some simple reaction in order to explain its abundance. Such was the case with molecular hydrogen (H_2), described in 2.2.2 on page 13.

2.2 Astrochemistry

Before we start with the discussion of the formation of specific astromolecules, and how this then necessitates the involvement of grain-chemistry, we first give a brief description of the different reaction processes relevant to Astrochemistry. We hope a formal introduction of the terminology provides both clarity in later chapters, and serves as a reference point should the reader not be familiar with some of the terms. Appendix A contains a list notable examples of each of the discussed processes in interstellar environments, taken from the UMIST database (McElroy et al. 2013). Where applicable we've specifically included the formation and destruction of HCN for each of the process types.

2.2.1 Elementary reaction processes

We refer the reader to De Becker (2013) and references therein for a more in-depth discussion, since a thorough discussion of each of these processes isn't justified in this text — we will apply but a few of these reactions and our primary focus will involve grain-chemistry.

Photodissociation

Photodissociation refers to the interaction between a photon and molecule where the photon energy is sufficient to break the chemical bond of the molecule. Thus the photon energy must exceed the molecular *bond dissociation energy*. A notable energy threshold of this process is the energy required to ionise atomic hydrogen ($13.6\,\mathrm{eV}$) since the abundance of hydrogen causes most photons at and above this energy to be absorbed, i.e., photodissociation of molecules may be limited by the abundance of atomic hydrogen which absorb most photons of energy above $13.6\,\mathrm{eV}$. In terms of energy values found in chemistry $13.6\,\mathrm{eV}$ is quite large — typical bond dissocitation energies for simple organic molecules (such as H_2) are $\approx 100\,\mathrm{kcal/mol} = 4.33\,\mathrm{eV}$ (Blanksby & Ellison 2003), however, actual photon energies may need to be slightly higher due to the orientation of the molecule. Thus, although photodissociation may be quite relevant to many environments, it is most relevant to those regions surrounding massive stars, which are subject to strong radiation fields. The reaction process is of the form $AB + \gamma \rightarrow A + B$.

Photoionisation

Photoionisation refers to the direct ionisation of an atom or molecule via interaction with a photon, i.e., a photon of sufficient energy strips away an electron from the atom/molecule leaving it positively charged (referred to as a cation in chemistry). Since this process also requires a radiation field, it is subject to the same limitations as photodissociation, that is to say, the photons required for the process are attenuated by atomic hydrogen and dust. Due to the attenuation by H, one is unlikely to find elements being ionised that have an ionisation energy higher than $13.6\,\mathrm{eV}$. The reaction can be expressed as $Z + \gamma \rightarrow Z^+ + \mathrm{e}^-$.

partake in the reactions here. Grain chemistry is introduced with the formation of H_2.

Radiative association reactions

These are secondary reaction processes that stabilise an excited reaction product through the emission of a photon, governed by the expression $A + B \rightarrow AB^* \rightarrow AB + \gamma$. The occurrence of this process is thus highly dependent on the lifetime of the formed molecule (referred to as an activated complex — an intermediate structure that results at the maximum energy point along a reaction pathway) and the timescale in which the stabilising photon can be emitted. There are thus three possible scenarios: no photo-emission takes place and the activated complex dissociates; a third body, which is able to remove the excess energy from the reaction product, collides with the activated complex stabilising the molecule; the activated complex radiates a photon, resulting in the stabilisation of the product. The second case is highly unlikely in a low-density environment, but is an important concept to keep in mind when we discuss interstellar ice-grains.

Associative detachment reactions

In this type of reaction we have the collision of a negatively charged atom or molecule (referred to as an anion) with a neutral partner: $A^- + B \rightarrow AB + \mathrm{e}^-$. The excess energy of the newly formed product is removed by the emission of an electron, resulting in a stable configuration. We will encounter this reaction type again in our discussion of the formation of H_2. This process doesn't play a large role in Astrochemistry; if one consults Table 2.1, we see that only a handful of anions have been detected in space, and thus the necessary ingredients for this type of reaction to take place are scarce.

Neutral-neutral reactions

As the name would suggest, neutral-neutral processes refer to the interaction between neutral atoms or molecules through the van der Waals force ($\propto 1/r^6$), and is of the form $A + B \rightarrow C + D$. For this reason they are short-range and possess a high activation barrier. Due to the energy requirement of molecule formation, neutral-neutral reactions are only relevant to high-temperature regions.

Ion-molecule reactions

In ion-molecule reactions, the two species interact via the polarisation-induced interaction potential ($\propto 1/r^4$). This large increase in efficiency when compared to neutral-neutral interactions means it is far more relevant to Astrochemistry, and thus the chemical composition of the environment in which this process occurs. The reactions would be of the type $A^+ + B \rightarrow C^+ + D$.

Dissociative electron recombination reactions

In this type of reaction we have the capture of an electron by and ion: $A^+ + \mathrm{e}^- \rightarrow A^* \rightarrow C + D$. The newly formed neutral starts in an excited state, and as a result, the molecule dissociates. This reaction type is one of the primary production pathways via which small neutral molecules are formed in the gas-phase.

Cosmic ray induced reactions

Cosmic rays (CRs) are not a singular species of particles; they are highly energetic charged particles that were accelerated by extreme astrophysical events, and consist primarily of protons,

electrons, and ionised Helium (an alpha particle). We thus have several possibilities and give examples of each:

Interaction of a molecule with a CR can lead to its dissociation ($AB+$ $He^+ \rightarrow A^+ + B+$ He). Given the large energies of cosmic rays, it is possible for the dissociating progenitor to be either the primary CR, or any of the (secondary) ions produced *by* an interaction with a CR. This sort of 'particle shower' will be especially efficient in dense molecular clouds (or rather, its efficiency increases with increasing density). If we consider the limitations of photodissociation, i.e., the attenuation of the process by atomic hydrogen and dust particles, then we see that dissociation of molecules via cosmic-ray-interaction can populate the environment with ions that would otherwise rarely happen. Thus cosmic rays enhance ion-neutral interactions, and by extension interstellar chemistry, in regions opaque to most radiation fields. Cosmic rays can also be responsible for the direct ionisation of neutral species, acting as another source of ions in dense molecular clouds. This process is analogues to photoionisation except the interacting particle is a CR instead of a photon, and can be expressed as $A + CR \rightarrow A^+ + e^-$. Lastly, we can have the collisional excitation of molecular hydrogen by electrons emitted during the interaction of some chemical species with a cosmic ray. We can thus consider this one of the secondary processes mentioned at the start of this paragraph. Subsequently, molecular hydrogen may relax via photoemission, and, considering the opacity of dense molecular clouds to photons, CRs may thus provide a local source of radiation fields — the implications of which will be further discussed in Section 2.2.3. These so-called Cosmic-ray-induced photoreactions can either ionise its partner or lead to its dissociation, being of the form $A + CR \rightarrow A^+ + e^-$ or $AB + CR \rightarrow A + B$, respectively.

Charge transfer reactions

Charge transfer reactions refer to the simple exchange of an electric charge between two species: $A^+ + B \rightarrow A + B^+$. As such, there is no breaking of the chemical bonds of either of the species involved. The efficiency of this process, in the case of the participants being atomic, benefits from having similar ionisation potentials.

Collisional association reactions

These types of reactions are characterised by three-body collisions of the form $A + B + C \rightarrow AB + C$. As mentioned, any three-body collision will be exceedingly rare in typical astrophysical environments, and should only be considered in the case of regions of extreme density. The role of the third participant in the collision is to act as a outlet for the excess energy of the reaction product, stabilising it where it would otherwise dissociate.

Collisional dissociation reactions

The final reaction we wish to introduce is the destruction of a molecule by means of collision (i.e., the reaction is governed by the reversed formula for collisional association). This is more relevant in regions of high temperature, as the colliding molecule/atom will have a higher kinetic energy when compared to low-temperature regions. As one would expect, the most abundant species are also the primary contributors to this type of reactions (i.e., H, H_2, and He). It may happen that the chemical bonds of a molecule are not directly broken by the collision, and it is instead left in an excited state, from which it can dissociate.

2.2.2 H_2 formation

The abundance of molecular hydrogen and, by extension, its formation in interstellar environments was still an open question during the first half of the 20th century. It was, however, already hypothesised that ices might play an important role in the chemical composition of some interstellar environments by the researchers of the time (see, e.g., Eddington 1937, Strömgren 1939). It was understood that the direct radiative association of two atomic hydrogens, through

$$\mathrm{H} + \mathrm{H} \longrightarrow \mathrm{H_2}^* \longrightarrow \mathrm{H_2} + \gamma, \tag{2.1}$$

is a process highly unlikely to occur as the required photostabalisation would have to proceed through a forbidden transition of the molecule (Gould & Salpeter 1963). Alternatively we could have the interaction of hydrogen with its anion (also known as a hydride), expressed as

$$\begin{aligned} \mathrm{H} + \mathrm{e}^- &\longrightarrow \mathrm{H}^- + \gamma \\ \mathrm{H} + \mathrm{H}^- &\longrightarrow \mathrm{H_2} + \mathrm{e}^-. \end{aligned} \tag{2.2}$$

This process circumvents the need for the forbidden transition since the electron may act as a third-body and rid the molecule of its excess energy (Watanabe & Kouchi 2008). The probability of this reaction, unfortunately, suffers from the fact it requires the collision of the hydride and hydrogen in the low-density environment of space. The combination of the low overall density of these regions and the low relative abundance of hydrogen anions implies another candidate is required to account for the prevalence of molecular hydrogen.

Following the failure of the gas-phase processes in accounting for sufficient formation of H_2, it was found (with relatively few and very reasonable assumptions) that interstellar grains can increase the rate of production to the levels that we observe/require (Hollenbach & Salpeter 1970). This is because, as with the electron in Equation 2.2, the surface acts as a third body where the reaction necessary to stabilise the newly formed molecule can deposit energy. Thus if we use the notation of $\mathrm{H}^{(S)}$ to indicate a hydrogen atom bound to the surface of the dust-grain, and $\gamma^{(S)}$ to indicate energy deposited into the surface, then Hollenbach & Salpeter (1970) solved the molecular hydrogen problem by describing its formation as:

$$\begin{aligned} \mathrm{H} + (\mathrm{S}) &\longrightarrow \mathrm{H}^{(\mathrm{S})} \\ \mathrm{H}^{(\mathrm{S})} + \mathrm{H}^{(\mathrm{S})} &\longrightarrow \mathrm{H_2} + \gamma^{(\mathrm{S})} \end{aligned}$$

with $\gamma^{(S)} \approx 4.5\,\mathrm{eV}$. Thus this process avoids both of the shortcomings of the processes in the previous two equations: it doesn't require the improbable collision of H and H^- as in Equation 2.2, and it solves the need for a forbidden transition as in Equation 2.1 by allowing the energy to be absorbed by the surface. Below we give a systematic overview of the process by which ice-grains are proposed to contribute to not only H_2 formation, but the entire chemical composition of interstellar clouds.

2.2.3 Surface chemistry

The primary benefit provided by the presence of interstellar dust could be argued to not be its ability to act as a catalyst for reactions to take place, but rather the fact that they may act as chemical repositories. Thus, an understanding of the basic steps involved in chemical processes between interstellar dust and the chemical species surrounding it is needed. These steps can be listed as: *accretion* of the surrounding gas unto the dust-grain; *migration* of chemical species over the surface; *reactions* between surface species; and, finally, *ejection* from the surface back into the surrounding gas. We describe each of these stages next.

Accretion

The first step necessary for the dust grains to partake in the chemical evolution of its environment is for it to come into contact with another chemical species. This means that if the dust grains have the ability to let chemicals accumulate on them, they can partially overcome the hindrance of low collisional rates in their environment.

The questions we are left to answer when a molecule/atom in the gas-phase and the dust grain happen upon each other is, firstly, do they interact and, secondly, how? In the first case the interaction refers to whether or not the atom/molecule accretes onto the surface of the dust grain, and the second question refers to the actual mechanism by which the accreted atom or molecule is bound to the surface. The process of the particles moving from the gas-phase onto the surface is often referred to as *adsorption*, and the particle thereafter referred to as an *adatom* or *admolecule* depending on its structure. Formally adsorption is thus the adhesion of a particle to the surface of a substance, as opposed to absorption, which refers to the particle's migration *into* a substance. Articles on the subject deal with the probability of whether or not the encounter between a gas-phase particle and the dust grain results in the adsorption of the particle by means of the *sticking coefficient* ξ — which, having a numerical value in the range of $[0, 1]$, directly gives the likelihood of adsorption upon collision with the dust grain. From a physical point of view the sticking coefficient will depend on factors such as the structure of the chemical that wants to adhere to the surface, the chemical composition and morphology of the surface, the temperature of both of the participants, etc. Luckily the reasonable (Knowles & Suhl 1977) assumption of ($\xi \to 1$ as $T \to 0$ K) circumvents much of the need of detailed calculations just in order to obtain ξ. Furthermore, one only needs an accurate and realistic value for the sticking coefficient should you wish to work with chemical composition of both the gas-phase and surface, as well as the rate of exchange between them — if you wish to simply investigate the nature of reactions on the surface of the dust grains, then the sticking coefficient is irrelevant as you have already assumed the adsorption of whatever you are investigating to have taken place.

There are two possible mechanisms which keep the adsorbate on the surface: being weakly bound to the surface by the van der Waals force, often referred to as *physisorption* (physical adsorption), where the structure of the adsorbate and surface both remain unchanged[4]; or the actual overlap of the quantum mechanical wavefunctions, resulting in what we know as a

[4]Note this does not mean to imply properties such as the potential energy surface remain unchanged, merely that chemical bonds are not formed.

chemical bond. This literal chemical adsorption is often referred to as *chemisorption*. Of these two mechanisms chemical adsorption is stronger by roughly an order of magnitude (Zangwill 1988, Petucci et al. 2013), but also has a much shorter range (Barlow & Silk 1976, Cazaux & Tielens 2002, Petucci et al. 2013). Thus chemical adsorption requires so-called binding sites which, depending on the composition and morphology of the surface, as well as the change introduced by other adsorbates, may limit the number of possible chemically adsorbed atoms and molecules. Barlow & Silk (1976) points out that although binding sites are limited, the site is not rendered inert upon occupation by an adsorbate and that, in fact, chemically adsorbed atoms remain highly reactive due to the fact that they maintain unused valence electrons (and therefore unused potential bonds).

Residency

Now that we have the situation of an atom or molecule being bound to the surface of the dust grain, we must once again consider two key points: what is required for the accumulation of chemicals on the surface, and how does this take place? The bare minimum requirements for surface reactions to be allowed to occur are for reaction partners to be present on the surface simultaneously. Thus if we define the characteristic time required for potential reactions partners to evaporate from the surface τ_e, and compare that with typical time between collisions of the dust-grain with chemical species in the gas-phase τ_{in}, we must have that they are, at least, of comparable magnitudes: $\tau_e \sim \tau_{in}$. The characteristic time that chemicals spend on the surface is known as the *residency time*, and we will refer to τ_{in} as the *influx time-scale*. If we expand on our simple example, we can add another requirement: that potential reaction partners must meet within their residency time (they must have the opportunity to interact for any reaction process to take place), bringing us to the idea of the *mobility* of the adsorbates.

As the name would suggest, mobility is the ability to migrate along the surface. Adsorbates will attempt to move between binding sites on the surface, which includes both physical- and chemical-adsorption sites. Transference between sites is exactly equivalent to moving between potential wells on a potential energy surface (PES), and occurs by means of thermal diffusion and, for low mass adsorbates (essentially just atomic hydrogen), through quantum tunnelling between the sites. The next requirement we can then make is to say that in order for significant chemical evolution on the surface of the dust-grain, we must have that the *mobility time-scale*, which can be expressed as

$$\tau_m \propto e^{E_a/kT}, \tag{2.3}$$

is comparably *shorter* than the residency time ($\tau_m < \tau_e$). Here E_a is the activation energy needed for transference between binding-sites, and T is the dust-temperature. This requirement ensures that the adsorbate has adequate time to sample an appreciable area of the grain-surface before evaporating.

There has actually been some debate about the extent to which quantum tunnelling contributes to the mobility of adsorbates: A study by Katz et al. (1999) on the formation of molecular hydrogen on interstellar ice analogues — specifically olivine and amorphous carbon grains — seemed to suggest that, despite its low mass, atomic hydrogen would be unable to tunnel between physisorption sites (chemical adsorption site were excluded in the study). We refer

the reader to a paper by Cazaux & Tielens (2004) in which the conclusions of the model of Katz et al. (1999) are re-evaluated by using an expanded model which takes into account both physical- and chemical-adsorption sites. The results of Cazaux & Tielens (2004) indicate that both thermal diffusion and quantum tunnelling are, indeed, essential for an accurate description of molecular hydrogen formation on a grain surface, and that H_2 formation can persist even to high temperatures (several hundred K).

The essentials of the residency step, then, is as follows: Adsorbates can move between chemical- and physical-adsorption sites by means of thermal diffusion and quantum tunnelling according to their mass and the surface temperature. Adsorbates must then meet a potential partner — via their random walk — within their residency time in order for a chemical reaction to take place.

Surface chemical reactions

By saying an adsorbate meets a reaction partner, we are referring to them as coming into close enough proximity to be able to form and/or break chemical bonds. Intentionally or not, many articles which refer to reactions on a surface are phrased in such a way such as to imply that activation barriers are not present on the surface. We emphasise that this only holds true in general in the case of reactions involving radicals (Masel 1996), and examples of barrier-possessing surface-reactions are giving in Table 2.2. In the case of reactions that are in possession of an activation barrier it is often said, for example, that *since the reaction has an activation barrier these reactions will not occur in the gas-phase*. It is the author's opinion that such phrases may mislead the reader, and we wish to emphasise the *reasons* for the occurrence of barriered reactions on a surface: The simplest reason is the longer exposure time reactants have while on the surface. At any point during their residency adsorbates may interact, and they may thus have many more chances to do so when compared to the gas-phase. Secondly, as we have mentioned, the surface may act as a third body for the expenditure of energy. These reactions may well occur in the gas-phase (recall Section 2.2.1), though this would require a three-particle-collision, and this is why it is said these occur on the grain surfaces. Lastly, the barrier *may actually be removed or reduced* on the surface depending on how the constituents of the reaction are changed by the fact that they are bound to the surface (Masel 1996). In this case it may indeed be fair to say that these (barriered) reactions do not occur in the gas-phase, but do occur on the grain surface — though an analysis of the specific reaction would be required before making the claim.

Post-residency

We will refer to the period after the adsorbate has left the surface as the *post-residency* phase. One of the two mechanism that would cause the adsorbate to leave the surface is if it can gain enough energy, due to temperature-dependent vibration, to overcome its binding energy to the surface. The typical time-scale of evaporation can be expressed as (Ertl 2010)

$$\tau_e = \nu_0^{-1} e^{E_b/kT}, \tag{2.4}$$

where ν_0 is the vibrational frequency of the adsorbate, E_b its binding energy to the surface, and T the temperature of the dust-grain. This desorption process is known as *thermal evaporation*, and we will simply refer to it as *evaporation*. The second means by which the adsorbate can

leave the surface is if its reaction with another adsorbate results in the product obtaining a large amount of energy (compared to the binding energy E_b) and is flung from the surface, or the resulting product simply is no longer capable of remaining bound to the surface due to its structure. An example of the latter is molecular hydrogen, where all available electrons spent on the molecular bonds and none are left to bond with the surface, forcing the molecule to either leave the surface or move to a physisorption site. We will refer to desorption of the adsorbate or product via these two mechanisms as *ejection*. It is worth mentioning that the option for the product of the surface-reaction to remain bound to the surface exists (Masel 1996), where they can thus remain to enrich the reactions and increase the complexity of the molecules that can be formed on these ice-grains.

We quantify some of the above concepts by means of its application to atomic and molecular hydrogen. The vibrational frequency of atomic adsorbates is of the order of 1 THz for most systems, and thus ν_0 is typically taken to be in the range of 10^{12}–10^{13} Hz (Ratsch & Scheffler 1998). Assuming values of 50 meV and 1 eV for the binding energy of physical and chemical adsorption respectively, and using Equation 2.4, we see that for T less than 10 K the residency time remains essentially infinite. In the case of chemisorption this holds true even at 100 K — beyond what we typically expect for the environments relevant to us. Physical adsorption, however, already has its residency time fall off to 2.5×10^{-5} s at 30 K, implying that it, as a mechanism, may have a limited contribution to chemical evolution in 'hot' molecular clouds[5].

Table 2.2: Examples of reactions with an activation barrier that occur on the grain surface. Adapted from De Becker (2013).

Reactants		Product(s)	E_b [meV]
H + O_3	→	O_2 + OH	38.8
H + N_2H_2	→	N_2H_2 + H_2	56.0
H + N_2H_4	→	N_2H_3 + H_2	56.0
H + H_2S	→	SH + H_2	74.1
H + CO	→	HCO	86.2
H + C_2H_4	→	C_2H_5	94.8
H + O_2	→	HO_2	103.4
H + C_2H_2	→	C_2H_3	107.7
H + H_2O_2	→	H_2O + OH	120.6

Surface reaction types

We have discussed some of the details as to why and how surface reactions take place, but we have neglected to mention what many in the field would describe as *surface reaction mechanisms*. This is because the author considers these to be closer to reaction configurations than distinct 'mechanisms'; regardless, we will stick to convention and refer to them as mechanisms. The three mechanisms, then, all provide a way for chemical species to meet (and subsequently react) on the grain-surface, thus describing the *type* of migration the adsorbates have undergone. Note that at the end of each of the mechanisms we describe below, the discussions of Sections 2.2.3 and 2.2.3 again apply.

The first mechanism, known as the Eley-Rideal (Eley 1941) mechanism, in fact, describes the situation where one of the reactants completely skips residency on the surface. We thus have the direct collision between a gaseous atom/molecule with an adsorbate, with the scenario depicted in Figure 2.1. The adsorbate (target on the surface) can either be caught in a chemical adsorption site, and thus be stationary, or it can be migrating between sites, so long as there is a direct reaction between it and the gaseous species. This mechanism clearly requires a high surface-coverage of target adsorbates for it to be prevalent, and is thus more suited for

[5]Unless, of course, Equation 2.4 neglects other major contributions.

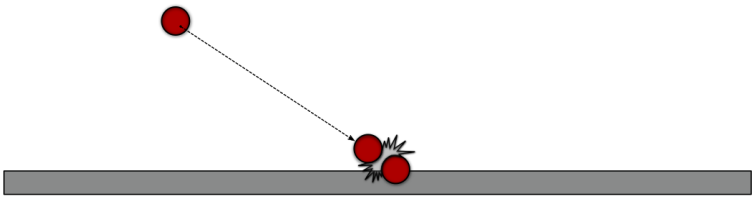


Figure 2.1: Eley-Rideal mechanism

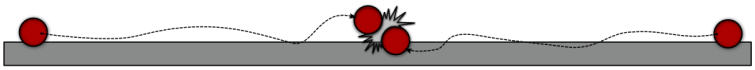


Figure 2.2: Langmuir-Hinshelwood mechanism

low-temperature environments — where adsorbates have a very long residency time — or for environments where you have an extremely large influx of chemical species onto the surface (or a combination of both).

The second mechanism, known as the Langmuir-Hinshelwood (Langmuir 1922, Hinshelwood 1930) mechanism, describes the situation where both adsorbates are migrating along the surface, eventually meeting and reacting. This is the scenario that was implicitly assumed in our discussions of surface reactions throughout this section. The prevalence of this mechanism thus relies on the adsorbates having a long residency time, but still having enough thermal energy to allow for migration between adsorption sites. The Langmuir-Hinshelwood mechanism is depicted in Figure 2.2.

The last mechanism worth mentioning is generally known as *thermal hopping*, though it will sometimes be referred to as the Harris-Kasemo (Harris & Kasemo 1981) mechanism. In this case we have that the gas-phase chemical has sufficient thermal energy for it to not simply be immediately bound to the surface, but *bounce* several times. Each of the bounces causes the particle to lose some of its energy, following which it encounters another adsorbate and reacts. This mechanism can simply be seen as a form of enhanced mobility for a short duration after adsorption, depicted in Figure 2.3.

2.2.4 Chemical reaction rates and chemical networks

With the majority of gas- and surface-mechanisms now in place, we wish to conclude this Section by discussing the relationship between the two. We have already hinted at this when we briefly discussed the formation of molecular hydrogen in Section 2.2.2, but we have yet to quantify the gas-surface synergy. To this end we first introduce the idea of reaction rates, along with the Arrhenius equation, and then discuss how this ties in with the concept of *chemical networks* and subsequently with astrobiological research.

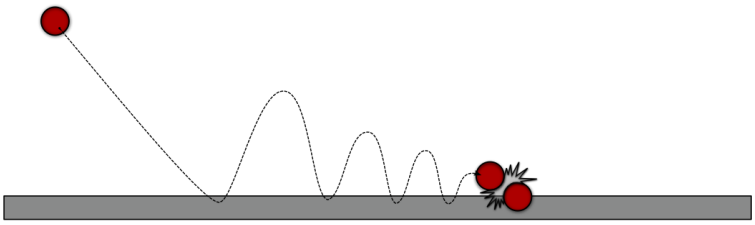


Figure 2.3: Thermal hopping

Let us assume we have chemical species A and B reacting to form a new species C, with densities n_A, n_B, and n_C respectively. If the reaction is of the form $c_1A + c_2B \rightarrow c_3C$, then we can say the *reaction rate* r has the form

$$r = k(T)\,(n_A)^i\,(n_B)^l\,, \tag{2.5}$$

where the values of i and l depend on the reaction mechanism and give the *reaction orders* for species A and B respectively, and k is the temperature-dependent *reaction rate coefficient*. This provides us with the literal rate at which this reaction will take place. Equation 2.5 can be generalised as

$$r = k(T)\,(n_A)^x\,(n_B)^y\,(n_C)^z\,\ldots\,, \tag{2.6}$$

and where the total reaction order will then be given by $x + y + z + \ldots$, determining the units of r. If we have that $c_1 = c_2 = c_3 = 1$, then we can also say that

$$r = -\frac{dn_A}{dt} = -\frac{dn_B}{dt} = \frac{dn_C}{dt},$$

showing the rate of destruction for species A and B, and the rate of formation for species C. In the case of reactions that have some activation barrier, such as those described in Section 2.2.3, the reaction rate coefficient $k(T)$ is described by the *Arrhenius equation* (Laidler 1984),

$$k(T) = Ae^{-(E_b/kT)}, \tag{2.7}$$

where A is called the pre-exponential factor (with units s^{-1}), E_b is the activation energy, k is Boltzmann's constant, and T is the temperature. Temperatures and densities of chemical species in astrophysical environments can be experimentally (or rather *observationally*) obtained, and thus lead to known reaction rates and rate coefficients[6]. The chemical evolution of the environment, however, poses the problem of trying to incorporate every type of reaction we have discussed thus far, and how these are all interconnected. Recalling our discussion of accretion in Section 2.2.3, we stated that detailed calculations of the rate of exchange between the gas-phase and the surface are of not much use if one merely wishes to look at surface-reactions alone. Chemical networks *will*, however, want to take this mechanism into account. Realistic calculations of this exchange rate can, however, be both quite difficult and often case-specific (see, e.g., Gould & Salpeter 1963, Schneider & Smith 1968, Hollenbach & Salpeter 1970,

[6]Indeed, the vast majority of articles that deal with chemistry which we have referenced include the experimentally obtained form and values of coefficients of the reactions.

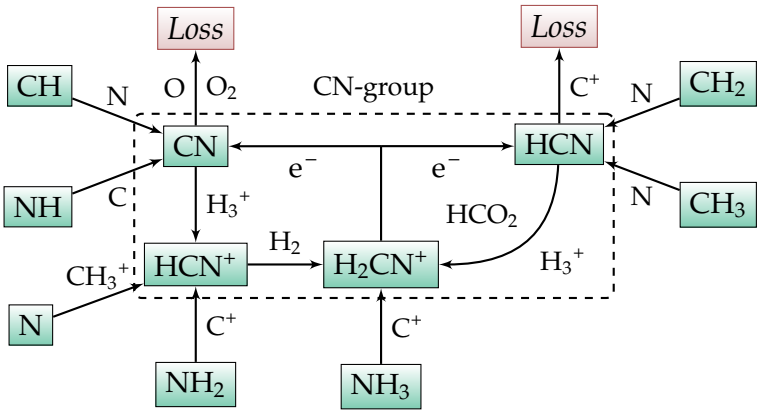


Figure 2.4: Example of a chemical network showing the principle chemical formation- and destruction routes of CN and HCN. Adapted from Prasad & Huntress Jr (1980b).

Leitch-Devlin & Williams 1985, Asnin et al. 2003). This, then, is the idea behind a chemical network — it is, essentially, an immense database of chemical reactions that is fed with data on the chemical composition (specifically the densities or concentrations) and temperature of the environment, and, using the reaction rates and exchange between gas-phase and surface populations, will then describe its chemical evolution. Chemical networks are thus a powerful tool with which to do research. The drawback, however, is that the reaction pathways are *pre-specified*. Thus if we, for instance, want to look at the formation of adenine in some astrophysical environment, and have no formation route for precursors leading up to adenine, then the chemical network will not consider it to be formed. Conversely, if we specify, say, a simple oligomerisation route with successive HCN addition and have values for $k(T)$ that are not physically realistic, then the chemical network will have no problem yielding an abundance of adenine (see, e.g., Chakrabarti & Chakrabarti 2000, Smith et al. 2001).

An example of a piece of such a chemical network is shown in Figure 2.4. In this schematic the network is focused on whether a molecule with a CN bond is formed or destroyed, and doesn't care about other reactants that are formed during a reaction[7]. The reactions are thus read as, for example, CH + N $\longrightarrow$ CN, not showing what happens to the H or the energies involved. The *Loss* channel indicates reactions in which the reactants will no longer contain a CN bond, such as the photodissociation reaction HCN + C^+. Earlier work by the same authors (Prasad & Huntress Jr 1980a) involving the same CN-group chemistry, but including CN^+ and C_2N^+, illustrates the rapid increase in complexity with more molecules and reactions taken into account.

The goal of the chapter up to this point has been to provide the reader with a sufficient amount of background information on what processes and parameters to considered when doing research involving any sort of extraterrestrial chemistry. Each section (and even some of the subsections) have entire textbooks dedicated to exploring their intricacies, but we hope to at least have explained each in such a way that the degree to which they are all connected has become clear. We have dedicated Chapter 3 to the quantitative description of how chemical systems are solved formally — of especial importance for Astrobiology — with the remainder

[7]Note that this is simply for the sake of simplicity of the schematic; the full details of the reactions are taken into account during calculations.

of Chapter 2 dedicated to acquainting the reader with prebiotic research in general, and some of the more interesting avenues of research (in the author's opinion) within the field.

2.3 Astrobiology

The exact meaning of the term 'Astrobiology' will become refined in the coming years. If one subscribes to the notion that life, whatever 'life' may mean, can be formed in space, then Astrobiology means the search for, and description of, the origins of life. If one tends more toward the belief that life itself isn't formed in space, then Astrobiology can be seen as the description of how and where prebiotic molecules are formed. Regardless of which answer is correct, research must strive to describe the data provided to us by experiments and observations. In this section we will attempt to elucidate what we mean by terms such as prebiotic molecules, precursors, what research in Astrobiology entails, and then some ongoing research directions we consider promising.

2.3.1 Complex prebiotic molecules

The building blocks of DNA and RNA are the nucleobases: cytosine ($C_4H_5N_3O$), thymine ($C_5H_6N_2O_2$), uracil ($C_4H_4N_2O_2$), adenine ($C_5H_5N_5$), and guanine ($C_5H_5N_5O$), as shown in Figure 2.5. The chemical structures of the nucleobases, in turn, are based on two nitrogen heterocycles [8] pyrimidine ($C_4H_4N_2$) and purine ($C_5H_4N_4$). The order of significance of discovery will be the same as we have just listed, i.e., DNA or RNA, then the nucleobases, and then prebiotic molecules closely related to the nucleobases. Any molecule that directly participates in the synthesis of the above-mentioned molecules can be considered to be a *prebiotic* molecule. This way of defining the term means it will also be subject to the same type of possible re-evaluation as the term Astrobiology itself. If we were to imagine that in the coming years a complete and unique formation route of DNA is found within which some molecule, which is currently considered to be prebiotic (HCN for example), doesn't participate *at all* then, presumably, it would no longer be considered a prebiotic molecule. This, admittedly pedantic, way of looking at things serves the purpose of emphasising the daunting task Astrobiology has ahead of it. Difficulties aside, observations have provided us with a reasonable picture both of what is happening and of how we are to proceed. Detection of extraterrestrial purine is well-established (Folsome et al. 1971; 1973, Hayatsu et al. 1975, Van der Velden & Schwartz 1977), whereas pyrimidine has yet to be detected, with only estimates of the upper-limits to the column density (in the interstellar medium) provided by Kuan et al. (2003). Of the nucleobases, only the detection of uracil has been confirmed[9] (Martins et al. 2008). Generally the reason for the lack of detection of the more complex prebiotic molecules is accepted to be that these molecules do not form in the gas-phase, and with ongoing research being directly towards their synthesis on the surface of icy grains both in the ISM and in molecular clouds (see, e.g., Ricca et al. 2001, Sandford et al. 2004, Bernstein et al. 2005, Nuevo et al. 2009, Sandford & Nuevo 2014).

As many of the nucleobases have yet to be detected, the search for, and description of, their

[8] Ring structures containing two or more different elements, one of which is nitrogen.
[9] This is to the author's knowledge and as of the time of writing.

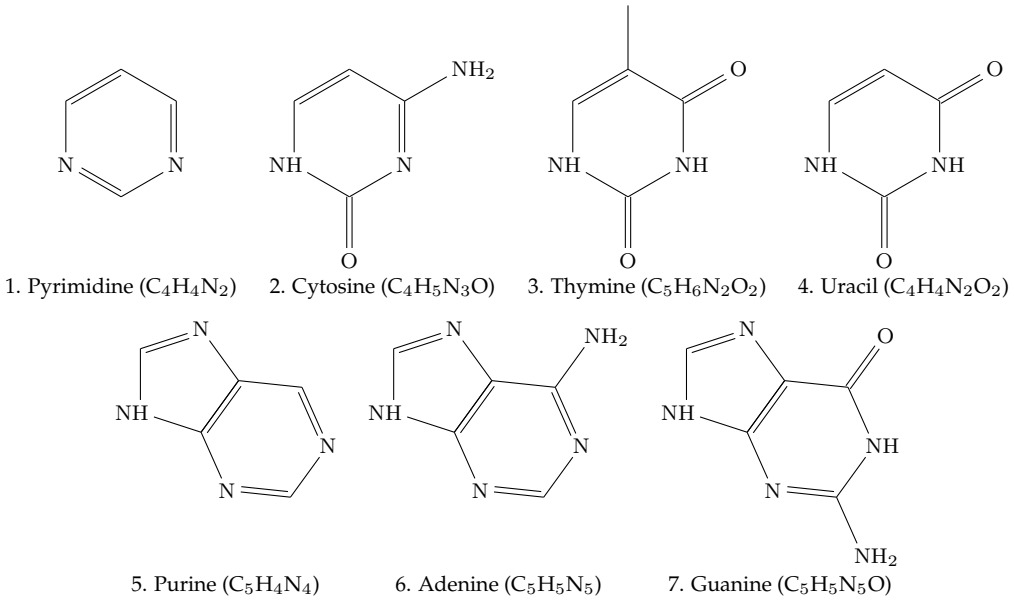


Figure 2.5: The two groups of nucleobases, based on their molecular structure, are shown in the top and bottom row respectively. The top row shows The pyrimidine-derivative nucleobases (2, 3, and 4) and their progenitor (1). The bottom row shows purine (5) and the nucleobases (6 and 7) derived from it.

precursors remains a highly active field of research. By the term 'precursor' we refer to those chemicals that eventually lead to some more complex molecule, i.e., it is likely that the dimer of HCN will somehow be involved in the formation of adenine (which is a *pentamer* of HCN), hence we refer to it as a precursor to adenine. The close chemical relationship between the nucleobases and the two mentioned N-heterocycles means a large amount of research is being directed towards the formation and synergies of these two groups, especially in astrophysical environments and on icy surfaces. The gas-phase formation of pyrimidine rings, and the subsequent formation of purine and adenine by means of photoexcitation in astrophysical environments, have been discussed by Glaser et al. (2007). Similarly, the synthesis of uracil via the photoexcitation of pyrimidine has been investigated by Nuevo et al. (2009), though in their study they assumed a surface-chemistry model — specifically that you have "H_2O:pyrimidine ice mixtures", i.e., pure water interstellar ice-grains with the single contaminant being pyrimidine. Given the importance of these types of studies (the description of *complex* prebiotic molecules), it is worth asking if these directions of research are well founded. Indeed, we have that *polycyclic aromatic hydrocarbons* (PAHs) and *polycyclic aromatic nitrogen heterocycles* (PANHs) are likely present throughout different astrophysical environments (Allamandola et al. 1989, Puget & Léger 1989). Polycyclic aromatic compounds have a structure that is both planar/flat and multi-ringed, with PAHs containing only hydrogen and carbon. As demonstrated by Ricca et al. (2001), the incorporation of nitrogen into the rings of PAHs may proceed by means of the photodesorption of hydrogen atoms, which, in turn, allows the incorporation of an HCN molecule into the ring. The largest energy barrier for this process being that of the desorption of the H atom, which the authors claim merely requires a UV photon, with the rest of the process having sufficiently small energy barriers so that it could readily proceed in an interstellar environment. Furthermore, it seems likely that both polycyclic aromatic hydrocarbons and nitrogen polycyclic aromatic heterocycles tend to be condensed onto the dust grains (Sandford

et al. 2004, Bernstein et al. 2005). Thus, not only is every aspect of these types of research avenues reasonable, but taken together they also start to form a larger picture as to the formation of molecules closely related to the nucleobases.

The prodigious importance of the surface chemistry taking place on dust grains, in regard to the description of the formation of complex prebiotic molecules, is thus clear. One of the hopes for this work was to answer the question of to what degree do ice-grains participate in the formation of prebiotic molecules, which, in the author's opinion, has been answered by the literature and discussions provided in this chapter.

2.3.2 HCN dimerisation

Comparing the chemical complexity of the dimer of hydrogen cyanide, cyanomethanimine $(HCN)_2$, with that of the nucleobases or related molecules from the previous section, one could undoubtedly state that $(HCN)_2$ is 'simpler'. This has not meant, however, that the description of its formation in an astrophysical environment has been easy: The particular problem of HCN dimerisation and its exact structure has received much attention in the past two decades (see, e.g., Heikkilä & Lundell 2000, Savee et al. 2009, Jobst et al. 2008; 2009, Jobst & Terlouw 2010). Interest in HCN and its dimer is born from the fact that adenine is formally the pentamer of HCN. This is, presumably, the reasoning behind the investigation of Chakrabarti & Chakrabarti (2000) into the formation of adenine via the oligomerisation of HCN by successive HCN addition:

$$\mathrm{HCN} + \mathrm{HCN} \longrightarrow \mathrm{CH(NH)CN\,(cyanomethanimine)}, \tag{2.8}$$

$$\mathrm{CH(NH)CN} + \mathrm{HCN} \longrightarrow \mathrm{NH_2CH(CN)_2}, \tag{2.9}$$

$$\mathrm{NH_2CH(CN)_2} + \mathrm{HCN} \longrightarrow \mathrm{NH_2(CN)C{=}C(CN)NH_2}, \tag{2.10}$$

$$\mathrm{NH_2(CN)C{=}C(CN)NH_2} + \mathrm{HCN} \longrightarrow \mathrm{C_5H_5N_5\,(adenine)}. \tag{2.11}$$

The first step of this synthesis route, Equation 2.8, runs into the problem that it is not viable in the gas-phase given typical astrophysical parameters, as shown by Smith et al. (2001). Smith et al. (2001) also note that there may well be reactions leading to adenine or other prebiotic molecules, however, these will involve ion-molecule or radical-molecule interaction processes. Among the investigations of a later study by Yim & Choe (2012) is the possible formation of one of the HCN dimers, iminoacetonitrile (HN=CHCN), by means of the dimerisation of HCN in the gas-phase — again found to not effectively occur under interstellar conditions. Yim & Choe (2012) also looked at the involvement of $HCNH^+$ in HCN dimerisation and formation of HNC (Hydrogen isocyanide, an isomer of HCN). A closely related study by Jung & Choe (2013) then reconsidered the idea of adenine synthesis via sequential addition of HCN, but differs from Equations 2.8–2.11 in that the oligomerisation starts from $HCNH^+$, and that both HCN and HNC are considered for the sequential steps. Although this was found to significantly reduce the required activation energies, the energies were still found to be too high for any of the oligomerisation steps to occur efficiently in interstellar space. The necessity for the description of the HCN dimerisation process was given credence by the discovery of E-cyanomethanimine towards Sgr B2 by Zaleski et al. (2013), who suggested the possible involvement of interstellar

ices given the fact that its discovery means it *can be* formed in astrophysical environments as well as the lack of suitable gas-phase reactions presented up until that point.

The problem of finding a gas-phase formation route for HCN dimers was, yet again, undertaken by Vazart et al. (2015). In this study, however, they abandoned the idea of the dimer forming via HCN oligomerisation and instead expanded on the work of Basiuk & Bogillo (2002), in which they investigate the reaction of methylenimine ($HN{=}CH_2$) with a CN radical, by using more accurate DFT calculations. In so doing Vazart et al. (2015) found that, indeed, the formation of cyanomethanimine isomers (i.e., HCN dimers and the different structural forms thereof) under interstellar conditions is possible:

$$\begin{aligned}
&\mathrm{CN} + \mathrm{CH_2{=}NH} \\
&\longrightarrow \mathrm{Z{-}NC{-}CH{=}NH\,(Z{-}cyanomethanimine)} + \mathrm{H} && (2.12)\\
&\longrightarrow \mathrm{E{-}NC{-}CH{=}NH\,(E{-}cyanomethanimine)} + \mathrm{H} && (2.13)\\
&\longrightarrow \mathrm{CH_2{=}N{-}CN\,(N{-}cyanomethanimine)} + \mathrm{H} && (2.14)\\
&\longrightarrow \mathrm{Z{-}CN{-}CH{=}NH\,(Z{-}C{-}cyanomethanimine)} + \mathrm{H} && (2.15)\\
&\longrightarrow \mathrm{E{-}CN{-}CH{=}NH\,(E{-}C{-}cyanomethanimine)} + \mathrm{H} && (2.16)
\end{aligned}$$

We re-emphasise this result: Vazart et al. (2015) shows that *the formation of HCN dimers can proceed through the gas-phase interaction between $HN{=}CH_2$ (methanimine) and CN, even considering the temperatures and densities found in space.* This is noteworthy for several reasons: if valid, which we will endeavour to test in this work (Sections 4.1 and 5.1), then the paradigm of having cyanomethanimine form through the dimerisation process is brought into question. This has the further implication of other prebiotic molecules — thought to be formed via oligomerisation — perhaps also being formed through processes involving multiple different chemical species, i.e., that the importance of oligomerisation has been overestimated by past research. Lastly, does this imply that, perhaps, the importance of ice-grain and surface chemistry has also been overemphasized with regard to Astrobiology? Given our discussions on H_2 formation, rates equations, and the formation of complex prebiotic molecules, we do not believe this to be the case. Thus, notwithstanding the above discussion, the ubiquity of hydrogen cyanide and hydrogen isocyanide in interstellar space means they will, almost certainly, be adsorbed onto ice-grains also present in these environments (recall the discussions on complex prebiotic molecules). As such, interactions between HCN and HNC on the ice-surfaces are likely to occur, and we are left to ask whether these interactions *can* lead to the formation of $(HCN)_2$ and, if this is indeed the case, then how does the energetic favourability of this compare to the gas-phase reactions of Equations 2.12–2.16?

2.4 Remarks

The fundamentals of Astrobiology have been presented. We have described the minimum required knowledge in order to be able to proceed with research on how the chemical origin of life, at least as we know it, might be formed in outer space. We have seen that the dimerisation of HCN remains an open question. A topic that has been neglected thus far is the quantitative details of how the *vast majority* of the results of articles presented in this chapter are obtained. Thus, in order for us to test or replicate any statements made by other researchers, and to subsequently comment on the validity thereof, we are required to understand these methods at a basic level (at least). This is the purpose of Chapter 3, and is a necessity for us to be able to make any further statement regarding the dimerisation of HCN.

Chapter 3

A Quantum view of Density Functional Theory

The goal of this chapter is to familiarise the reader with the concepts found in computational quantum chemistry — specifically the current methods used to solve chemical problems as well as what is called exchange-correlation. Particular care will be taken to do this in such a way as to keep the mathematical and physical definitions and meaning of the abovementioned clear. If one hopes to use physics to improve methods or understanding in another field, knowing how much physics is involved *already*, and where there are gaps or opportunities, is vital. We thus try to, where possible, introduce concepts and methods in this chapter from fundamental physics and will try to clarify all important points at the end.

We start by introducing a basic approximation needed to solve chemical systems in Section 3.1, which is needed to detail the Hartree-Fock method as was mentioned in Chapter 1. Section 3.2 culminates in the quantitative details of this method — the Hartree-Fock equations (HFE). The derivation of the HFE equations form a basis of knowledge for us to be able to introduce the prevailing method with which complex chemical systems are solved, density functional theory, in Section 3.3, which then allows us to introduce the concept of the exchange-correlation energy. The remainder of the chapter is used to flesh out the necessary details: exchange-correlation energy functionals are discussed in Section 3.3.3, and basis sets in Section 3.4. We conclude Chapter 3 with Section 3.4.3, containing another method used in chemistry which has relevance to this work, though which will perhaps not be encountered as often as those discussed before. Section 3.5 is provided to simply emphasise key points we wish the reader to have gleaned from this chapter.

3.1 Born-Oppenheimer approximation

As previously stated, the objective when trying to determine the properties of an atom or molecule is to solve for its electronic structure (the electron density) since from this we can directly get its energy. One could theoretically solve for all observable information of the system by finding a solution to the Schrödinger equation, but this is an overwhelming challenge. We still, however, have to *start* with the Schrödinger equation before approximations can be

made. The time independent Schrödinger equation of a system can always be written as

$$\hat{H}\Psi = E\Psi, \tag{3.1}$$

where $\hat{H}$ is the Hamiltonian of the system with the caret indicating that it is an operator, Ψ is the wavefunction of the system, and E is the energy. When we are faced with an atomic or molecular system, we will have both electrons and nuclei. For such a system we can explicitly write out the Hamiltonian as

$$\hat{H} = -\frac{\hbar^2}{2}\left(\sum_i \frac{\nabla_i^2}{m_e} + \sum_I \frac{\nabla_I^2}{m_I}\right) + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|r_i - r_j|} - \sum_{i,I}\frac{Z_I e^2}{|r_i - R_I|} + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|R_I - R_J|}, \tag{3.2}$$

where the mass, the position, and the charge, of the electron are given by m_e, r_i, and e, and the corresponding values for the nuclei are given by m_I, R_I, and Z_I. The kinetic energy of the electrons and nuclei are contained in the first two terms. The third and fifth terms (the positive ones) contain the inter-electronic and inter-nucleic repulsion. The remaining term (fourth term) contains the attraction between the electrons and the nuclei. Note that all forces involved here are coulombic forces. Considering, however, that the nuclei are many times more massive than the electrons ($m_p/m_e \approx 1836$), it is a reasonable approximation to make that we have only to consider the motion of the electrons, and regard the nuclei as stationary. This is called the Born-Oppenheimer approximation, proposed in 1927 (Born & Oppenheimer 1927). This allows us to simplify the Hamiltonian by neglecting the kinetic energy of the nuclei, and allowing us to consider the inter-nuclei repulsion term as constant. We can thus solve first for the electronic Hamiltonian

$$\begin{aligned}\hat{H}_e &= -\frac{\hbar^2}{2}\sum_i \frac{\nabla_i^2}{m_e} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|r_i - r_j|} - \sum_{i,I}\frac{Z_I e^2}{|r_i - R_I|} \\ &= \sum_i \hat{h}_i + \frac{e^2}{2}\sum_{i\neq j}\frac{1}{r_{ij}},\end{aligned} \tag{3.3}$$

where we have defined

$$\hat{h}_i \equiv -\frac{\hbar^2}{2}\sum_i \frac{\nabla_i^2}{m_e} - \sum_{i,I}\frac{Z_I e^2}{|r_i - R_I|}, \tag{3.4}$$

and $|r_i - r_j| \equiv r_{ij}$, which will yield the electronic energy E_e when substituted into Equation 3.1:

$$\hat{H}_e\Psi_e = E_e\Psi_e. \tag{3.5}$$

From this equation we can obtain the total internal energy by incorporating the nuclear repulsion term, yielding

$$E_{tot} = E_e + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|R_I - R_J|}. \tag{3.6}$$

A further implication of the Born-Oppenheimer approximation is that, since the electrons will be moving much faster than the nuclei, it would be a reasonable approximation to average over the electron-coordinates. This means we can construct the nuclear Hamiltonian for the motion of the nuclei in, what they see as, a smeared-out cloud of the electrons. The nuclear

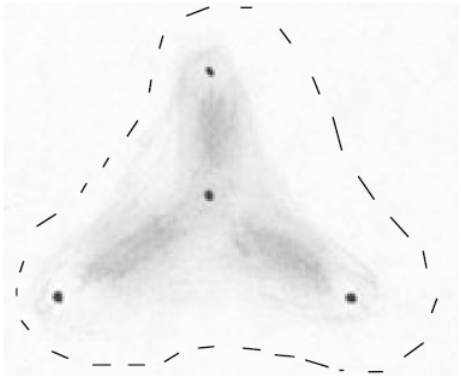


Figure 3.1: Taken from Lewars (2010). The equilibrium points about which the nuclei vibrate define the molecular geometry which can be expressed in either Cartesian nuclear coordinates, or as bond lengths r and angle a.

Hamiltonian H_n then is (Szabo & Ostlund 1996)

$$\begin{aligned}\hat{H}_n &= -\frac{\hbar^2}{2}\sum_I \frac{\nabla_I^2}{m_I} + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|R_I - R_J|} + \left\langle -\frac{\hbar^2}{2}\sum_i \frac{\nabla_i^2}{m_e} - \sum_{i,I}\frac{Z_I e^2}{|r_i - R_I|} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|r_i - r_j|}\right\rangle \\ &= -\frac{\hbar^2}{2}\sum_I \frac{\nabla_I^2}{m_I} + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|R_I - R_J|} + E_e(\vec{R}_I) \\ &= -\frac{\hbar^2}{2}\sum_I \frac{\nabla_I^2}{m_I} + E_{tot}(\vec{R}_I).\end{aligned} \tag{3.7}$$

This has the consequence of the molecule being given a shape by the smeared-out electron field (Lewars 2010), illustrated in Figure 3.1.

3.2 Hartree-Fock method

The simplest *ab initio* calculation of the solution to a chemical system is that of Hartree (1928). It is well-known that any system of three or more interacting objects is analytically unsolvable beyond special cases. Thus any chemical system where you have a nucleus (best-case scenario: a single proton) and more than one electron becomes impossible to solve exactly. Of course, any interesting system will fit this description. The Hartree-method thus starts with an *ansatz* (an initial guess) for the wavefunction of the system of n electrons:

$$\Psi_0 = \psi_0(1)\psi_0(2)\ldots\psi_0(n) = \prod_k \psi_0(k) \tag{3.8}$$

which is a simple product of individual single-electron wavefunctions, known as a Hartree product. Here $\psi_0(k)$ represents the wavefunction of electron k and will depend on the k-th electron's coordinates. The Schrödinger equation can then be solved in which the inter-electronic repulsion term contains one electron, say electron $k = 1$, and the average field of the rest of the electrons found from the remaining electron wavefunctions $\psi_0(2)$, $\psi_0(3)\ldots$, $\psi_0(n)$. This then yields an improved wavefunction for $\psi_1(k = 1)$. This process is then repeated for another electron, say $k = 2$, where the improved electron wavefunction $\psi_1(1)$ is incorporated into

the average field. Every iteration thus improves the global wavefunction Ψ_0, until we have obtained the first set of improved wavefunctions for all n electrons. This constitutes the first cycle, and we are left with

$$\Psi_1 = \psi_1(1)\psi_1(2)\dots\psi_1(n) = \prod_k \psi_1(k). \tag{3.9}$$

If we thus use the general notation in Equation 3.1,

$$\begin{aligned} \Psi_\alpha &= \prod_k \psi_\alpha(k) \\ \therefore \hat{H}\Psi_\alpha &= E_\alpha \Psi_\alpha, \end{aligned} \tag{3.10}$$

then we can say that the method requires iteration until the calculated energy converges: $|E_\alpha - E_{\alpha-1}| \leq \epsilon$, for some suitably small value of ϵ. This implies that the wavefunctions of iteration $\alpha - 1$: $\psi_{\alpha-1}(1)$, $\psi_{\alpha-1}(2)\dots$, $\psi_{\alpha-1}(n)$, and iteration α: $\psi_\alpha(1)$, $\psi_\alpha(2)\dots$, $\psi_\alpha(n)$, have essentially not changed, or (in Hartree's words) are *consistent* with each other. Hence the name self-consistent field (SCF) theory. For the purposes of mathematical completeness it is worth noting that these Hartree-product wavefunctions must satisfy the corresponding single-electron Schrödinger equation. Thus if we consider the Hamiltonian in Equation 3.10, we could write it in the separable form

$$\hat{H} = \sum_i \hat{H}_i = -\frac{\hbar^2}{2}\sum_i \frac{\nabla_i^2}{m_e} - \sum_{i,I} \frac{Z_I e^2}{|r_i - R_I|}. \tag{3.11}$$

Comparison of the above equation with Equation 3.3 shows that we are now using only the one-electron kinetic energy and the nuclear attraction terms.

If one ponders on the approximations made and discussed in Sections 3.1 and 3.2, the reasons for deviations of the predicted energy of the system from the experimental value of the energy becomes clear: When the smeared-out electron field is considered the exact positions of the electrons are lost. This means that the effects you usually expect to occur between electrons (more generally fermions), that is the Pauli exclusion principle due to the spin of the electrons, will no longer take place. It might happen that, due to the average field approximation, two electrons effectively move exactly through one another instead of a strong repulsive interaction taking place, which is non-physical.

The physical definitions, and mathematical treatment of fermions and bosons can be found in any introductory text on quantum mechanics, but for our current purposes it is sufficient to simply note that no more than two electrons can occupy a single atomic or molecular orbital (or state) at a time. This is due to an intrinsic magnetic moment of particles known as its spin, and thus needs to be incorporated into the Hartree method *post hoc*. In the case of the electrons, they have spin $s = \pm(1/2)\hbar$. Where it is usually given in units of $\hbar$, $+1/2$ is usually denoted as 'up' $\uparrow$, and $-1/2$ as 'down' $\downarrow$. The wavefunction should thus not only account for the spatial probability amplitude, but also for the probability of the different spin values. The resulting

wavefunction can thus be denoted by (Auletta et al. 2009)

$$\psi(\vec{x}, s) = \begin{pmatrix} \psi_\uparrow(\vec{x}) \\ \psi_\downarrow(\vec{x}) \end{pmatrix} = \psi_\uparrow(\vec{x}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \psi_\downarrow(\vec{x}) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.12}$$

known as the *spinor*. Before we proceed to incorporate this property into the Hartree method, let us use a simple example to both establish notation and elucidate what spin does to the wavefunction. Consider two electrons a and b, both in the $\uparrow$ spin state. In general the wavefunction will be written as $\psi_k(i, s)$, s indicating the spin state, k indicating some unique wavefunction, and i that the i-th particle is described by this wavefunction. Thus our example has the two wavefunctions $\psi_1(1, \uparrow)$ and $\psi_2(2, \uparrow)$, and we might expect to be able to write

$$\Psi = \psi_1(1, \uparrow)\psi_2(2, \uparrow), \tag{3.13}$$

however, this would be wrong. Electrons are fundamental particles, and care not for our notational and labelling conventions. Any two electrons are indistinguishable from one another, and if we thus decide to swap any of them around, no physical change should occur in the system. Thus if we define an operator $\hat{P}_{ij}$, called the Permutation operator, that is used to mathematically swap electrons i and j, we have that

$$\hat{P}_{12}\Psi = \hat{P}_{12}\psi_1(1, \uparrow)\psi_2(2, \uparrow) = \psi_1(2, \uparrow)\psi_2(1, \uparrow). \tag{3.14}$$

Since the wavefunction cannot be changed by this simple task, it can differ at most by a phase factor $e^{i\beta}$, hence

$$\begin{aligned} \because \hat{P}_{12}^2\Psi = e^{i2\beta}\Psi \quad &\stackrel{!}{=} \hat{I}^2\Psi \\ \implies \psi_1(1, \uparrow)\psi_2(2, \uparrow) \quad &= e^{i\beta}\psi_1(2, \uparrow)\psi_2(1, \uparrow) \\ &= \pm\psi_1(2, \uparrow)\psi_2(1, \uparrow), \end{aligned} \tag{3.15}$$

where the wavefunction is called symmetric if it yields a positive sign under particle exchange and antisymmetric if it becomes negative. The fact that fermions must have antisymmetric wavefunctions and bosons have symmetric wavefunctions is *far* from easy to prove. The proof, known as the *spin-statistics theorem*, must follow from a relativistic quantum theory. We refer the reader to a simplified version of the proof by Duck & Sudarshan (1998) which addresses the original formulation of Streater & Wightman (2000)[1].

The wavefunction Ψ of our example system must thus have the property that $\hat{P}_{12}\Psi = -\Psi$, which Equation 3.13 clearly doesn't fulfil. The solution requires only a slight modification, namely

$$\Psi = \frac{1}{\sqrt{2}} \left[\psi_1(1, \uparrow)\psi_2(2, \uparrow) - \psi_1(2, \uparrow)\psi_2(1, \uparrow)\right]. \tag{3.16}$$

Still sticking with the two-electron example — we can rewrite this in matrix notation as a determinant, so that

$$\Psi = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(1, \uparrow) & \psi_2(1, \uparrow) \\ \psi_1(2, \uparrow) & \psi_2(2, \uparrow) \end{vmatrix}. \tag{3.17}$$

[1] This is the republication of the original book published in 1964.

Which, if we define the product of the spatial and spin eigenfunctions $\psi_1(1,\uparrow) = \phi_1(1)s(1)$ as $\chi_s(n)$, can be expanded to n electrons as (see, e.g., Cramer 2013, Koch & Holthausen 2015)

$$\Psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \cdots & \chi_n(1) \\ \chi_1(2) & \chi_2(2) & \cdots & \chi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(n) & \chi_2(n) & \cdots & \chi_n(n) \end{vmatrix} . \tag{3.18}$$

Equations 3.17 and 3.18 are known as *Slater determinants*, first used by Slater (1930) and equations to keep in mind when we discuss *density functional theory* (DFT) in Section 3.3. To emphasise, Equation 3.18 thus provides a way of ensuring the wavefunction will be antisymmetric under particle exchange (both space and spin states). This can be seen by attempting to exchange two particles in Equation 3.18: particle-coordinate exchange equates to exchange of columns in the determinant formulation, and since column exchange changes the sign of a determinant we are ensured an antisymmetric wavefunction. The normalisation factors of $1/\sqrt{2}$ and $1/\sqrt{n!}$ in Equations 3.16 to 3.18 are due to the electron indistinguishability.

Thus, in the Hartree-Fock scheme the wavefunction of the system under consideration is written as a single Slater determinant. Each of the one-electron functions $\chi_i(\vec{x}_i)$ are called *molecular orbitals* (i.e., the description of the one-electron wavefuntion in a molecule), and the spin eigenfunction part of the orbitals have the important property of being orthonormal, i.e., $\langle\uparrow|\uparrow\rangle = \langle\downarrow|\downarrow\rangle = 1$, and $\langle\uparrow|\downarrow\rangle = \langle\downarrow|\uparrow\rangle = 0$. Furthermore, we have that $\langle\chi_i(\vec{x}_i)|\chi_j(\vec{x}_j)\rangle = \delta_{ij}$, where the Kronecker delta δ_{ij} is defined as

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \tag{3.19}$$

This implies that $d|\chi_i(\vec{x}_i)|^2$ yields the probability of finding the electron i with its specified spin state within a volume element described by $d\vec{x}_i$. It now remains to find the energy of the system. Given the definition of the wavefunction in Equation 3.18, we can then use the usual expression for the energy of a quantum mechanical system:

$$E_e = \langle\Psi|\hat{H}_e|\Psi\rangle \,, \tag{3.20}$$

from which we want to obtain a minimised energy by variation of the spacial parameters of the electrons. We can thus employ the well-known *variational method*. The correct or physically accurate molecular orbits are then those which minimise the energy in Equation 3.20. Note that we cannot improve our approximation beyond the inherent error in the electronic energy E_e caused by the single Slater assumption, but only find the best Slater determinant. This is a limitation of the method itself, not the mathematical treatment of the minimisation process, and is the reason why more and more sophisticated methods have been developed over the years.

3.2.1 Variational method

In general, for any normalisable wavefunction Ψ, we have that

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \tag{3.21}$$

We wish to reduce this equation to one over the individual single-electron orbitals $\chi_i(\vec{x}_i)$ for all n electrons. The full algebraic derivation of an approximated value for the energy is too lengthy to justify its replication in this text, but detailed calculations can be found in Lowe & Peterson (2011), or Cook (2012). A rough outline of the procedure is given below, which will be presented at different levels of complexity. The first and simplest case of the variational method will be used to present its core results, which, once understood, trivialises the proof of the second Hohenberg-Kohn (HK) theorem (discussed on page 38). The remainder of this section is used to show how the variational method is used to arrive at the Hartree-Fock equations, and how it can be further expanded to obtain the Hartree-Fock-Roothaan equations.

As usual, we can write down the expansion of $|\Psi\rangle$ as a linear combination of it's eigenstates:

$$|\Psi\rangle = \sum_n c_n |\psi_n\rangle , \tag{3.22}$$

with the c_n being the expansion coefficients and the eigenstates having the property of being orthonormal: $\langle \psi_n | \psi_m \rangle = \delta_{nm}$. Substitution of the expansion into Equation 3.21 then immediately yields

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{nm} \langle c_n \psi_n | \hat{H} | c_m \psi_m \rangle}{\sum_{nm} \langle c_n \psi_n | c_m \psi_m \rangle} = \frac{\sum_{nm} c_n^* c_m E_m \langle \psi_n | \psi_m \rangle}{\sum_{nm} c_n^* c_m \langle \psi_n | \psi_m \rangle} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2}. \tag{3.23}$$

There are two important things to note here: the expansion via Equation 3.22 assumed the eigenstates $|\psi_n\rangle$ of $\hat{H}$ form a *complete basis set*, an assumption we will do away with later in order to obtain an approximation of the wavefunctions of our system. Secondly, we can label the energies E_n any which way we want. Thus if we choose to write the lowest possible energy as E_0, the second lowest as E_1, and so forth, then we can obtain an artificial — yet useful — inequality:

$$E = \sum_n \frac{|c_n|^2 E_n}{|c_n|^2} \leq \sum_n \frac{|c_n|^2 E_0}{|c_n|^2} = E_0. \tag{3.24}$$

This inequality will only be satisfied in the case where the approximation of the groundstate wavefunction is equal to the true groundstate wavefunction, implying that *deviation from the true groundstate will yield energies greater than the groundstate energy.*

Let's now assume that, again, we want to approximate the wavefunction $|\Psi\rangle$. We can do

away with the complete basis set[2] assumption by truncating the infinite series at some N[3]:

$$|\Psi\rangle = \sum_{i}^{N} c_i \chi_i. \tag{3.25}$$

Now, looking at Equations 3.21 and 3.25, we can see that one approach would be to minimise the energy with respect to the expansion coefficients, i.e., taking $\partial E/\partial c^* = 0$. However, since we are working with an expansion series that is truncated, we can further generalise by including variations of the actual orbitals. Our conditions are then that for

$$\chi_i \to \chi_i + \delta\chi$$

we require that E remains constant for any small variation of the molecular orbitals

$$E(\chi_i) - E(\chi_i + \delta\chi) = 0, \tag{3.26}$$

and that any variation preserves the orthonormality condition as well. If we assume for a moment we were just using the variations of the expansion coefficients, this would mean we have to take

$$\frac{\partial}{\partial c^*} \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = 0,$$

which allows us to use *Lagrangian multipliers* to define some $\mathcal{L}$ as

$$\mathcal{L} \equiv \langle\Psi|\hat{H}|\Psi\rangle - \lambda\left(\langle\Psi|\Psi\rangle - 1\right). \tag{3.27}$$

Solving for $\partial\mathcal{L}/\partial\lambda$ is then equivalent to solving for $\partial\mathcal{L}/\partial c^*$, *with the added condition of then also preserving the orthonormality condition.* Now, if we return to the more general case where we add variations to the orbitals, then $\mathcal{L}$ is defined as[4]

$$\begin{aligned}
\mathcal{L}_i &\equiv \langle\Psi|\hat{H}_e|\Psi\rangle - \sum_{ij}\lambda_{ij}\left(\langle\chi_i|\chi_j\rangle - \delta_{ij}\right) \\
&= \sum_i \langle\Psi|\hat{h}_i|\Psi\rangle + \frac{e^2}{2}\sum_{i\neq j}\langle\Psi|r_{ij}^{-1}|\Psi\rangle - \sum_{ij}\lambda_{ij}\left(\langle\chi_i|\chi_j\rangle - \delta_{ij}\right) \\
&= \sum_i \langle\chi_i|\hat{h}|\chi_i\rangle + \sum_{ij}\left(\frac{e^2}{2}\left[\langle\chi_i\chi_j|r^{-1}|\chi_i\chi_j\rangle - \langle\chi_i\chi_j|r^{-1}|\chi_j\chi_i\rangle\right] - \lambda_{ij}\left[\langle\chi_i|\chi_j\rangle - \delta_{ij}\right]\right)
\end{aligned} \tag{3.28}$$

where we have to solve $\delta\mathcal{L}_i = 0$. To find $\delta\mathcal{L}_i$, we can follow the procedure of Echenique & Alonso (2007), where the variations can be found by rewriting the *functional derivative* of some $\mathcal{F}$ as

$$\lim_{\epsilon\to 0}\frac{\mathcal{F}[\Psi_0 + \epsilon\delta\Psi] - \mathcal{F}[\Psi_0]}{\epsilon} = \int \frac{\delta\mathcal{F}[\Psi_0]}{\delta\Psi}(\vec{x})\delta\Psi(\vec{x})d\vec{x}. \tag{3.29}$$

[2] More on basis sets in Section 3.4

[3] Note also that we revert back to the notation of using χ so as to emphasise that we are again working with orbitals.

[4] Note we return to using $\hat{H}_e$ since we will be using this to derive the Hartree-Fock equations.

When Equation 3.29 is applied to $\mathcal{L}_i$, we (after a little bit of algebra) arrive at

$$\frac{\delta \mathcal{L}_i}{\delta \chi_k^*} = \underbrace{\left[\hat{h} + \sum_j \left(\hat{J}_j[\vec{x}] - \hat{K}_j[\vec{x}]\right)\right]}_{\text{Fock operator } \hat{F}} \chi_k(\vec{x}) - \sum_j \lambda_{kj}\chi_j(\vec{x}), \tag{3.30}$$

were we have define $\hat{J}_j[\vec{x}]$ (Coulomb operator) and $\hat{K}_j[\vec{x}]$ (exchange operator) by the operations they perform on *some* $\phi(\vec{x})$

$$\hat{J}_j[\vec{x}]\phi(\vec{x}) = \left[\int \frac{|\chi_j(\vec{x}\,')|^2}{|\vec{r} - \vec{r}\,'|} d\vec{x}\,'\right]\phi(\vec{x}), \tag{3.31}$$

$$\hat{K}_j[\vec{x}]\phi(\vec{x}) = \left[\int \frac{\chi_j^*(\vec{x}\,')\phi(\vec{x}\,')}{|\vec{r} - \vec{r}\,'|} d\vec{x}\,'\right]\chi_j(\vec{x}). \tag{3.32}$$

We can then write down the eigenvalue equation

$$\hat{F}\chi_i(\vec{x}) = \sum_j \lambda_{ij}\chi_j(\vec{x}), \tag{3.33}$$

where $\hat{F}$ is defined in Equation 3.30, and we are interested in solutions which satisfy $\lambda_{ij} = \delta_{ij}\epsilon_i$. Thus the — very important — Hartree-Fock equations (HFE) can be written as

$$\hat{F}\chi_i(\vec{x}) = \epsilon_i\chi_i(\vec{x}). \tag{3.34}$$

Equation 3.34, and the method used to arrive at it, the variational principle, are the main points of concern of Section 3.2. For completeness we show how the *Hartree-Fock-Roothaan* equations (since they will likely be encountered when reading about computational methods used in chemistry) are obtained from the Hartree-Fock equations: We can now expand the orbitals in terms of some *known basis set* $\tilde{\chi}$, where each orbital is then written as[5]

$$\chi_i = \sum_\mu c_{\mu i}\tilde{\chi}_\mu, \tag{3.35}$$

with $c_{\mu i}$ being expansion coefficients. Substituting the expanded orbitals into the HFE, multiplying from the left by $\tilde{\chi}_\mu^*$ and then integrating over $d\vec{x}$ yields

$$\begin{aligned} \int \tilde{\chi}_\mu^* \hat{F} \sum_\mu c_{\mu i}\tilde{\chi}_\mu d\vec{x} &= \sum_\mu c_{\mu i} \int \tilde{\chi}_\nu^* \hat{F} \tilde{\chi}_\mu d\vec{x} &&= \epsilon_i \sum_\mu c_{\mu i} \int \tilde{\chi}_\nu^* \tilde{\chi}_\mu d\vec{x} \\ &= \sum_\mu c_{\mu i} \langle \chi_\nu | \hat{F} | \chi_\mu \rangle &&= \epsilon_i \sum_\mu c_{\mu i} \langle \chi_\nu | \chi_\mu \rangle \end{aligned} \tag{3.36}$$

By defining $F_{\nu\mu} \equiv \langle \chi_\nu | \hat{F} | \chi_\mu \rangle$ and $S_{\nu\mu} \equiv \langle \chi_\nu | \chi_\mu \rangle$, we obtain the Hartree-Fock-Roothaan equations:

$$\sum_\mu F_{\nu\mu} c_{\mu i} = \epsilon_i \sum_\mu S_{\nu\mu} c_{\mu i}. \tag{3.37}$$

[5]Note that we do not write out the explicit dependence of the orbitals χ on $\vec{x}$.

The Hartree-Fock-Roothaan equations are thus obtained via the basis set expansion of the orbitals, essentially allowing us to write the Hartree-Fock equations in matrix form. Recalling the discussions of shortcomings of the Hartree-Fock method on page 30, we may now have a useful method for iteratively solving for the system's energy, however, we still have the error introduced by the average-field approximation. Below we discuss density functional theory, which, although it has shortcomings of its own, is the gold standard in current research pertaining to modelling of complex chemical systems.

3.3 Density Functional Theory

3.3.1 Hohenberg-Kohn Theorems

Two theorems by Hohenberg & Kohn (1964) underpin the entire field of density functional theory. The theorems are concerned with the ground state of an interacting electron gas in an external potential, and prove that if the ground state is nondegenerate then its wavefunction is a unique function of the particle density. Note that, in the context of an electron gas, the external potential is the attraction to some positive charge distribution. In molecules this would be the charges of the nuclei. The theorems are described below.

Theorem 1. *The external potential, and thus the total energy, is uniquely dependent on the electron density.*

Proof. Consider a system of interacting electrons in an external potential $\hat{U}$. The Hamiltonian of the system can then be written, using $\hat{T}$ for the kinetic energy and $\hat{R}$ for the inter-electronic repulsion, as

$$\hat{H} = \hat{T} + \hat{U} + \hat{R}, \tag{3.38}$$

with the solution to the Scrödinger equation as given by Equation 3.1. The first of the Hohenberg-Kohn theorems now sets out to prove that the a single-electron potential $u(\vec{x}_i)$ in

$$\hat{U} = \sum_i u(\vec{x}_i) \tag{3.39}$$

is uniquely described by the single-electron density:

$$\rho(\vec{x}) = n \int |\Psi(\vec{x}, \vec{x}_2, \ldots, \vec{x}_N)|^2 d\tau_2 \ldots d\tau_n. \tag{3.40}$$

We first assume the opposite, i.e., that there are two single-particle potentials, $u(\vec{x})$ and $u'(\vec{x})$, that result in the same density $\rho(\vec{x})$, which will have the altered Hamiltonian:

$$\hat{H}' = \hat{T} + \hat{U}' + \hat{R}. \tag{3.41}$$

This implies the solution to the Schrödinger equation will also be altered, such that $\hat{H}'\Psi' = E'\Psi'$, but will have a corresponding electron density that is identical to Equation 3.40. Except for the case of $u(\vec{x})$ and $u'(\vec{x})$ differing by a constant, Ψ and Ψ' will also differ. Assuming normalised wavefunctions for Ψ and Ψ', we can use the method of Section 3.2.1 — the variational

method — to write down the two expressions

$$\begin{aligned} W &= \int \Psi'^{*} \hat{H} \Psi' d\tau &&= \langle \Psi' | \hat{H} | \Psi' \rangle > E \\ W' &= \int \Psi^{*} \hat{H}' \Psi d\tau &&= \langle \Psi | \hat{H}' | \Psi \rangle > E'. \end{aligned} \tag{3.42}$$

We can thus write

$$E < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle \,. \tag{3.43}$$

Consider now Equations 3.38 and 3.41, from which we can see that the $\hat{H} - \hat{H}'$ term will yield

$$\begin{aligned} \hat{H} - \hat{H}' &= (\hat{T} + \hat{U} + \hat{R}) - (\hat{T} + \hat{U}' + \hat{R}) \\ &= \hat{U} - \hat{U}', \end{aligned} \tag{3.44}$$

and thus

$$E < E' + \langle \Psi' | \hat{U} - \hat{U}' | \Psi' \rangle \,. \tag{3.45}$$

Similarly, but starting with E', we can find

$$E' < E - \langle \Psi' | \hat{U} - \hat{U}' | \Psi' \rangle \,. \tag{3.46}$$

Summation of Equations 3.45 and 3.46 yield the (clearly erroneous) expression

$$E + E' < E + E'. \tag{3.47}$$

There can thus not exist two single-particle potentials $u(\vec{x})$ and $u'(\vec{x})$ that will result in the same density $\rho(\vec{x})$, i.e., *the ground state electron density uniquely describes the external potential*. □

The claim was made that Hohenberg-Kohn theorems underpin DFT but, considering the above result, let us show this quantitatively: The density $\rho(\vec{x})$ uniquely describes the external potential $\hat{U}$, which determines the form of the Hamiltonian $\hat{H}$, which then describes the ground state wavefunction Ψ[6]

$$(\Psi \circ \hat{H} \circ \hat{U} \circ \rho)(\vec{x}) \tag{3.48}$$

Hence the ground state wavefunction must be uniquely described by the ground state electron density. Furthermore, the ground state total energy E is determined by Ψ and $\hat{H}$, and thus the sum of kinetic $\hat{T}$ and inter-electronic repulsion energies $\hat{G}$ must also be a function of the electron density:

$$(F \circ \rho)(\vec{x}) = \langle \Psi | \hat{T} + \hat{R} | \Psi \rangle = \int \Psi^{*} (\hat{T} + \hat{R}) \Psi d\tau \tag{3.49}$$

The external potential $\hat{U}$ contributes

$$\langle \Psi | \hat{U} | \Psi \rangle = \int \Psi^{*} \hat{U} \Psi d\tau = \int u(\vec{x}) \rho(\vec{x}) d\tau, \tag{3.50}$$

[6]Note that the notation of $(f \circ g)(\vec{x})$, for two general functions f and g, indicates that f is a function of g which in turn is a function of $\vec{x}$.

so that we finally have the full expression, as in Equation 3.21, for the total ground state energy:

$$\begin{aligned} E = \langle\Psi|\hat{H}|\Psi\rangle &= \int \Psi^*(\hat{T}+\hat{R})\Psi d\tau + \int u(\vec{x})\rho(\vec{x})d\tau \\ &= \underbrace{(F\circ\rho)(\vec{x})}_{\text{Universal}} + \underbrace{\int u(\vec{x})\rho(\vec{x})d\tau}_{\text{System dependent}}. \end{aligned} \tag{3.51}$$

Let us now write this in a form more commonly used in the literature. We can separate out the inter-electronic repulsion from $(F\circ\rho)(\vec{x})$, so that

$$(F\circ\rho)(\vec{x}) = (R\circ\rho)(\vec{x}) + \frac{e^2}{2}\iint \frac{\rho(\vec{x}_1)\rho(\vec{x}_2)}{|r_2-r_1|}d\tau_1 d\tau_2, \tag{3.52}$$

and then solve this for $(R\circ\rho)(\vec{x})$, and write $(F\circ\rho)(\vec{x})$ as how it was originally written in Equation 3.49. We end up with

$$\begin{aligned} (R\circ\rho)(\vec{x}) &= \langle\Psi|\hat{T}|\Psi\rangle + \langle\Psi|\hat{U}|\Psi\rangle - \frac{e^2}{2}\iint \frac{\rho(\vec{x}_1)\rho(\vec{x}_2)}{|r_2-r_1|}d\tau_1 d\tau_2 \\ &= \int \Psi^*\hat{T}\Psi d\tau + \int \Psi^*\hat{U}\Psi d\tau - \frac{e^2}{2}\iint \frac{\rho(\vec{x}_1)\rho(\vec{x}_2)}{|r_2-r_1|}d\tau_1 d\tau_2. \end{aligned} \tag{3.53}$$

As Cook (2012) points out, what is required is a form of $(R\circ\rho)(\vec{x})$ such that it yields the values of both the kinetic energy $(T\circ\rho)(\vec{x})$ and the non-Coulombic part of the inter-electronic repulsion energy, called the *exchange-correlation* energy E_{XC}, from the single-electron density $\rho(\vec{x})$. Thus we arrive at

$$(R\circ\rho)(\vec{x}) = (T\circ\rho)(\vec{x}) + (E_{XC}\circ\rho)(\vec{x}). \tag{3.54}$$

The mathematical notation of $(f\circ g)(\vec{x})$, where f and g are general functions, was used here to emphasise that all these values can be described by the electron density. They are called *functionals* of $\rho(\vec{x})$, and another notation is to write it as $f[\rho(\vec{x})]$. Thus the notational form of Equation 3.54 most widely used is

$$R[\rho(\vec{x})] = T[\rho(\vec{x})] + E_{XC}[\rho(\vec{x})]. \tag{3.55}$$

The above notation will be used throughout the rest of the text. The importance of E_{XC} is further discussed in Section 3.3.3.

Theorem 2. *The exact groundstate energy can be obtain via the variation of the electron density such that it minimises the total energy.*

Using Equation 3.51 we can write down the ground state energy as defined by some unique density $\rho^{(1)}(\vec{x})$. If we then have some different density $\rho^{(2)}(\vec{x})$ the variational principle states that it will yield a higher energy:

$$E[\rho^{(1)}(\vec{x})] \equiv E^{(1)} = \langle\Psi^{(1)}|\hat{H}^{(1)}|\Psi^{(1)}\rangle < \langle\Psi^{(2)}|\hat{H}^{(1)}|\Psi^{(2)}\rangle = E^{(2)} \tag{3.56}$$

Thus the total ground state energy can be found by minimising the energy with respect to the electron density $\rho(\vec{x})$. The density at which the energy is a minimum is the ground state density.

The proof of the second theorem is essentially presented in our discussion of the variational method in Section 3.2.1 (specifically Equation 3.23 and the accompanying text).

Summary

Thus, stated simply, the Hohenberg-Kohn theorems say that

1. Using the ground state electron density $\rho_0(\vec{x})$, all properties of a molecule can be found.
2. Any candidate electron density will yield an energy higher than the corresponding energy of the true ground state electron density of the molecule.

3.3.2 Kohn-Sham

We have seen that the Hohenberg-Kohn theorems thus provide us with a way to obtain the Hamiltonian, wavefunction, and thus the energy, but this unfortunately does not ease the difficulty of computation. At some point one is still faced with the daunting task of finding a solution to the Schrödinger equation. This is where a paper by Kohn & Sham (1965) provides us with a way forward: We can isolate the unknown and difficult-to-compute part of the energy of the system by starting with a system where the electrons are assumed to be non-interacting. The energy of this system is then decomposed into different terms, specifically with a term for the kinetic energy of the non-interacting electrons and another term with the deviation of the electronic kinetic energy from the former due to the inter-electronic interaction *as well as* all other non-classical effects. The latter of which will be concatenated into the exchange-correlation energy term E_{XC}[7]. We thus have a term in which an error will, hopefully, only contribute a small part to the overall energy. Furthermore, we can now use an initial guess of the electron density to iteratively improve on our orbitals $\chi_i(\vec{x})$. When we are confident in our choice of orbitals, we can compute the density $\rho(\vec{x})$ and then finally the energy. The separated (Kohn-Shan) energy functional is

$$E[\rho(\vec{x})] = T[\rho(\vec{x})] + U[\rho(\vec{x})] + J[\rho(\vec{x})] + E_{XC}[\rho(\vec{x})], \tag{3.57}$$

with

$$T[\rho(\vec{x})] = -\frac{\hbar^2}{2m_e}\sum_i \int \chi_i(\vec{x})\nabla_i^2\chi_i(\vec{x})d\tau$$

$$U[\rho(\vec{x})] = e^2\sum_{i,I}\int Z_I\frac{|\chi_i(\vec{x})|^2}{|r_i - R_I|}d\tau$$

$$J[\rho(\vec{x})] = \frac{e^2}{2}\sum_{i\neq j}\iint \frac{|\chi_i(\vec{x}_1)|^2|\chi_j(\vec{x}_2)|^2}{|r_2 - r_1|}d\tau_1 d\tau_2$$

and where we have assumed the density can be written in the single-determinant form,

$$\rho(\vec{x}) = \sum_i \chi_i(\vec{x})\chi_i^*(\vec{x}), \tag{3.58}$$

[7]Note that this value of E_{XC} is not the same as the one we spoke of in Theorem 1 of Section 3.3.1. Read carefully how both are defined.

allowing us to write Equation 3.57 using orbitals, instead of the density. There are some mathematical and conceptual intricacies due to the Hohenberg-Kohn theorems that are addressed by this representation of $\rho(\vec{x})$ that will not be discussed here. We will simply note that, in the single-determinant (or *natural orbital*) form, we have orthonormality,

$$\int \chi_i^*(\vec{x})\chi_j(\vec{x})d\tau = 0, \tag{3.59}$$

and that the density integrates to the number of electrons present:

$$\int \rho(\vec{x})d\tau = n. \tag{3.60}$$

We are now tasked with minimising Equation 3.57 with respect to the orbitals $\chi_i(\vec{x})$, which are subject to the orthonormality condition of Equation 3.59. We can, again, make use of the variational principle as in Section 3.2.1 on page 33. Since E_{XC} cannot be written as a function of the orbitals, we are forced to simply track its variation and keep it as an unknown. We thus define

$$V_{XC} \equiv \frac{\delta E_{XC}}{\delta \rho}, \tag{3.61}$$

with the use of V being due to the fact that it acts as a sort of 'exchange potential' even though it contains some kinetic energy — hence not sticking with the usual U representation for potentials. We will not employ the notation of Equation 3.61, but give it for the sake of completeness[8]. The orbitals that minimise the energy are found to obey

$$h_i^{KS}\chi_i = \epsilon_i\chi_i, \tag{3.62}$$

were we have define the *Kohn-Sham (KS)* operator as

$$h_i^{KS} = -\frac{\hbar^2}{2m_e}\nabla_i^2 - e^2\sum_I \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \int \frac{\rho(\vec{r}\,')}{|\vec{r}_i - \vec{r}\,'|}d\vec{r}\,' + \frac{\delta E_{XC}}{\delta\rho}. \tag{3.63}$$

Before moving on to the different ways of going about computing E_{XC}, it is worth reflecting on the differences between the methods thus far discussed in Chapter 3, namely Hartree-Fock (HF) and Density functional theory. HF makes use of the Born-Oppenheimer approximation, and thus leads to equations that, while exactly solvable, lead to an approximate energy for the system. DFT, on the other hand, contains *no* approximations in the derivation of its equations but leads to a term (the exchange-correlation term) that cannot be solved exactly. Thus the approximation of the energy of the system in DFT has to do with how its equations are solved. Considerable research thus goes into trying to find a balance between accuracy and computational cost when determining E_{XC} (see, e.g., Kurth et al. 1999, Riley et al. 2007, Sousa et al. 2007, Zhao & Truhlar 2007).

We further emphasise this point: *If an analytic expression for the exchange-correlation energy E_{XC} can be found, we would have the complete solution to any chemical system.*

[8] Because, as we will see, E_{XC} is important enough that we don't want it hidden away in another term.

3.3.3 Exchange-correlation

We have established that the main goal in computational quantum chemistry is finding a solution for, or approximate value of, the exchange-correlation energy. Below we discuss different ways of achieving this, each with different degrees of success. We first discuss the simplest and quickest of these, in terms of computational cost, and move towards increasing accuracy. At the end of the section we will briefly summarise and bring a few important points to the attention of the reader. Note that, similar to the reactions processes in Section 2.2.1 on page 10, a plethora of detail can be given about the different functionals and the basis sets in the subsequent section; excellent review articles on these subjects are Jones & Gunnarsson (1989), Baerends & Gritsenko (1997), von Barth (2004), Perdew et al. (2005), Cohen et al. (2008) and Blöchl (2011). More compact discussions can be found in Scuseria & Staroverov (2005), and Table 8.7 on page 295 of Cramer (2013).

Local Density Approximation (LDA)

The simplest method of estimating E_{XC} is the local density approximation. This method assumes the *exchange-correlation energy density* of the system n_ϵ^{sys} (E_{XC} per particle) is equal to the exchange-correlation energy density of a homogeneous electron gas n_ϵ^{fg}, as evaluated at some particular point (locally) in the electron gas, i.e., $n_\epsilon^{sys}(\vec{r}_i) = n_\epsilon^{fg}(\vec{r}_i)$. This is a system where the movement of the electrons is across a background positive charge distribution, and is a well-known system in Plasma physics (Chen & Smith 1984) and Quantum mechanics (Gasiorowicz 2007). The approximation finds its use in systems where there is a certain degree of homogeneity, such as an actual electron gas or, more usefully, in metals where the molecules have formed ordered structures. The approximation breaks down, however, when trying to describe the highly-irregular charge distributions of molecules.

Using the LDA, we can then write the exchange-correlation energy as[9]

$$E_{XC}^{LDA}[\rho(\vec{r})] = \int n_\epsilon[\rho(\vec{r})]\rho(\vec{r})d\vec{r}. \tag{3.64}$$

Thus if we have an electron gas with an electron density $\rho(\vec{r})$, we calculate E_{XC} at every point $\vec{r}$ of the electron gas and these values determine n_ϵ. As mentioned, the electron gas system is well known, and we split E_{XC} into two components:

$$E_{XC} = E_X + E_C, \tag{3.65}$$

where we can then use the analytically derived expression for the exchange energy density of an electron gas to obtain E_X:

$$n_{\epsilon,X}^{LDA} = -\frac{3}{4}\left(\frac{3\rho(\vec{r})}{\pi}\right)^{1/3} \tag{3.66}$$

$$E_X^{LDA}[\rho(\vec{r})] = -\frac{\pi}{4}\int\left(\frac{3\rho(\vec{r})}{\pi}\right)^{4/3} d\vec{r}. \tag{3.67}$$

[9]Note that we use the method-name superscript in all of the defining functions of the different methods and for comparison between different functionals, but will drop this notation in most of the text.

Thus we are only left with the need to obtain a value of the correlation energy E_C, which we only know the asymptotic behaviour of in low- and high-density limits (Wang & Perdew 1991, Levy & Perdew 1993) and with stochastic simulations (Ceperley & Alder 1980) providing us with accurate values of E_C in the intermediate range.

Local Spin Density Approximation (LSD)

The local density approximation can be extended to include the spins of the electrons. This is done by separating the density $\rho(\vec{r})$ into two *spin densities* $\rho(\vec{r}) = \rho_\uparrow(\vec{r}) + \rho_\downarrow(\vec{r})$. This allows for descriptions of the system where we can account for some spin-polarisation, i.e., $\rho_\uparrow(\vec{r}) \neq \rho_\downarrow(\vec{r})$. Comparing this with the LDA, we see that if we have a case where there are no unpaired electrons then the LSD reduces to the LDA. With this we can then define

$$\zeta \equiv \frac{\rho_\uparrow(\vec{r}) - \rho_\downarrow(\vec{r})}{\rho(\vec{r})}, \tag{3.68}$$

to yield the amount of polarisation of the system. Thus $\zeta = 0$ would imply there are equal amounts of polarisation (spin-unpolarised case) in the system and $\zeta = 1$ would indicate a system where all electrons are in the same spin state. The two spin densities define each other through $\rho(\vec{r})$, and instead of simply separating them and have n_ϵ depend on both, we can rewrite them both in terms of ζ. We thus have

$$\begin{aligned} \rho_\uparrow(\vec{r}) &= \frac{\rho(\vec{r})}{2}(1+\zeta) \\ \rho_\downarrow(\vec{r}) &= \frac{\rho(\vec{r})}{2}(1-\zeta) \end{aligned} \tag{3.69}$$

so that we can write the equivalent of Equation 3.66 for the local spin-density approximation as

$$n_{\epsilon,X}^{LSD} = -\frac{3}{8}\left(\frac{3\rho(\vec{r})}{\pi}\right)^{1/3}\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}\right] \tag{3.70}$$

The exchange-correlation energy for the LSD case would be

$$E_{XC}^{LSD}[\rho(\vec{r}),\zeta] = \int n_\epsilon[\rho(\vec{r}),\zeta]\rho(\vec{r})d\vec{r}, \tag{3.71}$$

and E_X, similar to Equation 3.67, is

$$E_X^{LSD}[\rho(\vec{r}),\zeta] = -\frac{\pi}{8}\int\left(\frac{3\rho(\vec{r})}{\pi}\right)^{4/3}\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}\right]d\vec{r}. \tag{3.72}$$

We thus consider some point $\vec{r}_i$ of a system and use the associated spin densities $\rho_\uparrow(\vec{r}_i)$ and $\rho_\downarrow(\vec{r}_i)$, along with the above equation, to find a value for E_{XC}^i corresponding to the value that an electron gas of equal density $\rho(\vec{r}_i)$ and spin-polarisation ζ_i would have. Doing so for all points of the system i yields E_{XC}.

We've mentioned that the LDA (and thus LSD) finds use in highly ordered systems, such as metals, but the *accuracy* in these cases is worth mentioning — Gunnarsson & Lundqvist (1976) shows that both the geometry of the system and energy values are found to be within a few percent of experimentally obtained values.

Generalised Gradient Approximation (GGA)

From both a mathematical and physical point of view the limitations of the abovementioned methods (LDA and LSD) come from the fact that we are extracting information only from a single point when determining E^i_{XC}. As with many numerical analysis methods we can greatly improve accuracy, while still only dealing with a local point $\vec{r}_i$, by taking into account the *change in density* — the gradient $\vec{\nabla}\rho(\vec{r})$. This leads to a class of functionals that are referred to as generalised gradient approximations, with their exchange energy being of the form

$$\begin{aligned} E_X^{GGA}[\rho(\vec{r}), \vec{\nabla}\rho(\vec{r})] &= \int n_{\epsilon,X}[\rho_\uparrow(\vec{r}), \rho_\downarrow(\vec{r}), \vec{\nabla}\rho_\uparrow, \vec{\nabla}\rho_\downarrow]\rho(\vec{r})d\vec{r} \\ &= \int n_{\epsilon,X}^{LDA}[\rho(\vec{r})]F_X(\lambda)d\vec{r} \\ &= \int n_{\epsilon,X}^{GGA}[\rho(\vec{r}), \lambda]d\vec{r} \end{aligned} \tag{3.73}$$

The newly introduced function $F_X(\lambda)$ is simply the correction terms to n_ϵ^{LDA} due to a Taylor-like expansion that incorporates the density gradient plus 1. Thus the exchange density in the generalised gradient approximation is

$$n_{\epsilon,X}^{GGA}[\rho(\vec{r})] = n_{\epsilon,X}^{LSD}[\rho(\vec{r})] + \Delta n_{\epsilon,X}\left[\frac{|\vec{\nabla}\rho(\vec{r})|}{\rho^{4/3}(\vec{r})}\right]. \tag{3.74}$$

The biggest change brought about by the GGA, on a physical level, is the more accurate electron behaviour farther away from the nucleus, which is exactly the improvement needed if one wants to more accurately describe chemical bonds. Thus the GGA made it possible to consider bond-breaking, and meant that it is much more useful for chemistry when compared to the LSD and LDA.

Hybrid functional approximations

One can also adopt the approach of using an exact expression for the exchange energy E_X^{HF} and supplementing it with the correlation energy, as treated in the Kohn-Sham approach, in order to approximate the exchange-correlation energy:

$$E_{XC}^{Hybrid} = (1-a)E_{XC}^{DFT} + aE_X^{HF} \tag{3.75}$$

These are also called *adiabatic connection methods*, because of their incorporation of the non-interactive and interactive states. Hybrid functionals are thus weighted sums of DFT exchange-correlation energy and HF energy. As the GGA improved on the LSD, hybrid functionals again substantially improve upon reaction energies and binding energies of GGA functionals. Of these class of functionals, one of the first — and very successful — was presented by Becke (1993a;b), with the exchange-correlation energy given by

$$E_{XC}^{B3PW91} = E_{XC}^{LSD} + a_0(E_X^{HF} - E_X^{LSD}) + a_X\Delta E_X^{B88} + a_C\Delta E_C^{PW91}, \tag{3.76}$$

where a_0, a_X, and a_C are *semiempirical* parameters, ΔE_X^{B88} is a gradient correction to the exchange energy of the LSD (Becke 1988), and ΔE_C^{PW91} is a gradient correction for the correlation

energy (Perdew et al. 1991). As we can see, however, this functional contains parameters used to fit the functional to the experimental values of simple molecules. This means that it cannot be considered a *parameter-free* functional. It show be noted, however, that even some well-known parameter-free functionals (see, e.g., the functional presented by Adamo & Barone 1999, PBE0) still use a mixture of different functionals which, even though the mixing ratios are not experimentally determined, still have to play the parameter-fitting game.

In Equation 3.76 we have already seen that functionals can also be classed into *empirical* or *semiempirical* functional groups. We can see that the higher up we go in terms of the complexity of functionals, the less clear the distinctions between them become — we thus end this section on functionals by giving one last functional: The most prodigious functional, by far, that one can encounter is the B3LYP functional (Stephens et al. 1994):

$$E_{XC}^{B3LYP} = (1-a_0)E_X^{LSD} + a_0 E_X^{HF} + a_X \Delta E_X^{B88} + E_C^{LSD} + a_C \Delta E_C^{PW91} \tag{3.77}$$

with $a_0 = 0.20$, $a_X = 0.72$, and $a_C = 0.81$. No study comparing the accuracies of different functionals would be complete without the inclusion of the B3LYP functional (see, e.g., Bauschlicher 1995, Sousa et al. 2007, Riley et al. 2007, Korth & Grimme 2009, Cramer & Truhlar 2009, Goerigk & Grimme 2011). Note this is not because the B3LYP functional is the most accurate one you can get, it is because of its versatility — it performs adequately well across most[10] types of chemical system, and thus works well as a benchmark.

Every functional, no matter what classification it is given or how complex it is, has the singular goal of being used to find a value for the exchange-correlation energy. If we imagine some complex chemical system we wish to solve completely, then we have seen from Sections 3.1 to 3.3.3 that the best way we to do so, that we know of, is the separation of E_{XC} from the things we can solve analytically, doing so, and then approximating E_{XC}. The *mathematical treatment of how we go about approximating the exchange-correlation energy*, then, is exactly what these functionals we have discussed, represent.

3.4 Basis sets

Basis sets, as used in Equation 3.35, are a well-known concept in quantum mechanics (see, e.g., Auletta et al. 2009, Le Bellac 2011, Gasiorowicz 2007). In computational quantum chemistry they are used as part of the *linear combination of atomic orbitals* (LCAO) approximation. We remind the reader of how orbitals were defined (Equation 3.18 on page 32) — orbitals are wavefunctions of the different constituents of the molecule, and when we refer to a basis set we are talking about the combination of orbitals. This means that by choosing how you approximate the orbitals, you are, by extension, choosing how to approximate the actual wavefunctions of the system. Thus, from the previous section, we see that in order to model a chemical system both an approximation for the wavefunctions and an approximation for how to solve them is needed. Hence the convention of saying "a B3LYP/6-311+G(2df,pd) level of theory was used", which indicates "Functional/Basis-set". Below we discuss some of the most common basis sets used, and remind the reader of the literature given at the end of the first paragraph of Section

[10] There are, of course, always exceptions, but we believe this statement is valid.

3.3.3.

3.4.1 Slater atomic orbitals

The general form of a Slater-type orbital (STO) is given by (Atkins & Friedman 2011)

$$\chi^{\zeta}_{nlm}(r,\theta,\phi) = Nr^{n-1}e^{-\zeta r}Y^m_l(\theta,\phi), \tag{3.78}$$

where n, l, m are quantum numbers, N is a normalisation factor, ζ governs the orbital size, and Y^m_l are the spherical harmonic functions (See any book on Quantum mechanics for more details). Note that STOs do not possess radial nodes, which can be introduced by making a linear combination of STOs. They do, however, give the correct exponential decay behaviour from the nucleus. This gives the exact orbitals when applied to the hydrogen atom. A fairly large disadvantage of STOs is that when you have the product of multiple of these orbitals, you are left to solve a multi-centre electron integral which is computationally expensive. Thus Slater-type orbitals are only really useful for atomic and diatomic systems.

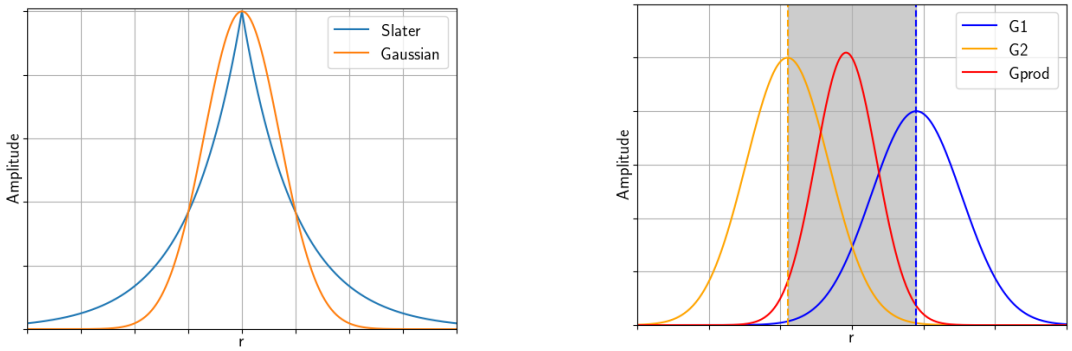
3.4.2 Gaussian orbitals

Gaussian-type orbitals are functions of the form (Atkins & Friedman 2011)

$$\chi^{\zeta}_{ijk}(x,y,z) = Nx^i y^j z^k e^{-\zeta r^2}, \tag{3.79}$$

where $(i,j,k) \in \mathbb{N}$ and their sum determines the type of orbital, ζ is a positive exponent, and N is, again, a normalisation factor. We can construct s-, p-, or d-type Gaussians by taking $i=j=k=0$, $i+j+k=1$, or $i+j+k=2$, respectively. The distinct difference between GTOs and STOs is the r^2 dependence in the exponent, the effects of which are seen in Figure 3.2. This has the disadvantage of causing GTOs to decay too rapidly in comparison to STOs, as well as having a slope equal to zero at the nucleus ($r=0$), meaning they produce unphysical behaviour near the nucleus. These properties mean that GTOs are less accurate than STOs and that more of them are needed to reach the same level of accuracy. The redeeming feature, however, of the Gaussian orbital approach is that any product of Gaussian functions at different centres is equal to another Gaussian function centred between the original ones. Therefore multiple orbital products can be reduced and relatively easily solved numerically — a property that more than compensates for the higher number of orbitals required, and the reason why the use of GTOs, instead of STOs, is common practise.

The simplest type of basis set is one in which each of the orbitals is represented by a single function, referred to as a *minimal basis set*. Thus elements such as Hydrogen and Helium would have their 1s-orbitals each represented by a single function, elements Lithium through Neon one function for each of their 1s-, 2s-, and 2p-ortibals (totalling five basis functions each), and so forth. Results of the wavefunction and energies of a system calculated using minimal basis sets can be improved upon by increasing the number of basis functions used to represent each orbital. By letting each orbital in the minimal basis set be represented by two basis functions, instead of one, you have what is called a *double-ζ* (DZ) *basis set*. Similarly if we choose to represent the orbitals using three basis functions we have a *triple-ζ* (TZ) *basis set*. By increasing



a) Radial forms of Slater- and Gaussian-type (1s) orbitals.

b) Product of two Gaussian functions with indicated region of new centrepoint.

Figure 3.2: Comparison of the properties of Slater- and Guassian-type orbitals. Note the behavioural differences of the orbitals near and far away from the nucleus ($r = 0$ and $r \to \infty$) in the left panel. In the panel on the right we have multiplied the product by a factor of ≈ 5 for ease of comparison with the original Guassians.

the number of basis functions used to represent each orbital you are thus increasing the level of accuracy at the cost of computational requirement. A compromise between the accuracy and computational requirements of the basis set is to use only one basis function for the inner atomic orbitals, and letting the outer valence orbitals be represented by two basis functions. Doing so you obtain what is called a *split-valence* (SV) *basis set*. Table 3.1 shows a comparison between the 3-21G and 6-311++G(d,p) basis sets. It shows the total number of Gaussian functions needed to describe the basis functions, as well as the number of basis sets used to attempt to construct the wavefunctions of the different elements — giving an idea of the different levels and accuracy (and computational cost) between the basis sets, as well as the growth of complexity between basis sets in general.

The nomenclature tries to incorporate some of the details of the basis set, though there are a number of different groups that have each established a notation and, furthermore, there are *so many* different basis sets that the author believes it would be too confusing to try and fully explain each in this text. The examples we have given fall into the category of the *Pople basis sets*, where the general 'formula' is A–$BC\ldots G$. A indicates the number of Gaussian functions used to describe each of the core basis functions, and the number of Gaussians for the valence orbitals are given between '–' and 'G'. Thus 3-21G indicates a split-valence double-ζ basis set, and the 6-311G part a split-valence triple-ζ basis set. We refer the reader to the website of the *Gaussian* software package[11] for a list of different basis sets and their details.

There are a plethora of basis sets to choose from, all of which follow the general trend of either being quick to use but not as accurate, or give a high amount of accuracy but require a large amount of computational resources. If we were to, theoretically, construct a basis set and use an infinite number of Gaussian functions to construct each orbital, we would have the perfect description of the wavefunctions of the system. However, note that given the discussion of the previous section, this would not mean we would get the exact description of the energy of the system, as there still exists an error in the form of the method we choose to solve for the

[11] http://gaussian.com/basissets/

Table 3.1: Comparison of the number of basis functions and the associated number of Gaussian functions used in the 3-21G and 6-311++G(d,p) basis sets.

Basis set	Elements	Basis functions	Gaussians per basis function	Total # of Gaussians
3-21G	H & He	1s	2	3
		1s′	1	
	Li to Ne	1s	3	15
		2s, $2p_{(x,y,z)}$	2	
		s2′, $2p'_{(x,y,z)}$	1	
6-311++G(d,p)	H & He	1s	3	9
		1s′, 1s″ 2s, $2p_{(x,y,z)}$	1	
	Li to Ne	1s	6	36
		2s, $2p_{(x,y,z)}$	3	
		2s′, $2p'_{(x,y,z)}$ 2s″, $2p''_{(x,y,z)}$ $3d_{(xx,yy,zz,xy,xz,yz)}$ 3s, $3p_{(x,y,z)}$	1	

exchange-correlation energy.

3.4.3 Plane waves

The last basis set we wish to comment on is the so-called *plane-wave* basis set, since this is the basis set that is implemented in the *Vienna Ab initio Simulation Package* (VASP, details provided in Hobbs et al. 2000). We allocate a separate section for this so as not to mislead the reader into thinking that one could simply swap out one of the basis sets of the previous section with a plane-wave basis set and be done with it. As an oversimplification we can simply state that the difference between the plane-wave basis set and the GTOs is that instead of describing the wavefunctions of the system using Gaussian functions, we instead use plane waves. Thinking about this implementation, it is clear that these plane waves must then be subjected to *some* potential[12], and subsequently we run into the questions of how to model this potential, how to efficiently solve it, and a whole list of questions that can be seen from the historical development of the method. Good reviews of this — what has now become known as the *projector augmented-wave* (PAW) method — as well its history and all required mathematical derivation can be found in the works of Rostgaard (2009), Blöchl (2011), Rangel et al. (2016), and Xiong & Yanai (2017). We do, however, present the core ideas of the method below.

We remind the reader of the discussion of the Kohn-Sham energy functional, Equation 3.57, and how we then reached the energy-minimising orbitals via Equations 3.62 and 3.63. Here we were introduced to the idea of the exchange-correlation energy, with the subsequent sections on the topic of approximating E_{XC} by means of different functionals and basis sets (the atomic-orbital based methods). The plane-wave method is still within the DFT framework in that it seeks to find the energy-minimising orbitals of the Kohn-Sham discussion. It does, however, take a different approach: we take the electronic wavefunctions to be represented by plane waves, and specify a region around the nuclei (with radius R) within which we make use of an

[12] Determined by the atoms in the molecule.

effective potential. The effective core potential is made to contain the core electrons (i.e., those electrons that would usually fall within this region), and is constructed in such a manner as to have the correct scattering potential beyond a certain distance from the centre. This has both the advantage of us having to calculate the effective core potentials of different atoms only once, and then using the result in subsequent calculations[13], as well as yielding bounds state from the scattering interference (destructive interference of the scattered waves indicate a bound state). Furthermore, this means we need only work with the valence electrons, and we thus have a reduction in the number of basis functions needed in the calculation. *Augmented-wave methods* (AWM) treat the wavefunctions within the central region as partial waves $|\phi_n\rangle$ by means of the plane-wave expansion (a plane wave can be represented by a number of spherically expanding waves), and the wavefunctions of the space between the core regions[14] as envelope functions $|\psi_n\rangle$, and then applying the boundary condition of matching values and derivatives at the edge of the central region R. These methods get the name from the fact that the central core region is referred to as an *augmentation sphere*. Both the *linear augmented-plane-wave* (LAPW) and *linear muffin-tin orbital* (LMTO) methods fall into this category, but will not be discussed here as the PAW method can be seen as the generalisation of these methods and yields each as a special case (Blöchl 1994).

We will not be adopting the original notation of Blöchl (1994) in his presentation of the PAW method, but will mention it when necessary. The basis of the method is to take the physical wavefunctions $|\Psi\rangle$, and transform them into what is called *auxiliary* wavefunctions $|\tilde{\Psi}\rangle$ via the transformation

$$\hat{\mathcal{T}}^{-1}|\Psi\rangle = |\tilde{\Psi}\rangle\,, \tag{3.80}$$

where $\hat{\mathcal{T}}$ is then a transformation operator, and all values with a tilde are the transformed auxiliary variables. The idea behind this being to have the auxiliary wavefunctions be numerically convenient — specifically having them be smooth functions that can be represented in a plane wave expansion. The physical and auxiliary wavefunctions outside of the core region are taken to coincide by taking $\hat{\mathcal{T}}$ to be

$$\hat{\mathcal{T}} = 1 + \sum_R \hat{\mathcal{T}}_R, \tag{3.81}$$

thus restricting the transformation to only the core region. This means that $\hat{\mathcal{T}}_R$ contributes the difference between the physical and auxiliary wavefunctions for every atom:

$$\hat{\mathcal{T}}_R|\tilde{\phi}_n\rangle = |\phi_n\rangle - |\tilde{\phi}_n\rangle\,. \tag{3.82}$$

Note that this implies that outside of the core region R we have that $|\phi_n\rangle = |\tilde{\phi}_n\rangle$, as required. In order to have the transformation operator be applicable to an *arbitrary* auxiliary wavefunction, the physical valence wavefunctions can be expressed as the superposition of the partial (core) wavefunctions with undetermined coefficients as

$$|\psi\rangle = |\phi_n\rangle\, c_n. \tag{3.83}$$

[13] Meaning it can be 'hard-coded' into the method or tabulated and looked-up/imported instead of spending computational resources on its calculation.

[14] Referred to as the interstitial region.

This allows us to define what Blöchl (1994) calls the *projector functions*[15], $\langle \tilde{p}_n|$, through the expansion of the auxiliary wavefunctions into auxiliary partial waves in the core region as

$$|\tilde{\psi}\rangle = |\tilde{\phi}_n\rangle\, c_n = |\tilde{\phi}_n\rangle\, \langle \tilde{p}_n|\tilde{\psi}\rangle\,. \tag{3.84}$$

Thus for every arbitrary auxiliary wavefunction $|\tilde{\psi}\rangle$ that can be expanded into auxiliary partial waves $|\tilde{\phi}_n\rangle$, there exists a projector function $|\tilde{p}_n\rangle$ such that $\langle \tilde{p}_n|\tilde{\phi}_m\rangle = \delta_{nm}$. From Equations 3.82 and 3.84 it then follows that

$$\begin{aligned} \hat{\mathcal{T}}_R\,|\tilde{\psi}\rangle = \hat{\mathcal{T}}_R\,|\tilde{\phi}_n\rangle\,\langle \tilde{p}_n|\tilde{\psi}\rangle &= \Big(|\phi_n\rangle - |\tilde{\phi}_n\rangle\Big)\,\langle \tilde{p}_n|\tilde{\psi}\rangle \\ \implies \hat{\mathcal{T}} &= 1 + \Big(|\phi_n\rangle - |\tilde{\phi}_n\rangle\Big)\,\langle \tilde{p}_n|\,. \end{aligned} \tag{3.85}$$

Which allows us to write the formal transformation between the physical wavefunction $|\Psi\rangle$ and the auxiliary wavefunction $|\tilde{\Psi}\rangle$ as

$$|\Psi\rangle = \hat{\mathcal{T}}\,|\tilde{\Psi}\rangle = \Big\{1 + \Big(|\phi_n\rangle - |\tilde{\phi}_n\rangle\Big)\,\langle \tilde{p}_n|\Big\}\,|\tilde{\Psi}\rangle\,. \tag{3.86}$$

So what does this all mean? This means that we can now write the expectation value of some physical observable A as

$$\langle A\rangle = \langle \Psi|A|\Psi\rangle = \langle \hat{\mathcal{T}}\tilde{\Psi}|A|\hat{\mathcal{T}}\tilde{\Psi}\rangle = \langle \tilde{\Psi}|\hat{\mathcal{T}}^\dagger A\hat{\mathcal{T}}|\tilde{\Psi}\rangle = \langle \tilde{\Psi}|\tilde{A}|\tilde{\Psi}\rangle\,, \tag{3.87}$$

where

$$\tilde{A} = A + |\tilde{p}_n\rangle\,\Big(\langle \phi_n|A|\phi_m\rangle - \langle \tilde{\phi}_n|A|\tilde{\phi}_m\rangle\Big)\,\langle \tilde{p}_m|\,. \tag{3.88}$$

Thus, this enables us to calculate the electron density, and subsequently the total energy of the system, using the numerically convenient auxiliary wavefunctions. We will not show the derivation of either of these here, as this involves substituting both Equations 3.84 and 3.85 into Equation 3.87, which becomes quite lengthy. We do, however, feel that this covers the conceptual workings of the projector augmented-wave method.

The question then naturally arises of how the PAW method compares to the use of functionals and a Slater/Gaussian basis set as mentioned in Sections 3.3 and 3.4. If we consider an electron close to the nucleus of some atom, it will have to possess a (relatively) large kinetic energy, meaning their description requires a fine grid to accurately solve numerically. However, we also know that the core electrons are not the ones of primary concern in chemistry — they are *inert*. Thus, by making use of an augmentation sphere, we avoid spending resources on their description while still having the relevant valence electrons. The projector augmented-wave method has been shown to perform reasonably well for calculations involving lattice structures due to the inherently periodic nature of plane waves (Mostofi et al. 2002, Mortensen et al. 2005).

[15] These projector functions are slightly different from the projection operators we encountered in Section 3.2, but function much the same way.

3.5 Remarks

When considering which basis set to use, specifically when choosing between the Gaussian-based basis sets and plane waves, it is important to note that the plane wave basis set is *fully delocalised* — meaning the electrons which they represent are not associated with a specific atom/molecule or any specific point in space. Gaussian-based basis sets, on the other hand, are localised around the atoms. These properties are inherent to the basis sets due to the mathematical treatment of the orbitals they are trying to describe, but were not specifically mentioned in their discussions. This 'localised' vs. 'delocalised' property thus means, irrespective of the completeness of the basis set you are using, that plane waves are more suited to highly periodic systems (such as solid-state systems), and Gaussian-based basis sets perform better for highly irregular systems (organic systems or complex molecules).

To familiarise the reader with the concepts found in computational quantum chemistry up to the point where the relationship between basis sets, exchange-correlation energy, and functionals can be seen has been the primary goal of this chapter. We have also tried to ensure that when one speaks of "*solving for the exchange-correlation energy*", the reader should now have an idea of where this comes from, as well as what its physical meaning is. The abundance of different methods in existence with which to solve chemical systems should also serve to illustrate the fact that there is no single best method to use for all systems.

The discussions we have presented here should also reveal the detail to which current methods for solving chemical models are fundamentally based on quantum mechanics — a major point that needed to be addressed by this work (recall the discussion on the development of this study in Section 1.4 on page 5).

Chapter 4

Methodology

This chapter details our approach to the validation of the complete gas-phase formation route of HCN dimers, the construction of the surface used as a model for dust-grain surfaces, as well as how we then tested for the dimerisation behaviour of HCN on said surface. As one of the outcomes of this work is to also serve as an introduction to prebiotic research, the author provides additional details of methods and approaches that proved troublesome or that did not work — specifically problems that were present regardless of the software that was used.

The extended gas-phase formation route of $(HCN)_2$, based on and containing the one originally presented by Vazart et al. (2015), is presented in Section 4.1. The steps followed to obtain a semi-realistic dust-grain surface, as well as difficulties encountered along the way, are described in Section 4.2. Finally, we detail the way in which we went about testing the chemical behaviour of adsorbed HCN, as well as how this was then compared to the gas-phase formation route in Section 4.3. Results that will not be used to draw conclusions from but nevertheless provide insight into our methodology will also, where relevant, be presented here.

4.1 Gas-phase dimerisation

The starting species for the gas-phase reactions were chosen as H_2 + 2 (HCN) because of the abundance of both of these molecules, as well as the fact that by using H_2 + 2 (HCN), compared to the H + 2 (HCN) of Vazart et al. (2015), we are able to look at more possibilities for the formation of the HCN dimer. This also has the consequence of being able to predict the formation of more complex prebiotic molecules that might lie further along the reaction path than $(HCN)_2$.

The full reaction pathway (based, in part, on the one proposed by Vazart et al. (2015)) that was considered for gas-phase formation can be seen in Figure 4.1. The molecules presented in blue constitute stable configurations, and those in red transition states. The calculation of each molecule of Figure 4.1 was done using the DMol3 module of the *BIOVIA Materials Studio 2016* software from *Dassault Systèmes Biovia Corp.*[1] (hereafter simply referred to as *Materials Studio*), which allows for the exploration of the potential energy surface of a molecule by means of, for example, searching for rotatable torsions of the molecule. This means we are provided some level of assurance that if we find a minimum-energy formation route for *whichever* molecule in our reaction path, that it is indeed a minimum-energy

[1] Material Studio Modelling Environment, Accelrys Software Inc., San Diego, 2012

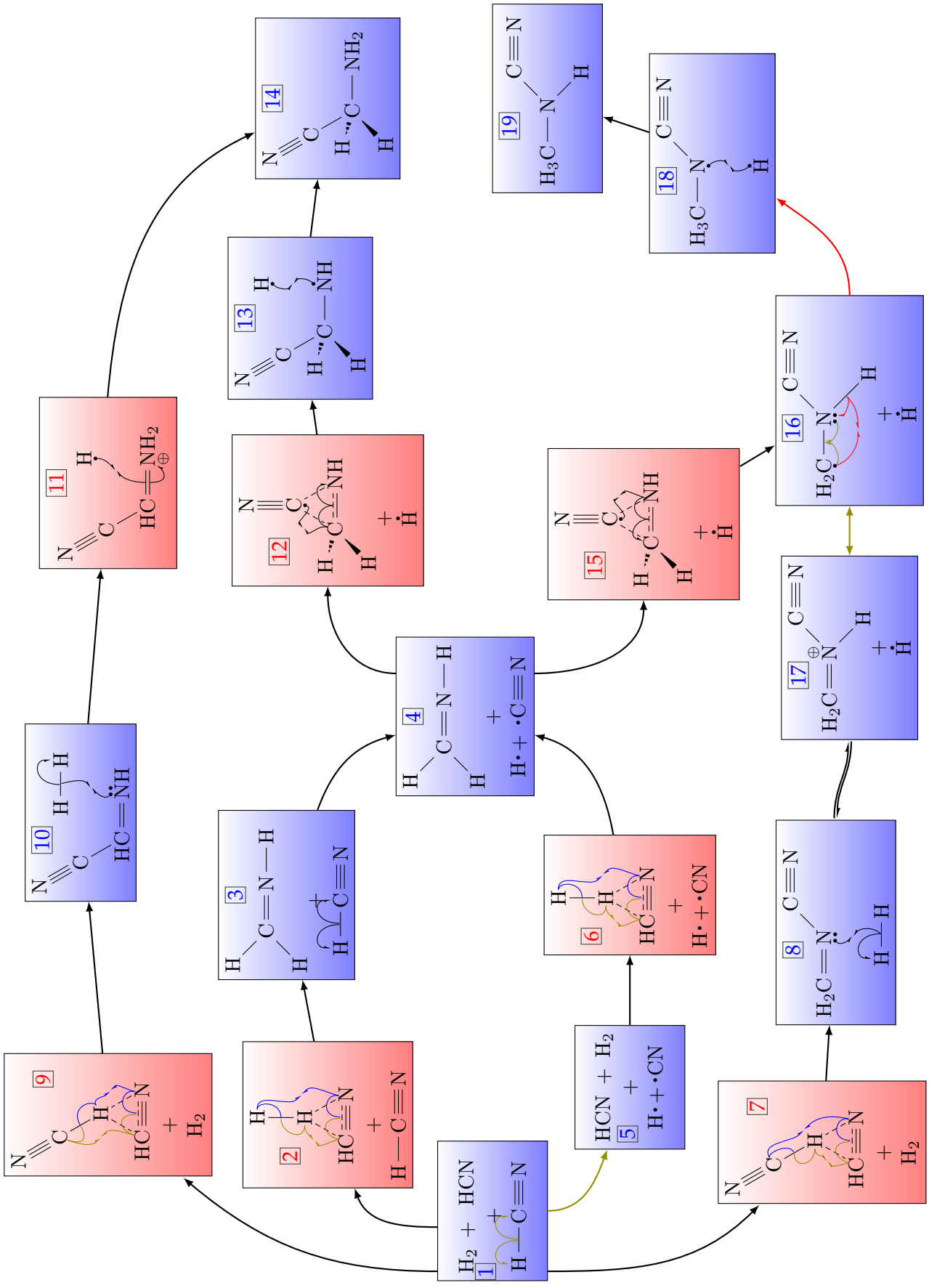


Figure 4.1: Gas-phase reaction path of H_2 + 2 (HCN), showing possible formation routes for the HCN dimer (nodes (8) and (10)) as well as aminoacetonitrile and methylcyanamide (both with the structure $H_4C_2N_2$).

route. Furthermore, the sampling of torsions allows us to find and use conformation isomers within the reaction pathway; molecules (**12**) and (**15**) are isomers of one another, but appear to be identical in Figure 4.1. The specifics of the energy differences can be found in Figure 5.1 of Section 5.1, and the structural differences are made clear by presenting them in the ball-and-stick format in Appendix B. A side-by-side comparison of Figure 5.1 and Figure 4.1 can be found at the end of Appendix C.

For all the calculations of the different molecules, geometry optimisation (Perdew & Wang 1992, Delley 2006) was initially done using the generalised-gradient approximation (GGA) with Perdew-Burke-Ernzerhof (PBE) correlation functional (Perdew et al. 1996) with basis set DNP, basis file 4.4. The core treatment parameter was set to "All Electron" and, therefore, the electrons were calculated as if they are valence electrons. After the geometry optimisation process, a conformation search was conducted using *Materials Studio's* Conformer module. All calculations were done at $0\,\mathrm{K}$, and the energy correction term was added to give Gibbs free energy at $298.15\,\mathrm{K}$. The zero-point vibrational energy (ZPVE) was accounted for in all calculations. Frequency calculations were used to confirm both optimised structures (an energy minimum with no imaginary frequency) and transition states (an energy maximum along with one imaginary frequency).

We believe these technical details have been sufficiently covered in Chapter 3, however, simply put: using *Materials Studio*, the molecules involved in each reaction are placed in roughly the orientation they are expected to react with each other. A trajectory can then be calculated and an energy diagram can be obtained: a minimum on the energy diagram corresponds to a new, stable configuration, while a maximum corresponds to a transition state (and thus also allows one to obtain the activation energy required for the reaction to take place). The molecular configurations at each energy minima were also found to not have imaginary frequencies, and thus correspond to true stable configurations (or rather, true minima on the potential energy surface, since that is what is implied by the lack of imaginary frequencies).

4.2 Ice-grain surface construction

The original rationale behind the modelling of the dust-grain surface was to obtain sufficient data to proceed with the OQS-model, as well as constrain variables to further this goal. As such, we attempted to use Computational chemistry to model an ice-surface which is, by and large, *the* method of investigation when it comes to chemical reactions in astrophysical environments, with the goal of improving on the results using open quantum systems. It thus required an understanding of how this modelling is done, and what results to expect from it.

Initial attempts at constructing an ice-surface were restricted to taking hexagonal crystalline H_2O (1h ice) from the *InfoMaticA* database (designation Pearson.1905259), which is a $7.82 \times 7.82 \times 7.36\ \text{Å}^3$ rhomboidal cell (see Figure 4.5), and creating a (3x3x3) supercell to form the bulk from which the slab will be cut. Following the procedure of Asgeirsson (2015), the bulk was cut along the (001) Miller plane, after which a $15\,\text{Å}$ vacuum gap was added above the cut. The use of the (001) Miller index for the cut allows us to have the ice-bilayers lie along the depth of the surface. The resulting *dangling H atoms* on the surface yield a much too ordered surface charge distribution, however, removing a random selection of them can greatly improve the

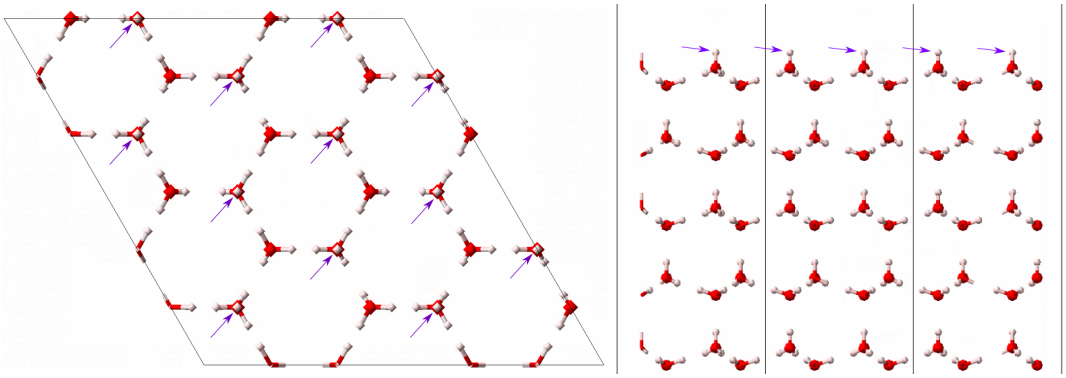


Figure 4.2: Fletcher phase of 1h ice with dangling hydrogen atoms indicated.

physical realism of the ice surface. The ordered proton structure of the surface is referred to as the Fletcher phase (Buch et al. 2008), seen in Figure 4.2. Afterwards, we used the structure optimisation routine of VASP, setting the convergence energy to $0.05\,\mathrm{eV/\AA}$ along with a maximum number of steps of 150. Again, the GGA-PBE exchange-correlation functional was used, along with manual selection of the number of k-points: 3 per cell of the original 1h-structure (as in Figure 4.5), for each axis — corresponding to selecting each constraint parameter to 9 in *Materials Studio*. These calculations did not converge due to a variety of reasons: care was not taken to have an integer number of bi-layers (see Figure 4.3), causing the structure to have a imbalanced charge distribution[2] and the optimisation to fail. Furthermore, great care needs to be taken with the selection of the parameters for the k-mesh in the *Brillouin Zone*[3](an introduction to this concept can be found in Eichler (2008)) otherwise the calculation does not converge and VASP returns an error.

Without in-depth knowledge of both the programming details behind VASP, and the mathematics used in density functional theory, attempts to remedy the convergence of the procedures were limited to freezing several of the bottom bilayers of the slab (i.e., although included in the evaluation of the dynamics of the system, they are treated as unmoving.) and playing around with parameters such as the K-mesh selection, convergence energy, integration scheme, and number of iterations used to compute. When these types of procedures were finally successful in converging and yielding results — using spacing of $0.25\,\mathrm{\AA}$ between k-points, with the rest of the settings as before — it was found that the reordering of the molecules were such that they appear to have either been optimised at a much higher temperature than was intended, or that the non-integer number of bilayers caused a dipole moment around the slab, morphing the structure in a way where it appears the hydrogen atoms on one side of the slab merely migrate towards the other side. This was likely a consequence of the total size of the cell in which the calculations were performed. Because of the periodicity in *all* directions, should the gap between the slab surface and the boundary of the cell be too small, it would cause the structure to interact with its own periodic copies. Although this is required for the axes running parallel to the surface (henceforth the x and y axis), it was not known at the time that the axis normal to the surface (z-axis) is also treated as being periodic. In both these cases the H atoms were ripped from the ice and expelled away from the surface, giving the appearance that the ice

[2]specifically, a dipole across the slab.
[3]a concept from solid state physics, hence not cover in Chapter 3

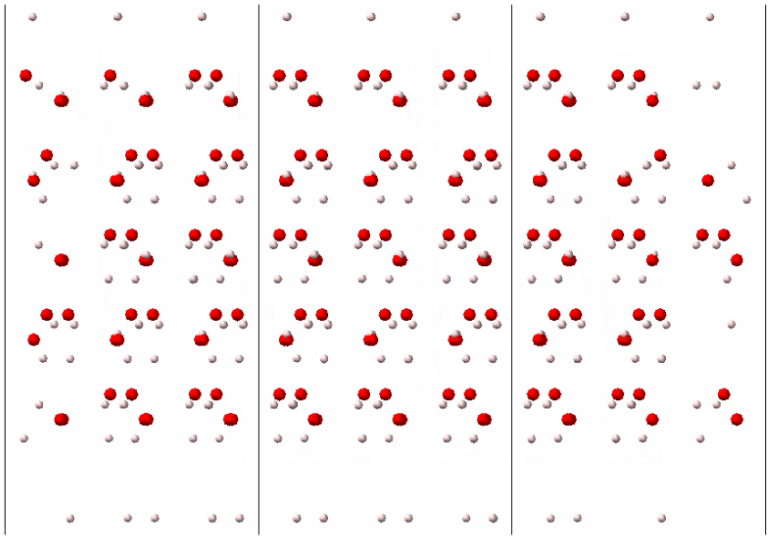


Figure 4.4: Hydrogen being expelled away from the surface when a 1h ice-structure is used to directly with no prior optimisation.

evaporated/exploded (see Figure 4.4). To reiterate: the calculations themselves were thus successful, the results were just unrealistic. This prompted us to switch over from direct structure optimisation as the method of investigation to Molecular dynamics, in the hopes of the system not developing some numerical intricacy.

Those computation parameters that were found to work for the structure optimisation (as in the previous paragraph) were kept identical for the rest of the MD simulations. Due to the small size of initial cell-sizes, as well as the initialisation temperatures, the MD simulations either yielded the same type of evaporation/explosion as some of the previous structure optimisations, or no movement of the molecules were found. We thus attempted to increase the number of molecules of the simulation by increasing the area of the slab, however, the problems persisted. We next tried to work with the uncut bulk structure (Figure 4.5), again creating a a supercell (3x3x3). We used the micro canonical and temperature scaling routines (number of molecules n, volume V, and energy E are kept as constants), however, these runs were found to produce equilibrium structures vastly different from the expected hexagonal form at low temperatures (Asgeirsson 2015). Possibly due to the aforementioned periodicity of the z-axis. Repeating the simulations using the previous configurations as input simply yielded identical results, and the structures that were found were thus not simply a local minimum that can be overcome by running the simulation multiple times.

The reasons for the strange results of the MD simulations were not immediately clear, and the selection of the k-points were suspected. Different integration schemes for the approximation, such as *Tetrahedron with Blöchl corrections*, manual selection of the k-points, and running the simulations at higher temperatures in an attempt to avoid unrealistic local equilibria, were all attempted. From this it was found that using the temperature scaling routine, along with manual selection of Density of states (DOS) k-mesh parameters and leaving the integration

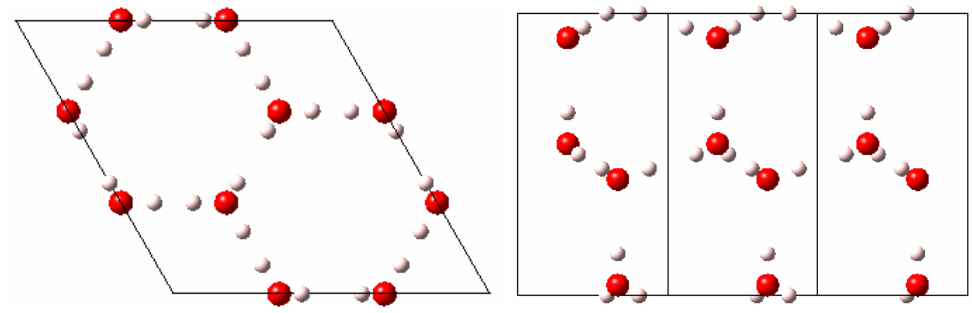


Figure 4.5: Top- and side-view of the Pearson.1905259 database ice (1h) structure, used as the bulk structure for the slab-formation.

schemes for both DOS and SCF as default yields not only a successful execution of the simulation, but also results consistent with what the physical situation for hexagonal ice at extremely low temperatures would be (i.e, that the molecules will only shift slightly when the temperature is raised from, say, 10 K to 30 K). This is thus the starting point from which we could finally construct an amorphous ice slab.

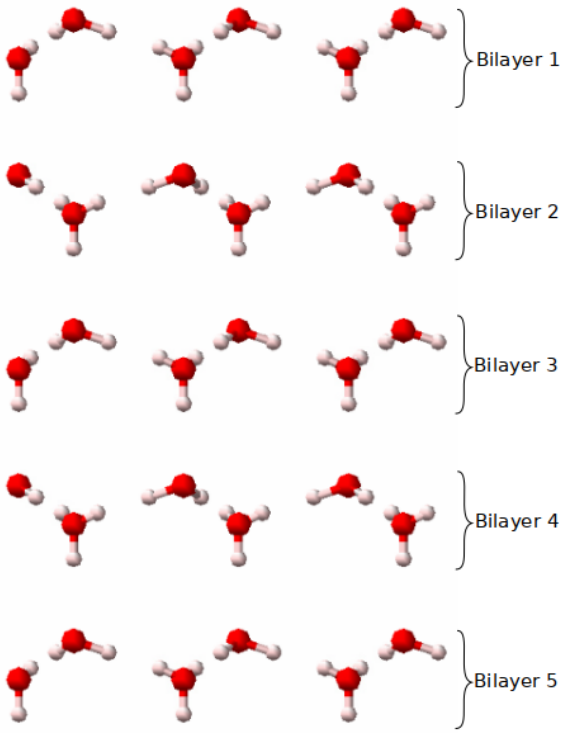


Figure 4.3: Side-view of the hexagonal ice structure showing the bilayers.

We started with the hexagonal ice-structure, as previously defined, and an assumed temperature of 8 K. The structure was allowed to evolve for 250 ps at constant temperature, after which the temperature was increased. These steps were repeated, with sequential temperature increases of roughly 50 K, up to 350 K. The structure was kept at 350 K for roughly 1000 ps and then cooled back down to 8 K using the same method[4]. This procedure of temperature scaling yields a bulk structure of amorphous solid water, which can then be cut, again along the (001) Miller plane, to form the ice slab (the final product is presented in Figure 5.4 on page 64). A hexagonal 1h ice-slab was also constructed to use as a reference. In this case the ice need not be annealed: a slab was cut from the bulk structure that is of the same dimensions as our annealed surface. To calculate the hexagonal slab energy involves first running an energy calculation where all atoms are both present and frozen[5], then removing a random selection of the dangling H atoms and redoing the energy calculation. At this point we can now simulate the adsorption of a molecule on the surface of an ice-grain by adding it to the model of the anneal slab, as described below.

[4] the process of heating up and cooling back down is known as *annealment*.
[5] Meaning they are forced to remain in place during the calculation.

4.3 Surface dimerisation

In order to obtain a quantitative value for the energy required for the surface dimerisation of HCN to take place, the relevant molecules of the reaction must be manually placed on or near the surface. The first step taken was to add a single HCN molecule to the ice-surface, however, several problems may arise (and, indeed, did arise) during this procedure. The periodicity of the z-axis, as mentioned in Section 4.2, may again cause unwanted behaviour: the new forces introduced by the addition of an HCN molecule slightly away from the ice-surface can cause a perturbation large enough that the entire surface, once again, appears to want to spread out in the z-axis — specifically into the vacuum above and below the surface, homogenising the forces throughout the cell. It was thus required to further increase the size of the gap above and below the surface to well above a distance that any possible periodic interaction can occur. The second difficulty one encounters with the adsorption procedure is due to the morphology of the surface itself, however, the discussion thereof will be left to the end of the next paragraph.

For the introduction of an HCN molecule to the surface, we chose several sites (at random) on the surface, several orientations of the molecule, as well as several distances from the surface to place the molecule before starting another calculation as in Section 4.2. A few of the different configurations for one such site is shown in Figure 4.6. At each chosen distance from the surface the HCN molecule was rotated such that at one end we have the Hydrogen atom closest to the surface, and on the other end we have the Nitrogen atom closest to the surface, with six rotations in-between. Each configuration was also run at several temperatures, starting from $8\,\mathrm{K}$ and incrementally increasing up to $25\,\mathrm{K}$. This temperature range was justified in our discussions in Section 2.1, with the upper limit of $25\,\mathrm{K}$ having been chosen because we found it to be sufficient for enough molecular movement to take place that it should be capable of facilitating the dimerisation process. There is, however, a flaw in the reproducibility of this method of investigation: with all the free-parameters chosen for the configuration of the molecule relative to the surface, as well as the complexity of the surface morphology, that the problem has essentially become a chaotic[6] one. This means that, should we indeed find HCN adsorption onto dust-grains facilitates their dimerisation, we have no real way to determine exactly what the physical process is that causes this. Furthermore, should we find that the aforementioned adsorption has no effect, we can only conclude that it has no effect for the specific morphology of our, relatively small, ice-surface model.

The addition of a single HCN molecule to the surface, and looking at what happens to it is, however, as far as this study was taken. The extremely time-consuming nature of modelling chemical systems means we have, regrettably, have had be content with the progress made thus far. We do, however, discuss potential pitfalls and recommendations for the addition of the second HCN molecule, as well as other possible approach that may be better suited to the task in Section 6.5. The one thing that can probably not be avoided though, is computational cost.

[6] In the mathematical sense of sensitivity to initial conditions.

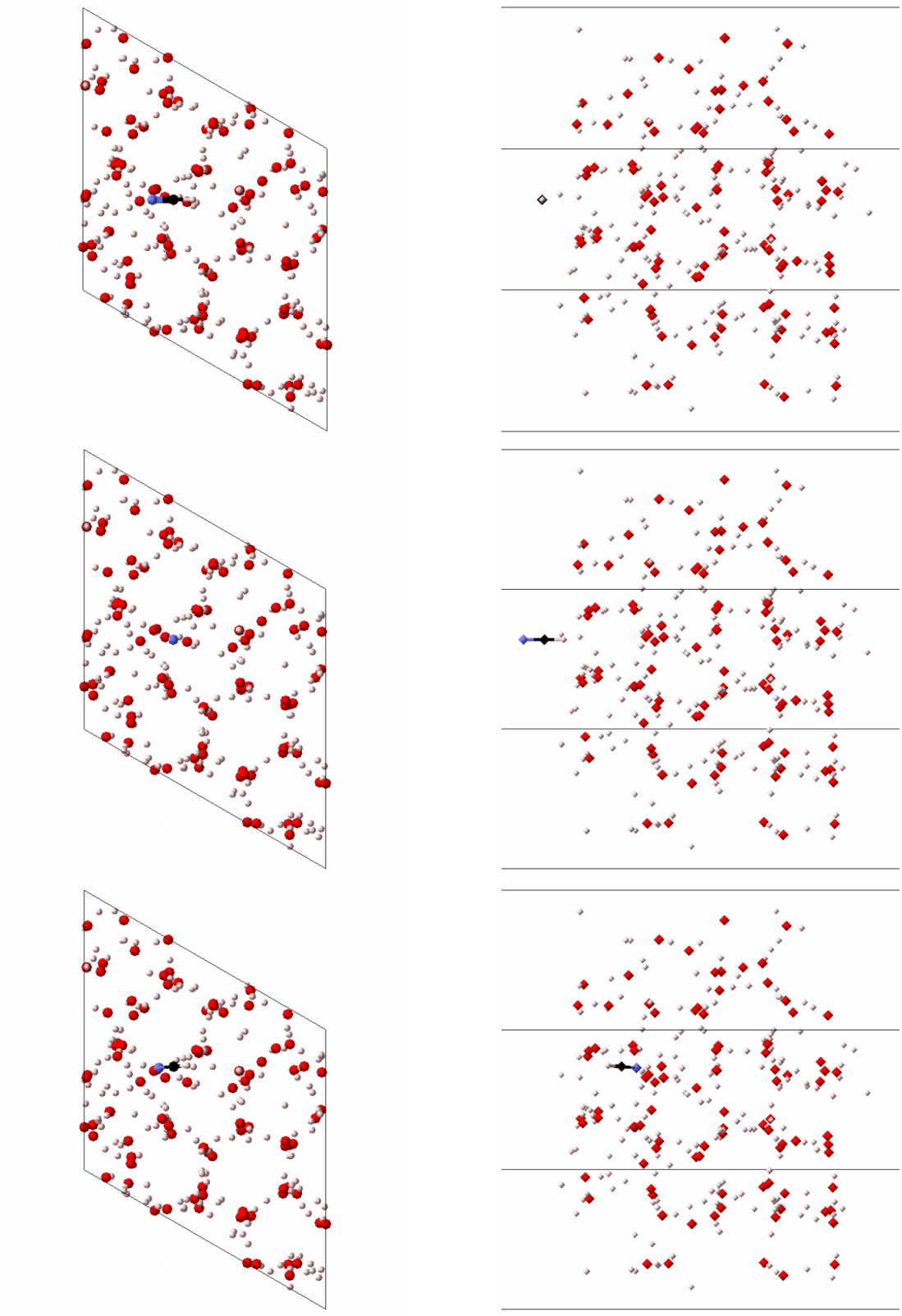


Figure 4.6: Top- and side-views of some HCN adsorption configurations.

Chapter 5

Results

The results that can be considered as answering our primary research objectives, as well as those that have direct implications on them, are presented in this chapter. Note that, although the majority of the discussion of the results and their implications are found in the next chapter, Chapter 6, we will endeavour to explain everything presented to the point where it can be understood and simply used as a reference when reading the accompanying discussion.

The results are presented in the same sequential manner as in the previous chapter: The complete chemical reaction pathway of the HCN dimer, cyanomethanimine, as well its accompanying energy level diagram, are given in Section 5.1. The ice-surface that was used to model the dust-grain surface, as well as all results which follow directly from it, are found in Section 5.2.

5.1 Gas-phase formation route

The main results of the gas-phase formation route (Figure 4.1) are found in its corresponding energy level diagram, Figure 5.1. The molecules (**2**), (**7**), (**9**), (**6**), (**11**), (**15**), and (**12**) all indicate transition states. Of the transition states, (**6**), (**2**), (**7**), (**9**), and (**11**) are barriers, with (**15**) and (**12**) being spontaneous. The only other reactions that posses an activation energy are (**1**) to (**5**), (**3**) to (**4**), and between (**8**) and (**17**).

Figure 5.1 shows that cyanomethanimine formation can occur through the direct interaction of two HCN molecules, through either the transition states (**9**) or (**7**), with respective barrier energies of $399.86\,\mathrm{kJ/mol}$ and $297.53\,\mathrm{kJ/mol}$. The use of temperature in low-density environments can often be misleading, as its statistical definition under these circumstances can lead to temperatures of tens of thousands to several million Kelvin. The energy referred to in Figure 5.2 should thus be thought of as all forms of energy that can be provided by the environment to facilitate these reactions. The figure shows that we can expect our formation route to start when this 'environmental energy' reaches roughly $110\,\mathrm{kJ/mol}$ [1], whereafter we would expect the rapid formation of the molecules found in nodes (**8**), (**10**), (**3**), (**19**) and (**14**). Node (**8**) contains the N-cyanomethanimine and Z-C-cyanomethanimine dimers[2], (**10**) contains the E-C-cyanomethanimine, (**3**) contains methanimine (H_2CNH), with (**19**) and (**14**) yielding

[1] equivalent to 20% of the largest energy barrier present in the reaction pathway, or 0.2 fractional energy.
[2] The energy difference between the two is only about $6\,\mathrm{kJ/mol}$.

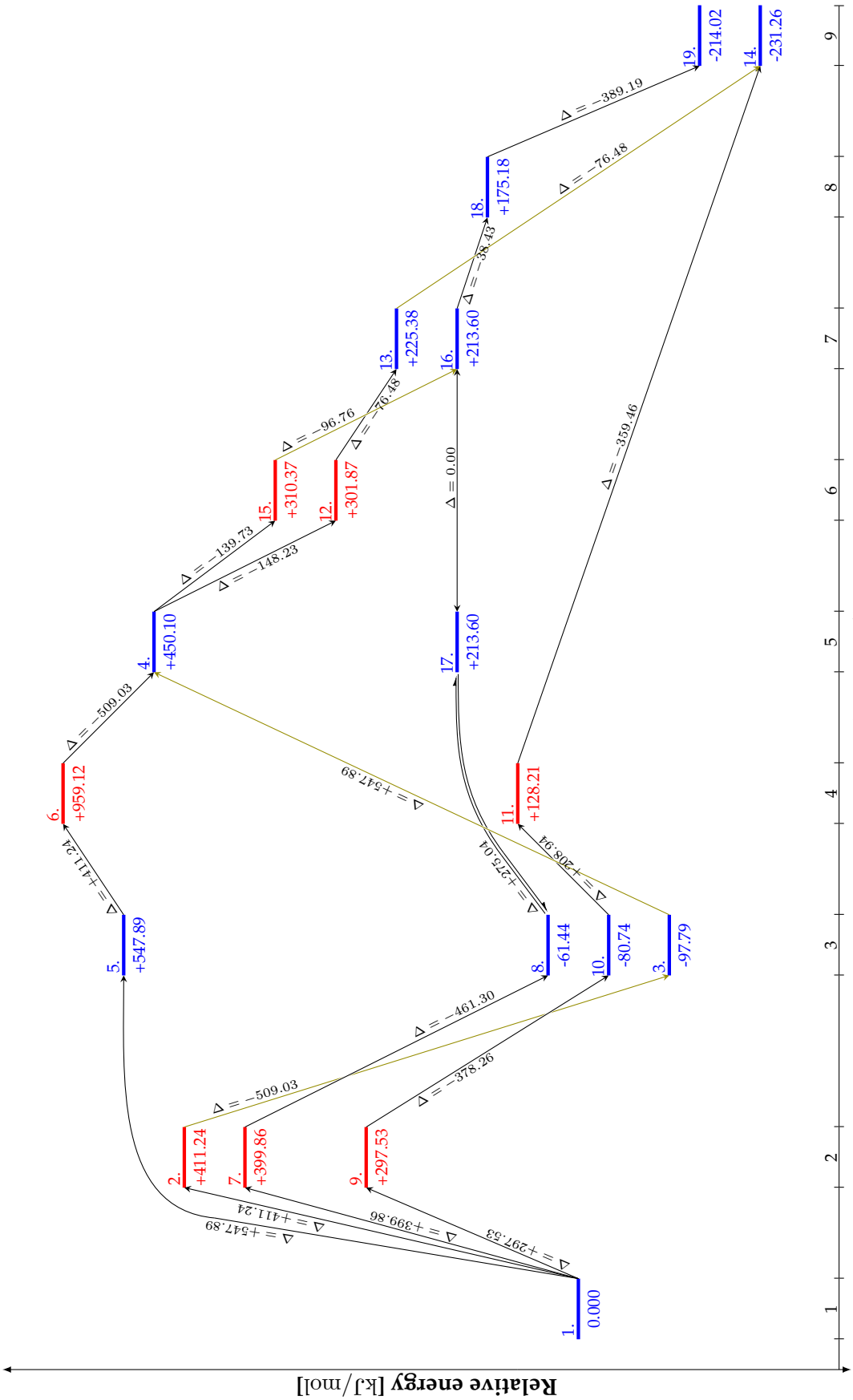


Figure 5.1: Energy level diagram (energies in kJ/mol) of the reaction pathway as discussed in Chapter 4 and Figure 4.1. Transition states are marked as red. Additional use of colour is to improve readability.

methylcyanamide and aminoacetonitrile respectively. All molecules proposed for formation by this gas-phase reaction pathway have been confirmed as detected in space (see Table 2.1 on page 8). Note, however, this does not imply that they are all formed via the reaction pathway of Figure 4.1. We find our overall energy barriers to be comparable to those found by Vazart et al. (2015), specifically those for the formation of cyanomethanimine through the interaction of a CN radical with methanimine (H_2CNH). This is in support of their hypothesis that HCN dimers should be able to form in the gas-phase under the constrains of astrophysical environments, and also suggest that should the environment lack sufficient energy, that the direct dimerisation of HCN through HCN + HCN $\longrightarrow$ $(HCN)_2$ (node (**5**) of Figure 5.1) is unlikely to occur — supporting the claim of Smith et al. (2001).

A qualitative analysis of the likelihood of molecules taking different reaction pathways can be obtained by means of treating each chemical configuration along the reaction route as a thermodynamic state. The implicit assumptions of such an analysis is discussed in Section 6.2, but the validity thereof, at least for a rudimentary investigation, is justified when comparing Maxwell-Boltzmann statistics to the Arrhenius equation (see Equation 2.7 on page 19). In the realm Maxwell-Boltzmann statistics, the expected probability for a particle to have an energy of ϵ_i is given by

$$P(\epsilon_i) = \frac{1}{Z} e^{-\epsilon_i / kT}, \tag{5.1}$$

where

$$Z = \sum_i e^{-\epsilon_i / kT}. \tag{5.2}$$

While the above equations and the Arrhenius equation compare different things (one is a probability and the other has units of s^{-1}), multiplying Equation 5.1 with the collisional rate of the molecules will yield approximately the same answers; the complex physical details contained in the pre-exponential factor should be the biggest difference. The transitional probability of each state to each other available state is presented in Figure 5.2. Note that there are no such graphs for nodes (**14**) or (**19**), as these are both end points of the reaction routes[3]. Further note the energy axis is in units of a fraction of the largest energy barrier present in the route, namely the barrier between (**1**) and (**5**) of $547.89\,\mathrm{kJ/mol}$. The most important thing to look at in Figure 5.2 is the top-left panel which shows the value of Equation 5.1 (thus the probability) of having molecules (**1**), thus H_2 + 2 (HCN), go to any of those described by (**5**), (**2**), (**7**), (**9**), or simply remaining at (**1**), all as a function of the energy available to these molecules (the 'temperature' of the molecules). All panels show both the probability of remaining in the current molecular setup (i.e., not reacting in a way predicted by the reaction path), as well as the probability to move along the reaction pathway to one of the permitted combination of molecules. What then can be gleaned from this figure is, that, should roughly $130\,\mathrm{kJ/mol}$ of energy be available to the molecules, the reaction path will produce non-negligible amounts of the molecules mentioned previously.

[3]That is not to say we believe no other reactions take place involving these molecules, just that this is where our calculations end.

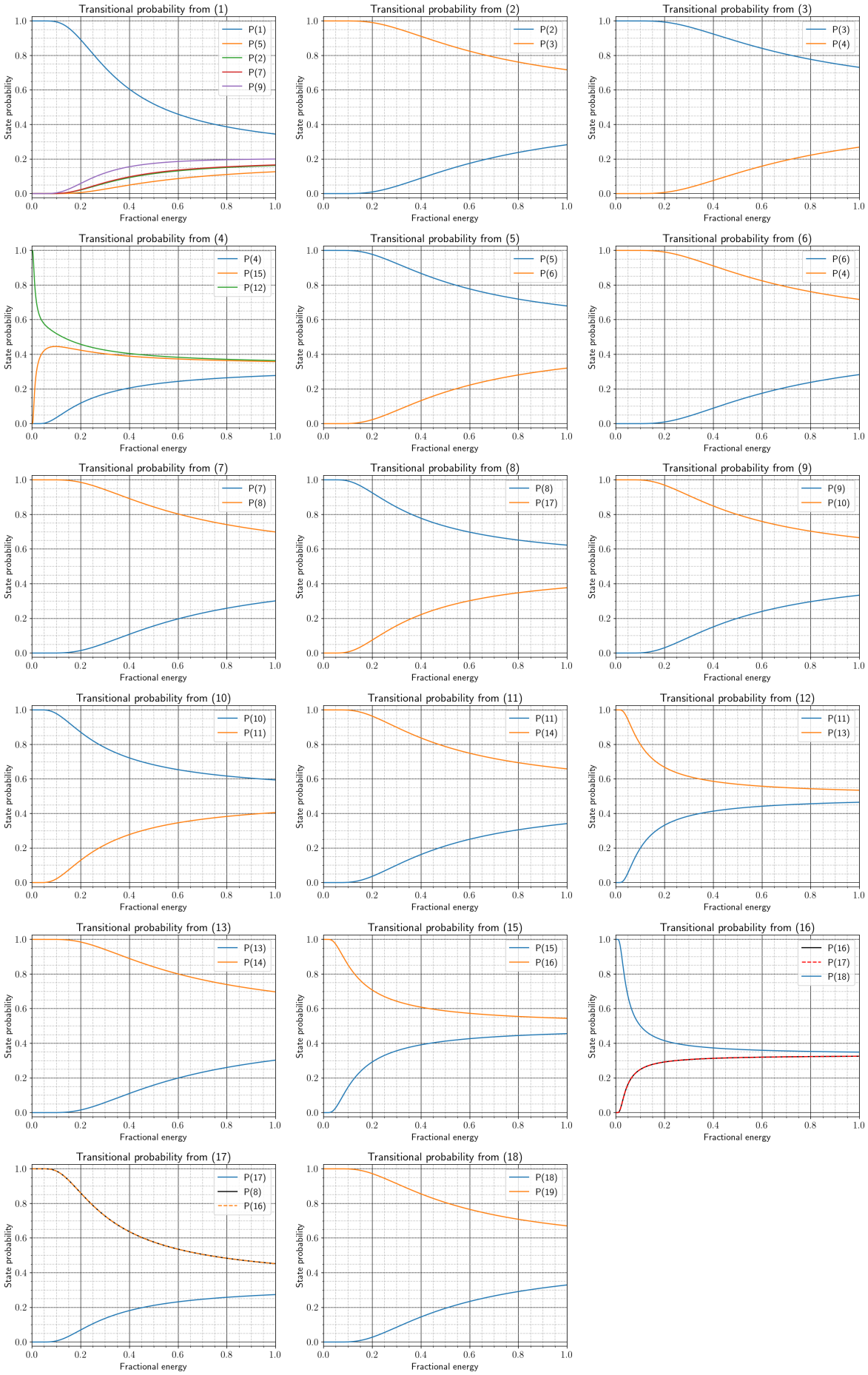


Figure 5.2: Transitional probabilities of each gas-phase states as a function of a fraction of the largest barrier energy, constructed from treatment of the chemical configurations as thermodynamic states.

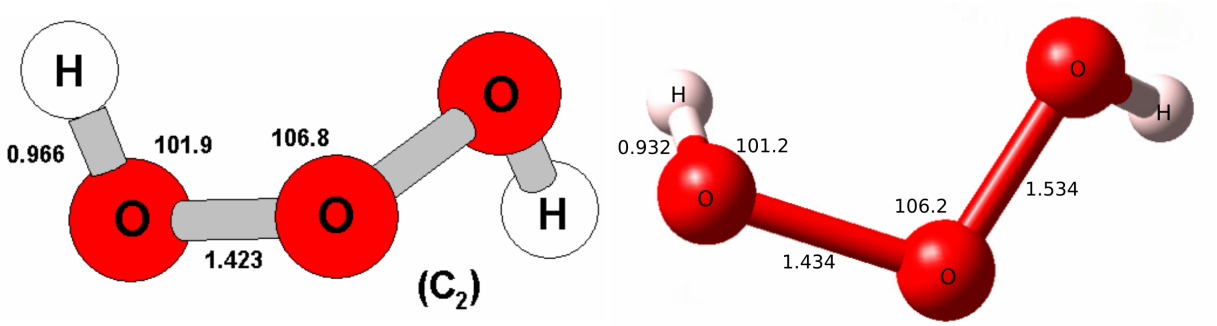


Figure 5.3: Geometrical parameters (angles in degrees, and distances in angstrom) for HOOOH, taken from Plesničar (2005), as compared to those of the apparent trioxidane found in our ice-slab.

5.2 Ice-grain surface

We were able to construct an ice-slab under conditions suitable for it to serve as a crude model for the surface of an extraterrestrial dust grain. It was found that the annealment process produces a surface that appears to behave in a realistic manner at low temperatures: little to no molecular movement is found when doing molecular dynamic simulations at temperatures below 20 K, with noticeable movement occurring for temperatures of 25 K and above. Choosing the vacuum gap (which defines the surface) to be sufficiently large is essential in avoiding unwanted periodic interaction. A gap-size of 30–40 Å was found to be sufficient, although over-selecting the size of the gap should have no considerable down-side. The construction process described in Section 4.2 led to the ASW structure as seen in Figure 5.4. Hints of the hexagonal pattern of the 1h-ice structure remain but in terms of what an adsorbed molecule will experience, the slab, as we have constructed it, should serve as an approximation for the surface of dust-grain as found in space.

Several interesting molecules appeared to form during the annealment process, shown in Figure 5.5, and are thus present in the final ice-slab. The two outer products of Figure 5.5 are not molecules — the way *Materials Studio* displays bonds is misleading: all atoms within a certain distance of one another have a bond displayed between them. In the case of Figure 5.5, this 'bond display length' seams to be 1.5 Å. The middle product of Figure 5.5 appears to be trioxidane (HOOOH), not a molecule that has been detected in space (refer to Table 2.1), although the comparison presented in Figure 5.3 shows it to have roughly the correct geometry. The fact that the 'bonds' are displayed simply based on the atomic separation, instead of being somehow calculated, means that we can take note of these structures that appear to form but will not claim this is somehow evidence that these molecules should form in the ice layers of dust-grains. It is also possible that we what we assumed to be molecules could be looked at as transition states. In this case the hydrogen atom with two displayed bonds could belong to either of the water molecules. It would be worth investigating whether, with a larger system or one at a slightly higher temperature, we see the formation of a polymeric chain.

As mentioned in Section 4.3, we did not reach a point where we could conclusively test for the dimerisation of HCN on the ice-surface: In the case of the extremely low temperature calculations (8–15 K), no chemical interaction (between HCN and the surface) was found. Whether this is a consequence of the lack of energy available to the system, or due to some numerical

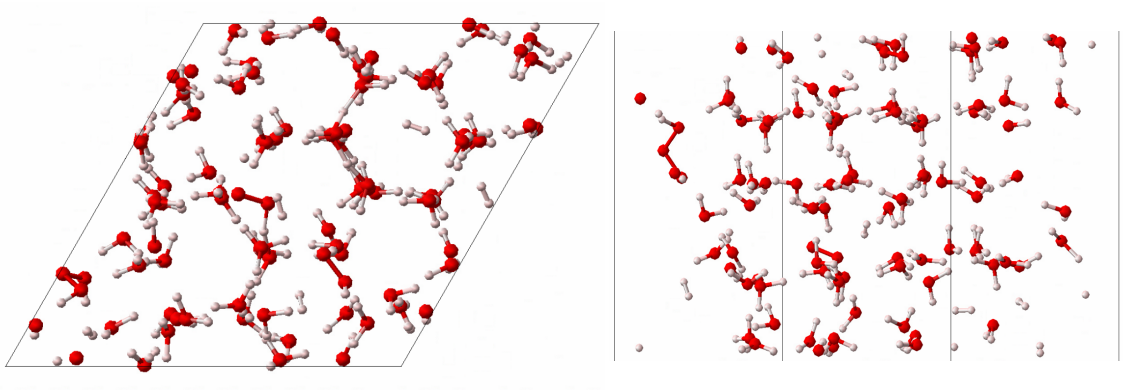


Figure 5.4: Top- and side view of the annealed ice-slab. Vacuum gap above the surface not depicted. Coordinates can be found in Table C.1.

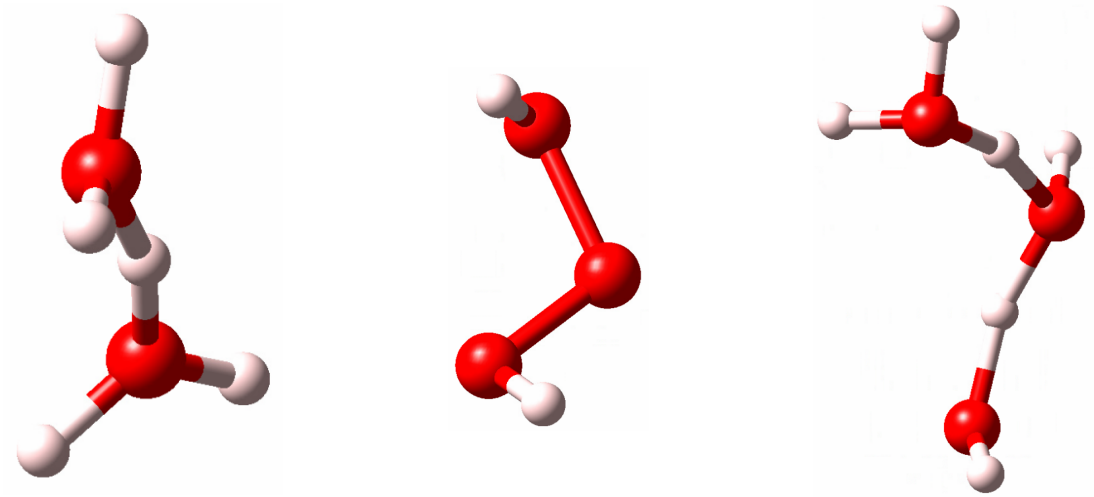


Figure 5.5: Apparent structures formed in the bulk during the annealment process.

intricacy of the model (and software) is still uncertain, however, the former is suspected. As with the case without the HCN molecule in the system, noticeable atomic movement occurs for temperatures at and above $25\,\mathrm{K}$, with what appears to be the separation of HCN $\longrightarrow$ H + CN, however, the bond-display issue we mentioned makes it difficult to conclusively state what exactly is happening to the hydrogen cyanide molecule.

Chapter 6

Discussions and conclusion

The general take-aways of this study and of prebiotic research are presented in Section 6.1. The rest of the Chapter follows the same structure as those before it, with the discussions of the gas-phase HCN dimerisation found in Section 6.2, those relating to the slab surface, and the incorporation and comparison of HCN dimerisation thereupon in Section 6.3. For the sake of emphasis we again mention the use of open quantum systems as a replacement for density functional theory in Section 6.4. We conclude this text in Section 6.5.

6.1 General discussion

We believe the interdisciplinarity of Astrobiology has been properly demonstrated in this work. It is far from an absolute necessity to have expertise in all fields involved, as focused research will merely lead to an iterative research approach where key ideas can be tested and systematic progress can be made. However, it is the author's opinion that, should one wish to dedicate oneself to becoming an astrobiologist, the importance of a comprehensive knowledge of all fields mentioned in this text cannot be understated.

The concept of a chemical network has been discussed, and while this allows one to test for a tremendous amount of complex chemical reactions, it must be remembered that it should be handle in the same way as a 'database of previous results'. Simply because one's chemical network does not contain a certain reaction, it should not be thought that the reaction does not take place or, indeed, could be of great importance. Thus any Astrobiology research group using an astrochemical network should be constantly refining it with thoroughly investigated and detailed chemical reactions. The retroactive-research component of this is comprised of looking at which molecules have been observed and which we should be lending our attention towards, requiring a knowledge of Astrochemistry for the investigation of the detection and possible formation. When this is then applied to prebiotic molecules, one either needs experiments that mimic their astrophysical environments, or to use chemical modelling for a theoretical study. Should this avenue of research have been followed, we are still left with a decision of whether to include the effects that the presence of dust-grains, and thus an available surface for chemical reactions, has on both allowed/viable formation routes and the changes these routes introduce to the gas-phase reactions. The study of chemical reactions that take place on some surface is an entire field of study on its own, known as surface science. The

more one tries to improve one's understanding of the problem, the more the benefit of having an extensive knowledge of the different fields is highlighted.

In terms of what, quantitatively, astrobiological research entails, that which has been presented in this work should serve as a quintessential example — both in terms of the actual problem, namely HCN dimerisation in different phases, as well as the methodology we have applied in our attempt to resolve the problem. The strategy for doing astrobiological research is liable to change. Until a definitive answer to the formation of DNA or RNA, or something as yet unknown, has been found, the approach we have outlined in this work should be sufficient. Should it be discovered how DNA and RNA can be formed in space, then Astrobiology will cease being *prebiotic*, and the former Chemistry component will slowly have to yield to a field closer to Molecular Biology; until such time we hope this work can be of some use to those embarking on the journey that is prebiotic research.

6.2 On the Gas-phase dimerisation of HCN

The most basic conclusion we can draw from Section 5.1, is that it appears the formation of the HCN dimer, cyanomethanimine, is possible in the gas-phase. Furthermore, the means by which the dimerisation occurs, as presented in Figure 4.1 and Figure 5.1, appears to support the original criticism of Smith et al. (2001) — at least for the first step of the oligomerisation process suggested by Chakrabarti & Chakrabarti (2000), as shown in Equation 2.8 on page 23. Figure 5.1 gives a quantitative answer to the energy barriers restricting direct gas-phase dimerisation of HCN, and Figure 5.2 gives an estimate of the likelihood of formation based on available energy. The fact that the products in our formation route — (**19**) methylcyanamide, (**14**) aminoacetonitrile, (**3**) methanimine and the HCN dimers (**8**) and (**10**) — have also been detected in astrophysical environments lends confidence to the validity of our results (refer to Table 2.1 and Borget et al. (2012)).

We believe that, in this work, we have presented a viable formation route for methylcyanamide, aminoacetonitrile, methanimine, and cyanomethanimine — all of which have a sufficiently low energy cost for formation that they remain viable given typical astrophysical parameters. The occurrence of atypical events may lead to direct HCN dimerisation and, as such, it shouldn't be neglected from chemical networks that are focussed on prebiotic chemistry. Our results are aligned with those of Vazart et al. (2015), and we believe the HCN dimerisation problem has been solved. Using our reaction path (Figure 4.1), further investigation of prebiotic molecule formation beyond what we have presented can now be made. We have not looked into what possible roles methylcyanamide or aminoacetonitrile play in Astrobiology, but neither molecule is far removed from the general make-up of the nucleobases.

6.3 Dust-grain surfaces

Although we ultimately failed to show the dimerisation of adsorbed HCN, we believe a large step towards achieving this goal has been made. We succeeded in creating an ice-surface with roughly the properties one would expect of the surface of dust-grain — at least as a first step towards a truly realistic model. A very powerful means of investigation would be to use the same

procedure of annealment, as in Section 5.2, to construct multiple surface models. We suggest this instead of merely creating a much larger surface, since *the* limiting factor we found was computation cost (and thus time). The use of multiple surface models allows one to remove some of the variance that might occur in the calculations due to the surface's morphology. We believe the inclusion of impurities (such as carbon atoms and other molecules) in the model is not yet of vital importance. These molecules should be sparse enough as to not contribute significantly to the overall abundance of surface-reactions. When one reaches the point of having a satisfactory model for the surface of a dust-grain, careful consideration is needed when deciding the HCN orientation on the surface: The addition of the second HCN molecule to the surface, specifically for the purpose of testing dimerisation, is not as trivial as simply adding it in close proximity to the first, as this in no way ensures they interact in such a way as to form cyanomethanimine. An elegant approach, in the author's opinion, would be to use a Monte Carlo method for the way in which the HCN molecule approaches the surface. This would allow for the testing of all the surface reaction types as discussed in Section 2.2.3 on page 17, and is easily implemented.

6.4 Open quantum systems and molecular modelling

At the very start of this text we talked about the original idea of using open quantum systems as a way to simplify the problem of surface dimerisation (Section 1.3 on page 4), and that the original intention of modelling the problem using DFT was to merely provide information to test this idea against. As should be evident by how we have presented this text, the original idea was not deemed feasible for the time being. The manner in which Chapter 3 was presented clearly demonstrates the degree to which DFT methods take quantum mechanics into account — and that the incorporation of OQS and density functional theory is not a trivial task. Further development of the method by D'Agosta & Di Ventra (2013) will likely lead to a more advanced method of investigation than the one we have used in this text, but for the time being ordinary density functional theory still has plenty to contribute.

6.5 Conclusions and recommendations

This project was the author's first encounter with Astrobiology. We hope this text has provided the reader with useful information regarding the interdisciplinarity of the field, and enough detail that not much of how to go about doing prebiotic research is left unanswered. Recommendations have been included in each respective section of this chapter, but we would like to point out that these ideas should not be kept in isolation — the eventual end-goal of an Astrobiology research group should be to not only investigate the formation of *some* lone molecule, such as the HCN dimer, but to investigate how different prebiotic molecules form and eventually interact, leading to more complex structures. This can only be done through the use of all available resources: chemical networks, both gas-phase reactions and those that involve adsorbed molecules, and an indefatigable research ethic.

Life *did* somehow form, and it is the goal of research such as this to narrow down how and where exactly it happened.

Appendix A

Notable interstellar reaction processes

Presented are some notable interstellar reaction process that are could be of potential use in this study. The processes are focused around the following molecules: HCN, HNC, H_2O, CH, CN, and NH. The reaction types and corresponding tables we have included are:

1. Neutral-neutral reactions
2. Ion-molecule reactions
3. Associative detachment
4. Dissociative electron recombination
5. Photoionisation
6. Photodissociation
7. Radiative association
8. Charge transfer

Each reaction type includes the formula used to calculate the reaction rate coefficient k, described in Section 2.2.4, along with other required parameters.

Class:		Neutral-neutral reactions			
Form:		$k = \alpha\,(T/300)^{\beta}\,e^{-\Gamma/T}$			
Reactants		Product	α	β	Γ
$HCN + CN$	$\rightarrow$	$NCCN + H$	2.50×10^{-17}	1.71	770.0
$HCN + H$	$\rightarrow$	$CN + H_2$	6.20×10^{-10}	0.00	12 500.0
$CN + H_2$	$\rightarrow$	$HCN + H$	4.04×10^{-13}	2.87	820.0
$CH + N_2$	$\rightarrow$	$HCN + N$	5.60×10^{-13}	0.88	10 128.0
$HNC + H$	$\rightarrow$	$HCN + H$	1.14×10^{-13}	4.23	−114.6
$HNC + CN$	$\rightarrow$	$NCCN + H$	2.00×10^{-10}	0.00	0.0
$N + CH_2$	$\rightarrow$	$HNC + H$	3.95×10^{-11}	0.17	0.0
$C + NH_2$	$\rightarrow$	$HNC + H$	3.26×10^{-11}	−0.10	−9.0
$H_2O + H$	$\rightarrow$	$OH + H_2$	1.59×10^{-11}	1.20	9610.0
$H_2O + NH$	$\rightarrow$	$OH + NH_2$	1.83×10^{-12}	1.60	14 090.0
$NH + OH$	$\rightarrow$	$H_2O + N$	3.11×10^{-12}	1.20	0.0
$OH + HCN$	$\rightarrow$	$H_2O + CN$	1.87×10^{-13}	1.50	3887.0
$CH + N_2$	$\rightarrow$	$HCN + N$	5.60×10^{-13}	0.88	10 128.0
$CH + N$	$\rightarrow$	$CN + H$	1.66×10^{-10}	−0.09	0.0
$H_2 + C$	$\rightarrow$	$CH + H$	6.64×10^{-10}	0.00	11 700.0
$OH + C$	$\rightarrow$	$O + CH$	2.25×10^{-11}	0.50	1 480 000.0
$CN + OH$	$\rightarrow$	$HCN + O$	1.00×10^{-11}	0.00	1000.0
$CN + NH$	$\rightarrow$	$HCN + N$	2.94×10^{-12}	0.50	1000.0
$H + NCCN$	$\rightarrow$	$CN + HCN$	1.48×10^{-10}	0.00	3588.0
$O + HCN$	$\rightarrow$	$CN + OH$	6.21×10^{-10}	0.00	12 439.0
$NH + H_2$	$\rightarrow$	$NH_2 + H$	5.96×10^{-11}	0.00	7782.0
$NH + C$	$\rightarrow$	$CN + H$	1.20×10^{-10}	0.00	0.0
$H_2 + N$	$\rightarrow$	$NH + H$	1.69×10^{-9}	0.00	18 095.0
$OH + N$	$\rightarrow$	$NH + O$	1.88×10^{-11}	0.10	10 700.0

Class:		Ion-molecule			
Form:		$k = \alpha\,(T/300)^{\beta}\,e^{-\Gamma/T}$			
Reactants		Product	α	β	Γ
$HCN + HCN^+$	$\rightarrow$	$HCNH^+ + CN$	1.60×10^{-9}	−0.50	0.0
$HCN + H_2O^+$	$\longrightarrow$	$HCNH^+ + OH$	2.10×10^{-9}	−0.50	0.0
$NH_3 + C_2N^+$	$\rightarrow$	$HCN + HCNH^+$	1.8×10^{-9}	−0.50	0.0
$HC_3N + C^+$	$\rightarrow$	$HCN + C_3{}^+$	2.50×10^{-10}	−0.50	0.0
$HNC + HCN^+$	$\rightarrow$	$HCNH^+ + CN$	1.00×10^{-9}	−0.50	0.0
$HNC + NH^+$	$\rightarrow$	$HCNH^+ + N$	1.80×10^{-9}	−0.50	0.0
$HNC_3 + H^+$	$\rightarrow$	$HNC + C_2H^+$	3.23×10^{-9}	−0.50	0.0
$HCNH^+ + HC_3N$	$\rightarrow$	$HNC + HC_3NH^+$	1.70×10^{-9}	−0.50	0.0
$H_2O + C_2N^+$	$\rightarrow$	$HCN + HCO^+$	2.30×10^{-10}	−0.50	0.0
$H_2O + CN^+$	$\rightarrow$	$HCN^+ + OH$	1.60×10^{-9}	−0.50	0.0
$HCN + OH^-$	$\rightarrow$	$H_2O + CN^-$	1.20×10^{-9}	0.00	0.0
$CH_2 + O_2$	$\rightarrow$	$H_2O + CO$	2.48×10^{-10}	−3.30	1443.0
$CH + HCNH^+$	$\rightarrow$	$HCN + CH_2{}^+$	3.15×10^{-10}	−0.50	0.0
$CH + N^+$	$\rightarrow$	$CN^+ + H$	3.60×10^{-10}	−0.50	0.0
$CH_2{}^+ + NH_3$	$\rightarrow$	$CH + NH_4{}^+$	1.26×10^{-9}	−0.50	0.0
$HNC_3 + C^+$	$\rightarrow$	$CH + C_3N^+$	1.55×10^{-9}	−0.50	0.0
$CN + H_2{}^+$	$\rightarrow$	$HCN^+ + H$	1.20×10^{-9}	−0.50	0.0
$CN + NH^+$	$\rightarrow$	$HCN^+ + N$	1.60×10^{-9}	−0.50	0.0
$C_2 + HCN^+$	$\rightarrow$	$CN + C_2H^+$	8.40×10^{-10}	0.00	0.0
$C + HCN^+$	$\rightarrow$	$CN + CH^+$	1.10×10^{-9}	0.00	0.0
$NH + HCN^+$	$\rightarrow$	$CN + NH_2{}^+$	6.50×10^{-10}	−0.50	0.0
$NH + O_2{}^+$	$\rightarrow$	$HNO^+ + O$	3.20×10^{-10}	−0.50	0.0
$NH_2{}^+ + HCN$	$\rightarrow$	$HCNH^+ + NH$	1.20×10^{-9}	−0.50	0.0
$NH_2{}^+ + HNC$	$\rightarrow$	$HCNH^+ + NH$	1.20×10^{-9}	−0.50	0.0

Class:		Dissociative electron recombination			
Form:		$k = \alpha\,(T/300)^{\beta}\,e^{-\Gamma/T}$			
Reactants		Product	α	β	Γ
$HCNH^+ + e^-$	$\rightarrow$	$HCN + H$	9.50×10^{-8}	-0.65	0.0
$HCNOH^+ + e^-$	$\rightarrow$	$HCN + OH$	1.00×10^{-7}	-0.50	0.0
$CH_2CN^+ + e^-$	$\rightarrow$	$HCN + CH$	1.50×10^{-7}	-0.50	0.0
$HC_3N^+ + e^-$	$\rightarrow$	$HCN + C_2$	3.60×10^{-7}	-0.60	0.0
$H_2NC^+ + e^-$	$\rightarrow$	$HNC + H$	1.80×10^{-7}	-0.50	0.0
$HCNH^+ + e^-$	$\rightarrow$	$HNC + H$	9.50×10^{-9}	-0.65	0.0
$NCCNH^+ + e^-$	$\rightarrow$	$HNC + CN$	3.00×10^{-7}	-0.50	0.0
$CH_3CNH^+ + e^-$	$\rightarrow$	$HNC + CH_3$	2.85×10^{-7}	-0.69	0.0
$C_2H_5OH_2^+ + e^-$	$\rightarrow$	$H_2O + C_2H_4 + H$	1.77×10^{-6}	-0.73	0.0
$H_2OCN^+ + e^-$	$\rightarrow$	$H_2O + CN$	1.50×10^{-7}	-0.50	0.0
$H_3CO^+ + e^-$	$\rightarrow$	$H_2O + CH$	1.40×10^{-8}	-0.78	0.0
$H_3O^+ + e^-$	$\rightarrow$	$H_2O + H$	7.09×10^{-8}	-0.50	0.0
$C_2H^+ + e^-$	$\rightarrow$	$CH + C$	1.53×10^{-7}	-0.76	0.0
$CH_2^+ + e^-$	$\rightarrow$	$CH + H$	1.60×10^{-7}	-0.60	0.0
$C_2H_2^+ + e^-$	$\rightarrow$	$CH + CH$	9.00×10^{-8}	-0.50	0.0
$H_2C_4N^+ + e^-$	$\rightarrow$	$HC_3N + CH$	3.00×10^{-7}	-0.50	0.0
$C_2N^+ + e^-$	$\rightarrow$	$CN + C$	1.50×10^{-7}	-0.50	0.0
$HONC^+ + e^-$	$\rightarrow$	$CN + OH$	1.50×10^{-7}	-0.50	0.0
$HCNH^+ + e^-$	$\rightarrow$	$CN + H + H$	9.30×10^{-8}	-0.65	0.0
$H_2NC^+ + e^-$	$\rightarrow$	$CN + H_2$	1.80×10^{-8}	-0.50	0.0
$HNCO^+ + e^-$	$\rightarrow$	$NH + CO$	1.50×10^{-7}	-0.50	0.0
$HNCOH^+ + e^-$	$\rightarrow$	$NH + HCO$	1.00×10^{-7}	-0.50	0.0
$HN_2^+ + e^-$	$\rightarrow$	$NH + N$	2.09×10^{-8}	-0.74	0.0
$NH_2^+ + e^-$	$\rightarrow$	$NH + H$	9.21×10^{-8}	-0.79	17.1

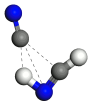
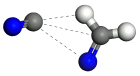
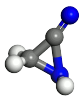
Class:		Photodissociation		
Form:		$k = \alpha e^{-\Gamma A_v}$		
Reactants		Product	α	Γ
$HCN + \gamma$	$\rightarrow$	$CN + H$	1.30×10^{-17}	1557.0
$H_2CN + \gamma$	$\rightarrow$	$HCN + H$	5.48×10^{-10}	2.0
$HNC + \gamma$	$\rightarrow$	$CN + H$	1.30×10^{-17}	1500.0
$CH_2NH + \gamma$	$\rightarrow$	$HNC + H_2$	3.50×10^{-9}	1.60
$H_2O + \gamma$	$\rightarrow$	$OH + H$	1.30×10^{-17}	485.5
$CH + \gamma$	$\rightarrow$	$C + H$	1.30×10^{-17}	365.0
$CH_2 + \gamma$	$\rightarrow$	$CH + H$	1.30×10^{-17}	250.0
$HCNO + \gamma$	$\rightarrow$	$CH + NO$	1.30×10^{-17}	1500.0
$HCSi + \gamma$	$\rightarrow$	$CH + Si$	1.30×10^{-17}	750.0
$CN + \gamma$	$\rightarrow$	$C + N$	1.30×10^{-17}	5290.0
$C_2H_5CN + \gamma$	$\rightarrow$	$CN + C_2H_5$	1.30×10^{-17}	2388.0
$OCN + \gamma$	$\rightarrow$	$CN + O$	1.30×10^{-17}	750.0
$SiNC + \gamma$	$\rightarrow$	$Si + CN$	1.30×10^{-17}	750.0
$NH + \gamma$	$\rightarrow$	$N + H$	1.30×10^{-17}	250.0
$CH_2NH + \gamma$	$\rightarrow$	$NH + CH_2$	1.30×10^{-17}	500.0
$HNCO + \gamma$	$\rightarrow$	$NH + CO$	1.30×10^{-17}	1500.0
$NH_2 + \gamma$	$\rightarrow$	$NH + H$	1.30×10^{-17}	40.0
Class:		Photoionisation		
Form:		$k = \alpha e^{-\Gamma A_v}$		
Reactants		Product	α	Γ
$H_2O + \gamma$	$\rightarrow$	$H_2O^+ + e^-$	3.10×10^{-11}	3.9
$CH + \gamma$	$\rightarrow$	$CH^+ + e^-$	7.60×10^{-10}	3.3
$CH^- + \gamma$	$\rightarrow$	$CH + e^-$	1.30×10^{-17}	250.0
$CN^- + \gamma$	$\rightarrow$	$CN + e^-$	2.22×10^{-9}	2.0
$NH + \gamma$	$\rightarrow$	$NH^+ + e^-$	1.30×10^{-17}	250.0

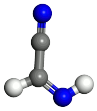
Class:		Radiative Association			
Form:		$k = \alpha\,(T/300)^{\beta}\,e^{-\Gamma/T}$			
Reactants		Product	α	β	Γ
$HCN + CH_3^+$	$\rightarrow$	$CH_3CNH^+ + \gamma$	9.0×10^{-9}	-0.5	0.0
$HCN + Si^+$	$\rightarrow$	$SiNCH^+ + \gamma$	6.0×10^{-15}	-1.5	0.0
$H_2O + HCO^+$	$\rightarrow$	$HCOOH_2^+ + \gamma$	4.0×10^{-13}	-1.3	0.0
$H_2O + CH_3^+$	$\rightarrow$	$CH_3OH_2^+ + \gamma$	2.0×10^{-12}	0.0	0.0
$H + OH$	$\rightarrow$	$H_2O + \gamma$	5.26×10^{-18}	-5.22	90.0
$CH + H_2$	$\rightarrow$	$CH_3 + \gamma$	5.09×10^{-18}	-0.71	11.6
$C + H$	$\rightarrow$	$CH + \gamma$	1.00×10^{-17}	0.00	0.0
$CN + C_2H$	$\rightarrow$	$HC_3N + \gamma$	1.00×10^{-16}	0.0	0.0
$CN + CH_3$	$\rightarrow$	$CH_3CN + \gamma$	1.00×10^{-16}	0.0	0.0
$C + N$	$\rightarrow$	$CN + \gamma$	5.72×10^{-19}	0.37	51.0
Class:		Associative detachment			
Form:		$k = \alpha\,(T/300)^{\beta}\,e^{-\Gamma/T}$			
Reactants		Product	α	β	Γ
$C^- + NH$	$\rightarrow$	$HCN + e^-$	5.00×10^{-10}	0.00	0.0
$H^- + CN$	$\rightarrow$	$HCN + e^-$	1.00×10^{-10}	0.00	0.0
$H_2O + C^-$	$\rightarrow$	$H_2CO + e^-$	5.00×10^{-10}	0.00	0.0
$H_2 + O^-$	$\rightarrow$	$H_2O + e^-$	7.00×10^{-10}	0.00	0.0
$H^- + OH$	$\rightarrow$	$H_2O + e^-$	1.00×10^{-10}	0.00	0.0
$H + OH^-$	$\rightarrow$	$H_2O + e^-$	1.40×10^{-9}	0.00	0.0
$CH + C^-$	$\rightarrow$	$C_2H + e^-$	5.00×10^{-10}	0.00	0.0
$CH + O^-$	$\rightarrow$	$HCO + e^-$	5.00×10^{-10}	0.00	0.0
$CH + O$	$\rightarrow$	$HCO^+ + e^-$	1.09×10^{-11}	-2.19	165.1
$CH + OH^-$	$\rightarrow$	$H_2CO + e^-$	5.00×10^{-19}	0.00	0.0
$H^- + C$	$\rightarrow$	$CH + e^-$	1.00×10^{-9}	0.00	0.0
$H + C^-$	$\rightarrow$	$CH + e^-$	5.00×10^{-10}	0.00	0.0
$C^- + N$	$\rightarrow$	$CN + e^-$	5.00×10^{-10}	0.00	0.0
$N + C_4H^-$	$\rightarrow$	$CN + C_3H + e^-$	3.00×10^{-12}	0.00	0.0
$NH + H^-$	$\rightarrow$	$NH_2 + e^-$	1.00×10^{-10}	0.00	0.0
$H^- + N$	$\rightarrow$	$NH + e^-$	1.00×10^{-9}	0.00	0.0

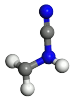
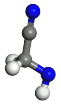
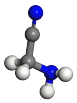
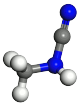
Class:		Charge transfer		
Form:		$k = \alpha\,(T/300)^{\beta}$		
Reactants		Product	α	β
$HCN + CN^+$	$\rightarrow$	$HCN^+ + CN$	1.79×10^{-9}	-0.5
$HCN + H_2^+$	$\rightarrow$	$HCN^+ + H_2$	2.70×10^{-9}	-0.5
$H_2O + HCN^+$	$\rightarrow$	$HCN + H_2O^+$	1.80×10^{-9}	-0.5
$C_2H_2 + HCN^+$	$\rightarrow$	$HCN + C_2H_2^+$	1.50×10^{-9}	0.0
$H_2O + H^+$	$\rightarrow$	$H_2O^+ + H$	6.90×10^{-9}	-0.5
$H_2O + H_2^+$	$\rightarrow$	$H_2O^+ + H_2$	3.90×10^{-9}	-0.5
$H_2O^+ + CH_2$	$\rightarrow$	$H_2O + CH_2^+$	4.70×10^{-10}	0.0
$H_2O^+ + C_2$	$\rightarrow$	$H_2O + C_2^+$	4.70×10^{-10}	0.0
$CH + CN^+$	$\rightarrow$	$CH^+ + CN$	6.40×10^{-10}	-0.5
$CH + OH^+$	$\rightarrow$	$CH^+ + OH$	3.50×10^{-10}	-0.5
$CH^+ + HCO$	$\rightarrow$	$CH + HCO^+$	4.60×10^{-10}	-0.5
$CH^+ + NH_3$	$\rightarrow$	$CH + NH_3^+$	4.59×10^{-10}	-0.5
$CN + N^+$	$\rightarrow$	$CN^+ + N$	1.10×10^{-9}	-0.5
$CN + O^-$	$\rightarrow$	$CN^- + O$	1.00×10^{-9}	0.0
$CN^+ + H$	$\rightarrow$	$CN + H^+$	6.40×10^{-10}	0.0
$CN^+ + NH$	$\rightarrow$	$CN + NH^+$	6.50×10^{-10}	-0.5
$NH + H^+$	$\rightarrow$	$NH^+ + H$	2.10×10^{-9}	-0.5
$NH + N^+$	$\rightarrow$	$NH^+ + N$	3.70×10^{-10}	-0.5
$NH^+ + H_2O$	$\rightarrow$	$NH + H_2O^+$	1.05×10^{-9}	-0.5
$NH^+ + O_2$	$\rightarrow$	$NH + O_2^+$	4.51×10^{-10}	0.0

Appendix B

Chemical Structures

Reaction number:	Reactants:	Transition state:
(9)	HCN + HCN	
(2)	HCN + H_2	
(7)	HCN + HCN	
(6)	HCN + H_2	
(12)	H_2CNH + CN	
(15)	H_2CNH + CN	

Reaction number:	Description:	Structure:
(10)	NCCHNH	
(3)	H_2CNH	
(8)	H_2CNCN	
(8a)	NCCHNH	

Reaction number:	Description:	Structure:
(17)	H_2CNHCN	
(13)	$NCCH_2NH$	
(13a)	NCH_2CNH	
(14)	$NCCH_2NH_2$	
(14a)	$NCCH_2NH_2$	
(19)	H_3CNHCN	
(18)	H_3CNCN	

Appendix C

Structure Coordinates

Tabulated are the coordinates (a, b, c) of Figure 5.4 on page 64. The first two lines of the table give the lengths of (a, b, c) and the opposing angles (A, B, C). All systems have the vertical axis perpendicular to the other coordinates and the coordinates in the plane at a 120° angle to each other. The total lengths of each of the axis is measured in angstroms with all atomic coordinates being in a fraction of the total axis-length. The *.sci* files of selected structures are imbedded in the captions of their associated figures, which also contain the coordinate data.

104.672	15.64	15.64	(Å)
120	90	90	(angle)
0.537407134944	0.0310009711806	0.239045784577	H
0.501106975048	0.694661114313	0.201212647718	H
0.534864150434	0.0870460458936	0.40021553926	H
0.492705290129	0.853219195111	0.25047798261	H
0.537280168228	0.340404814732	0.431097958673	H
0.507180106728	0.8113663729	0.108274357308	H
0.503896079644	0.0170767772737	0.159942527788	H
0.538042094635	0.729331787783	0.146860874347	H
0.499251287174	0.906186701573	0.206700934272	H
0.538206725702	0.974229831233	0.350567806828	H
0.501792882483	0.310923125452	0.451252727364	H
0.536136718876	0.9838152014	0.123099283331	H
0.569680601405	0.839217428042	0.982524428505	O
0.5333611458	0.0423523466187	0.185762156724	O
0.56480464249	0.139259369593	0.0393335125491	O
0.529921926875	0.777077950027	0.0796242110919	O
0.566103107002	0.000483300109789	0.29407793107	O
0.534197819184	0.346253621677	0.331386434608	O
0.471169106906	0.73352261108	0.937878610627	H

0.461529609792	0.896440207701	0.154641472293	H
0.541265406425	0.12880437736	0.174552529107	H
0.533656815584	0.814062973819	0.0479858990296	H
0.556846518772	0.0141053946563	0.303520142269	H
0.518090233412	0.313031467851	0.346066517263	H
0.542836139484	0.879728564046	0.0166617256192	O
0.509833441818	0.975135743572	0.166831644202	O
0.547194708886	0.186961472388	0.183461094762	O
0.519781896594	0.729445902922	0.00998598104686	O
0.539790423015	0.0319449919586	0.343558173042	O
0.508977038056	0.310574234561	0.363602128366	O
0.436006903755	0.714524850028	0.977752532165	H
0.432503095473	0.671797299554	0.546746561947	H
0.430124460917	0.278379941293	0.277958048287	H
0.551947283672	0.864081966393	0.99560766452	H
0.569168945666	0.0169334326864	0.246119443509	H
0.540005923982	0.293906116796	0.280871416384	H
0.455212620745	0.908370088921	0.174738768627	H
0.553897377769	0.507858288359	0.440996692639	H
0.474427921667	0.0762389820522	0.416951485816	H
0.4853489253	0.161808501646	0.650723200505	H
0.452953495498	0.474912322507	0.940269221505	H
0.432325652499	0.907124892231	0.0269593565488	H
0.566763455118	0.0954006308838	0.967389461646	H
0.470984818059	0.663214336449	0.077706623408	H
0.432529615696	0.0814279281499	0.400226164007	H
0.55267149991	0.559888470849	0.479698646364	H
0.569057531909	0.322446614278	0.433501795197	H
0.431702832139	0.904315910802	0.331376483353	H
0.50771876175	0.742746924096	0.0711069434511	O
0.472352207675	0.524306713015	0.425539001163	O
0.494170597315	0.143029651949	0.197120708929	O
0.469165474935	0.703460284165	0.979442568523	O
0.490791893247	0.886677292854	0.211663450597	O
0.463507250542	0.335155682003	0.351412331169	O
0.475416785938	0.89013252263	0.968254336861	H
0.444037141537	0.0274238996048	0.114088809164	H
0.485192574994	0.122163462837	0.197163543897	H
0.459486609593	0.699041256823	0.986372490056	H
0.481486882719	0.961304418769	0.300376261825	H
0.454320349475	0.356050390624	0.351061281726	H
0.476077400523	0.824470668952	0.909913977662	O
0.435164111369	0.0205076755351	0.0912608789861	O

0.466302456225	0.118350518231	0.184112649337	O
0.442388043121	0.662873178012	0.95769999081	O
0.474556479095	0.995000331738	0.348746261568	O
0.432926120901	0.345659496257	0.323457351072	O
0.487110073858	0.833393998547	0.666642636789	H
0.46723280665	0.954818543064	0.356246063607	H
0.505345169294	0.255147028446	0.306350266155	H
0.485327282512	0.81130055531	0.900261537515	H
0.518504727478	0.0014738337659	0.170327675919	H
0.459762330557	0.165627195724	0.225361751556	H
0.516205481838	0.140373250511	0.480905092821	H
0.51111669557	0.149981400399	0.61525868914	H
0.537005279013	0.490059524748	0.769730424157	H
0.505373263407	0.386058630364	0.847275984681	H
0.539620139382	0.883315448401	0.959726993782	H
0.504744185878	0.413723143753	0.607672348807	H
0.503931329696	0.528465539975	0.644328982133	H
0.534476551112	0.178543986127	0.547276075426	H
0.503347138319	0.560069394297	0.88755217851	H
0.538184149883	0.461235933833	0.852456459884	H
0.51357770256	0.83571358344	0.891786457273	H
0.535432388336	0.544228256583	0.669603503159	H
0.566545275794	0.319387113643	0.495165897003	O
0.533348114369	0.470752473135	0.65059438201	O
0.563412310281	0.672529893317	0.640742460839	O
0.530419816804	0.182612316807	0.49194494906	O
0.565457706365	0.516229239843	0.829385119021	O
0.53235136519	0.873745481186	0.851798392827	O
0.55707762827	0.321113950692	0.489561488287	H
0.523325813036	0.469225328828	0.650358593947	H
0.488858351242	0.552280385537	0.564950585683	H
0.467563997534	0.333573399184	0.453993248445	H
0.556101704704	0.513817049533	0.841603964254	H
0.537274711079	0.082980817866	0.916236049576	H
0.539201384739	0.324504414781	0.486332974038	O
0.508831062552	0.480192706751	0.658241953848	O
0.537052348139	0.648732783369	0.661834289057	O
0.508196167592	0.0990672995092	0.47513912338	O
0.539967939014	0.516667861314	0.841852563954	O
0.505852773703	0.792167300601	0.869923945941	O
0.448912959272	0.293213830251	0.661115098432	H
0.537095202908	0.408705315865	0.568667933539	H
0.545914098736	0.66295693214	0.658315354418	H

0.532896992642	0.255673034752	0.481961468506	H
0.566604092652	0.517868344196	0.765720084685	H
0.537902434457	0.82619948633	0.80351056794	H
0.464058214622	0.610092100126	0.679367227482	H
0.565500666674	0.134414723432	0.573112662341	H
0.472319732909	0.58244721413	0.877840302432	H
0.568396086741	0.406791036961	0.825789852172	H
0.448724346776	0.805676518967	0.785853772544	H
0.431499381327	0.417239007409	0.580220180227	H
0.43392218048	0.524762246251	0.610978859954	H
0.472822532049	0.226129250098	0.507050798198	H
0.4376558305	0.608909833913	0.910920201231	H
0.472495075006	0.415995323289	0.851746664816	H
0.548100730439	0.838829591204	0.634490508484	H
0.432703148716	0.820357692856	0.920756010733	H
0.497448044485	0.312330285859	0.515119069136	O
0.461508193046	0.493134294044	0.681732600142	O
0.494511864411	0.6013036895	0.616944018141	O
0.471371517308	0.155007972912	0.490965851436	O
0.499318878963	0.502744370928	0.829706780316	O
0.441572466011	0.843719293552	0.784264285011	O
0.488460440846	0.332195862858	0.519404427501	H
0.452793351798	0.476258280312	0.684536477923	H
0.487097006607	0.646736853634	0.656140211643	H
0.470529678593	0.162724444668	0.558311221685	H
0.484461592316	0.520120395221	0.845994134855	H
0.570190515496	0.787719981097	0.752802553251	H
0.472189894211	0.339596697176	0.5114496661	O
0.433659484275	0.482162409351	0.63999189693	O
0.46855840123	0.671831671404	0.685368786175	O
0.437556751531	0.122739753537	0.456884641062	O
0.474979597055	0.516333444499	0.831693232433	O
0.561590375242	0.754455049623	0.735835269093	O
0.500649076521	0.137217805759	0.488144259744	H
0.466274354927	0.489571971221	0.737186034664	H
0.505601818842	0.772414357017	0.92059345538	H
0.483994485909	0.154450441175	0.879183898701	H
0.504503201283	0.493341863298	0.76804793028	H
0.468360848633	0.399306309104	0.579103420627	H
0.536226589567	0.586231717317	0.152339567108	H
0.493760466068	0.163648100907	0.147317436361	H
0.536134704867	0.5706512086	0.368415511908	H
0.50634374893	0.502133599641	0.230574616412	H

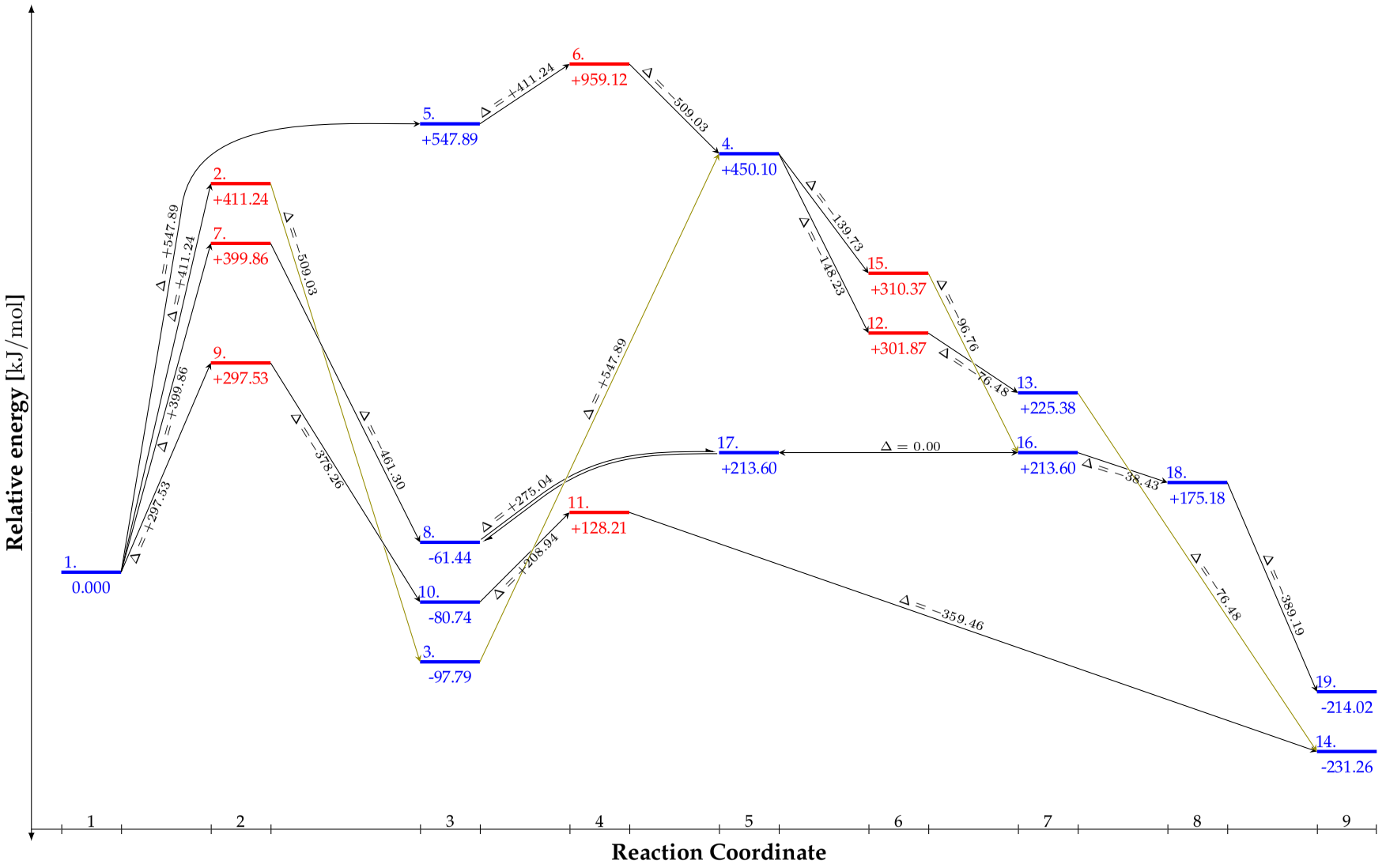
0.537225722226	0.869699288403	0.447056672	H
0.497596104741	0.825663162715	0.443127508077	H
0.4976457937	0.622523922418	0.234946539869	H
0.535060750929	0.209563514932	0.0957677648559	H
0.503790725605	0.598483454635	0.403433109078	H
0.53540418911	0.409337261379	0.332537456575	H
0.482684729749	0.790933466461	0.63868733355	H
0.539641793024	0.414117281756	0.0565886775433	H
0.567602187734	0.296784777661	0.974019421668	O
0.533003974616	0.520095557102	0.13823783874	O
0.565180395661	0.649565574011	0.201953542554	O
0.530056120476	0.178008275241	0.0366812580965	O
0.560920484391	0.434324224314	0.232837217133	O
0.532864581264	0.861769239646	0.388631168461	O
0.558550947321	0.320351163844	0.993361532835	H
0.504113044697	0.452724058589	0.0594263014536	H
0.551352282329	0.688979348995	0.178270523889	H
0.520407031325	0.193404641789	0.0495560553196	H
0.545865054583	0.493425136799	0.284008111506	H
0.504307496769	0.973362172947	0.393802520488	H
0.543526636678	0.346639860214	0.0146979389796	O
0.50684404931	0.504674376834	0.127100456698	O
0.54201543814	0.705025225062	0.190418391754	O
0.505402906247	0.204505229792	0.0518795445147	O
0.536776269356	0.508866021093	0.301203550503	O
0.505038124383	0.89637010256	0.359478855754	O
0.441626166528	0.296360612023	0.067168464968	H
0.534261009835	0.525521561475	0.204215085197	H
0.431858570484	0.698556628787	0.243672201491	H
0.537337638473	0.300994513346	0.0169325167147	H
0.567308231615	0.448972268095	0.193942307564	H
0.535922359476	0.797352055365	0.332347196211	H
0.431882285979	0.0800133511831	0.0899742744341	H
0.430646001001	0.190419157357	0.0680000181457	H
0.467872261817	0.605805915302	0.455836446106	H
0.569977263222	0.392293014175	0.299843295381	H
0.461604109483	0.858806356579	0.430754428773	H
0.436295825964	0.422964409835	0.0726126741798	H
0.568831697313	0.579111210483	0.162531699467	H
0.455934036255	0.114369046761	0.958241929181	H
0.557734862689	0.37189902969	0.176946803836	H
0.46877483134	0.366784194191	0.324595238212	H
0.430956457866	0.872472689467	0.409038375313	H

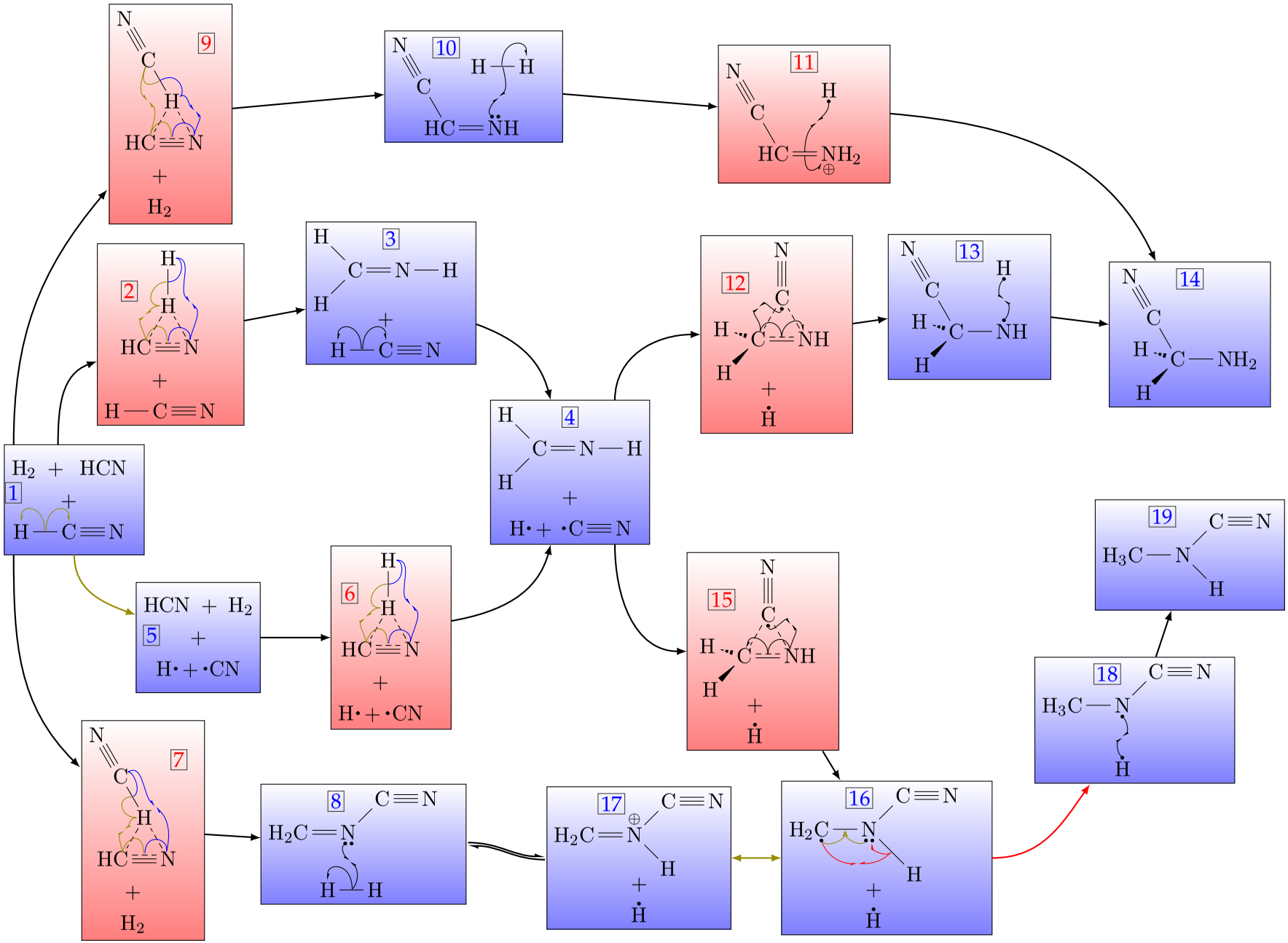
0.463137523626	0.390276143904	0.0592513711155	H
0.492783396248	0.306253413629	0.994485830028	O
0.46060286355	0.524252584898	0.157706771934	O
0.495493469007	0.69126836878	0.248412795821	O
0.462500755283	0.153792856729	0.00260967483235	O
0.505018435647	0.494945996043	0.289172041452	O
0.459567601047	0.861553764474	0.364600307294	O
0.483501808586	0.319548245291	0.00126805103631	H
0.445660266116	0.496707810538	0.148813639597	H
0.480284914863	0.67077892673	0.171400221023	H
0.450427846974	0.223666287884	0.0682176714668	H
0.497170395587	0.453923323079	0.271089476883	H
0.450162343365	0.867853381632	0.364074179069	H
0.46755060273	0.332014078274	0.0102742586045	O
0.436540538005	0.475758773549	0.139965374072	O
0.474203928569	0.627201319547	0.112041469955	O
0.4428296237	0.264070844125	0.10808898313	O
0.477077923184	0.497526028574	0.337344942294	O
0.434091734966	0.847203760953	0.341624378229	O
0.50230827706	0.25778058241	0.0329432822499	H
0.465724282821	0.53940980564	0.215245624683	H
0.513629341441	0.871450324152	0.334120841954	H
0.46403241595	0.26981765931	0.00553097734315	H
0.516714207746	0.510461288367	0.120861326292	H
0.465533920157	0.560461951092	0.126988064005	H
0.534902213167	0.0958686744298	0.656517887557	H
0.502941000757	0.631366597356	0.520916921266	H
0.551562155126	0.161644211193	0.11875660942	H
0.502067590204	0.130791096341	0.00921146848252	H
0.536053807927	0.217117342348	0.707809523926	H
0.50422560354	0.920234435892	0.620006924462	H
0.503020509655	0.117005225746	0.676423136401	H
0.533867081437	0.644074117789	0.598966210887	H
0.497042029176	0.913193463381	0.835597542576	H
0.534069765837	0.932067764127	0.846462303515	H
0.495766608266	0.322401315378	0.941138894157	H
0.53948438562	0.908768230698	0.615013318271	H
0.431220331287	0.854645549545	0.534625174001	O
0.530734266564	0.00370207119586	0.674346878383	O
0.563725836339	0.173143677269	0.646414812937	O
0.531144888117	0.653575693043	0.472907205814	O
0.568704672291	0.0800723245464	0.843008151668	O
0.541960653711	0.325682525702	0.823273164433	O

0.562648447413	0.858114734745	0.54591738692	H
0.521021259968	0.998460936883	0.670281022515	H
0.554139242942	0.165256887737	0.646306406948	H
0.521292146592	0.659438883298	0.470883900229	H
0.55040161423	0.0690890418658	0.856681462046	H
0.450819299133	0.524186389743	0.984208579428	H
0.546058931807	0.849488848309	0.578563964117	O
0.506741890307	0.980759492429	0.684060359298	O
0.536866203162	0.152456395191	0.646967596855	O
0.50471520076	0.656343259108	0.472416422521	O
0.5412856424	0.0461978490862	0.8495495484	O
0.508483867431	0.341357060201	0.869120488055	O
0.434955770154	0.791898722316	0.516402732189	H
0.532073712225	0.033495574608	0.743681811423	H
0.565605141453	0.28319999346	0.756920831041	H
0.534054035776	0.721508081709	0.496509331253	H
0.433196444065	0.03471166168	0.78272563868	H
0.539151055584	0.319428394867	0.880315658858	H
0.470451997755	0.0560192507507	0.679393492929	H
0.431030042898	0.615721228857	0.432103819231	H
0.471794018789	0.133367531699	0.94737444944	H
0.465255977844	0.110555304243	0.115797815918	H
0.469124096831	0.347872909818	0.902359483866	H
0.440905176202	0.914573115827	0.636006610088	H
0.434925465647	0.188862449949	0.489510994801	H
0.465295462605	0.673306980264	0.627414928932	H
0.570051221856	0.147274478543	0.857790110959	H
0.445109725429	0.890318913794	0.759217263307	H
0.430022538218	0.309102115064	0.920570997037	H
0.469412790161	0.903236666053	0.604083644626	H
0.494203247544	0.820894373788	0.5011655626	O
0.468519529929	0.0195793307185	0.714628254459	O
0.505491924143	0.176487814816	0.677902302214	O
0.461669126109	0.682061488799	0.508838867399	O
0.498411026855	0.977667508284	0.842436235452	O
0.470284907587	0.346797645995	0.836834916258	O
0.479812764523	0.841220979866	0.520107003725	H
0.45433248616	0.975779356233	0.697823687366	H
0.449077008134	0.312156470274	0.717453904111	H
0.452426615213	0.667396384021	0.493885033193	H
0.490194848098	0.00961353284193	0.843394966658	H
0.551525185245	0.330664217185	0.836665913661	H
0.470007967566	0.859697380318	0.532485088651	O

0.44793962809	0.925329567546	0.680406089412	O
0.476193332315	0.168732458976	0.663390172198	O
0.436527999459	0.664033023356	0.484475134963	O
0.479697143502	0.102834762274	0.881113887872	O
0.567656733594	0.344216952965	0.826667492727	O
0.498124446263	0.760294065067	0.498354211914	H
0.472986957472	0.0509948369154	0.785465529013	H
0.507367171675	0.272464974187	0.80680698264	H
0.464713392537	0.750732402347	0.510008565631	H
0.505245705501	0.963067584594	0.737338617604	H
0.47448218951	0.235668433037	0.724914419991	H

Table C.1: Atomic coordinates of Figure 5.4.


Relative energy [kJ/mol]
Reaction Coordinate
1
2
3
4
5
6
7
8
9
1.
0.000
2.
+411.24
3.
-97.79
4.
+450.10
5.
+547.89
6.
+959.12
7.
+399.86
8.
-61.44
9.
+297.53
10.
-80.74
11.
+128.21
12.
+301.87
13.
+225.38
14.
-231.26
15.
+310.37
16.
+213.60
17.
+213.60
18.
+175.18
19.
-214.02
Δ = +547.89
Δ = +411.24
Δ = +399.86
Δ = +297.53
Δ = +411.24
Δ = −509.03
Δ = −509.03
Δ = −139.73
Δ = −148.23
Δ = −96.76
Δ = −76.48
Δ = −38.43
Δ = 0.00
Δ = −461.30
Δ = −378.26
Δ = +275.04
Δ = +208.94
Δ = +547.89
Δ = −359.46
Δ = −76.48
Δ = −389.19


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
H2 + HCN
HCN + H2
H2
+

References

Adamo, C., & Barone, V. 1999, The Journal of chemical physics, 110, 6158

Allamandola, L., Tielens, A., & Barker, J. 1989, The Astrophysical Journal Supplement Series, 71, 733

Andre, P., Ward-Thompson, D., & Barsony, M. 1993, The Astrophysical Journal, 406, 122

Asgeirsson, V. 2015, Master's thesis, University of Iceland

Asnin, L., Chekryshkin, Y. S., & Fedorov, A. 2003, Russian chemical bulletin, 52, 2747

Atkins, P. W., & Friedman, R. S. 2011, Molecular quantum mechanics (Oxford university press)

Auletta, G., Fortunato, M., & Parisi, G. 2009, Quantum Mechanics (Cambridge University Press), 513–540

Awad, Z., Viti, S., Bayet, E., & Caselli, P. 2014, Monthly Notices of the Royal Astronomical Society, 443, 275

Baerends, E. J., & Gritsenko, O. V. 1997, The Journal of Physical Chemistry A, 101, 5383

Barlow, M., & Silk, J. 1976, The Astrophysical Journal, 207, 131

Basiuk, V., & Bogillo, V. 2002, Advances in Space Research, 30, 1445

Bauschlicher, C. W. 1995, Chemical Physics Letters, 246, 40

Becke, A. D. 1988, Physical review A, 38, 3098

—. 1993a, The Journal of chemical physics, 98, 5648

—. 1993b, The Journal of chemical physics, 98, 1372

Belloche, A., Garrod, R., Müller, H., et al. 2009, Astronomy & Astrophysics, 499, 215

Belloche, A., Menten, K., Comito, C., et al. 2008, Astronomy & Astrophysics, 482, 179

Bernstein, M., Mattioda, A., Sandford, S., & Hudgins, D. 2005, The Astrophysical Journal, 626, 909

Biham, O., Furman, I., Katz, N., Pirronello, V., & Vidali, G. 1998, Monthly Notices of the Royal Astronomical Society, 296, 869

Blanksby, S. J., & Ellison, G. B. 2003, Accounts of chemical research, 36, 255

Blöchl, P. E. 1994, Physical Review B, 50, 17953

—. 2011, arXiv preprint arXiv:1108.1104

Borget, F., Danger, G., Duvernay, F., et al. 2012, Astronomy & Astrophysics, 541, A114

Born, M., & Oppenheimer, R. 1927, Annalen der Physik, 389, 457

Brasil, C. A., Fanchini, F. F., & Napolitano, R. d. J. 2013, Revista Brasileira de Ensino de Física, 35, 01

Brillouin, L. 1934, Les champs" self-consistents" de Hartree et de Fock No. 159 (Paris)

Brown, P. D. 1990, Monthly Notices of the Royal Astronomical Society, 243, 65

Buch, V., Groenzin, H., Li, I., Shultz, M., & Tosatti, E. 2008, Proceedings of the National Academy of Sciences, 105, 5969

Burke, K. 2012, The Journal of chemical physics, 136, 150901

Capriotti, E. R., & Kozminski, J. F. 2001, Publications of the Astronomical Society of the Pacific, 113, 677

Cazaux, S., & Tielens, A. 2002, The Astrophysical Journal Letters, 575, L29

—. 2004, The Astrophysical Journal, 604, 222

Ceccarelli, C., Castets, A., Loinard, L., Caux, E., & Tielens, A. 1998, Astronomy & Astrophysics, 338, L43

Ceperley, D. M., & Alder, B. 1980, Physical Review Letters, 45, 566

Chakrabarti, S., & Chakrabarti, S. K. 2000, Astronomy & Astrophysics, 354, L6

Chang, Q., Cuppen, H., & Herbst, E. 2005, Astronomy & Astrophysics, 434, 599

Chen, F. F., & Smith, M. D. 1984, Plasma (Wiley Online Library)

Cohen, A. J., Mori-Sánchez, P., & Yang, W. 2008, Science, 321, 792

Cook, D. B. 2012, Handbook of computational quantum chemistry (Courier Corporation)

Cramer, C. J. 2013, Essentials of computational chemistry: theories and models (John Wiley & Sons)

Cramer, C. J., & Truhlar, D. G. 2009, Physical Chemistry Chemical Physics, 11, 10757

Cuppen, H., & Herbst, E. 2005, Monthly Notices of the Royal Astronomical Society, 361, 565

D'Agosta, R., & Di Ventra, M. 2008, Physical Review B, 78, 165105

—. 2013, Physical Review B, 87, 155129

De Becker, M. 2013, Bulletin de la Société Royale des Sciences de Liége, 82, 33

Delley, B. 2006, The Journal of Physical Chemistry A, 110, 13632

d'Hendecourt, L., Allamandola, L., & Greenberg, J. 1985, Astronomy & Astrophysics, 152, 130

Di Ventra, M., & D'Agosta, R. 2007, Physical review letters, 98, 226403

Dirac, P. A. 1929, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 123, 714

Du, F., Parise, B., & Bergman, P. 2012, Astronomy & Astrophysics, 538, A91

Duck, I., & Sudarshan, E. C. G. 1998, American Journal of Physics, 66, 284

Echenique, P., & Alonso, J. L. 2007, Molecular Physics, 105, 3057

Eddington, A. 1937, The Observatory, 60, 99

Eichler, A. 2008, Lecture-Universit" at Wien-Accessed, 19

Eley, D. D. 1941, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 178, 452

Ertl, G. 2010, Reactions at solid surfaces, Vol. 14 (John Wiley & Sons)

Farebrother, A. J., Meijer, A. J., Clary, D. C., & Fisher, A. J. 2000, Chemical Physics Letters, 319, 303

Folsome, C., Lawless, J., Romiez, M., & Ponnaamperuma, C. 1973, Geochimica et Cosmochimica Acta, 37, 455

Folsome, C., Lawless, J., Romiez, M., & Ponnamperuma, C. 1971, Nature, 232, 108

Freyer, T., Hensler, G., & Yorke, H. W. 2003, The Astrophysical Journal, 594, 888

Friedel, D., Snyder, L., Turner, B., & Remijan, A. 2004, The Astrophysical Journal, 600, 234

Fromherz, T., Mendoza, C., & Ruette, F. 1993, Monthly Notices of the Royal Astronomical Society, 263, 851

Gasiorowicz, S. 2007, Quantum physics (John Wiley & Sons)

Glaser, R., Hodgen, B., Farrelly, D., & McKee, E. 2007, Astrobiology, 7, 455

Goerigk, L., & Grimme, S. 2011, Physical Chemistry Chemical Physics, 13, 6670

Gould, R., & Salpeter, E. 1963, The Astrophysical Journal, 138, 393

Gunnarsson, O., & Lundqvist, B. I. 1976, Physical Review B, 13, 4274

Gupta, V., Tandon, P., & Mishra, P. 2013, Advances in Space Research, 51, 797

Gupta, V., Tandon, P., Rawat, P., Singh, R., & Singh, A. 2011, Astronomy & Astrophysics, 528, A129

Güttler, A., Zecho, T., & Küppers, J. 2004, Carbon, 42, 337

Harris, J., & Kasemo, B. 1981, Surface Science, 105, L281

Hartree, D., & Hartree, W. 1935, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 150, 9

Hartree, D. R. 1928, Mathematical Proceedings of the Cambridge Philosophical Society, 24, 89

Hasegawa, T. I., Herbst, E., & Leung, C. M. 1992, The Astrophysical Journal Supplement Series, 82, 167

Hayatsu, R., Studier, M. H., Moore, L. P., & Anders, E. 1975, Geochimica et Cosmochimica Acta, 39, 471

Heikkilä, A., & Lundell, J. 2000, The Journal of Physical Chemistry A, 104, 6637

Heitler, W., & London, F. 1927, Zeitschrift für Physik, 44, 455

Hinshelwood, C. 1930, Annual Reports on the Progress of Chemistry, 27, 11

Hobbs, D., Kresse, G., & Hafner, J. 2000, Physical Review B, 62, 11556

Hohenberg, P., & Kohn, W. 1964, Physical review, 136, B864

Hollenbach, D., & Salpeter, E. 1970, The Journal of Chemical Physics, 53, 79

Hornekær, L., Baurichter, A., Petrunin, V., Field, D., & Luntz, A. 2003, Science, 302, 1943

Hornekær, L., Baurichter, A., Petrunin, V. V., et al. 2005, The Journal of chemical physics, 122, 124701

James, H. M., & Coolidge, A. S. 1933, The Journal of Chemical Physics, 1, 825

Jobst, K. J., Hanifa, M. R., Ruttink, P. J., & Terlouw, J. K. 2009, Chemical Physics Letters, 473, 257

Jobst, K. J., Hanifa, M. R., & Terlouw, J. K. 2008, Chemical Physics Letters, 462, 152

Jobst, K. J., & Terlouw, J. K. 2010, Chemical Physics Letters, 497, 7

Jones, R. O., & Gunnarsson, O. 1989, Reviews of Modern Physics, 61, 689

Jung, S. H., & Choe, J. C. 2013, Astrobiology, 13, 465

Kalvāns, J. 2015, The Astrophysical Journal, 806, 196

Katz, N., Furman, I., Biham, O., Pirronello, V., & Vidali, G. 1999, The Astrophysical Journal, 522, 305

Kemble, E., & Zener, C. 1929, Physical Review, 33, 512

Knaap, H., Van den Meijdenberg, C., Beenakker, J., Van de Hulst, H., et al. 1966, Bulletin of the Astronomical Institutes of the Netherlands, 18, 256

Knowles, T., & Suhl, H. 1977, Physical Review Letters, 39, 1417

Koch, W., & Holthausen, M. C. 2015, A chemist's guide to density functional theory (John Wiley & Sons)

Kohn, W., Becke, A. D., & Parr, R. G. 1996, The Journal of Physical Chemistry, 100, 12974

Kohn, W., & Sham, L. J. 1965, Physical review, 140, A1133

Korth, M., & Grimme, S. 2009, Journal of chemical theory and computation, 5, 993

Kuan, Y.-J., Yan, C.-H., Charnley, S. B., et al. 2003, Monthly Notices of the Royal Astronomical Society, 345, 650

Kurth, S., Perdew, J. P., & Blaha, P. 1999, International journal of quantum chemistry, 75, 889

Laidler, K. J. 1984, Journal of Chemical Education, 61, 494

Langmuir, I. 1922, Transactions of the Faraday Society, 17, 621

Lattelais, M., Ellinger, Y., & Zanda, B. 2007, International Journal of Astrobiology, 6, 37

Le Bellac, M. 2011, Quantum physics (Cambridge University Press), 507–543

Leitch-Devlin, M., & Williams, D. 1985, Monthly Notices of the Royal Astronomical Society, 213, 295

Levy, M., & Perdew, J. P. 1993, Physical Review B, 48, 11638

Lewars, E. G. 2010, Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics (Springer Science & Business Media)

Li, A., & Draine, B. 2001, The Astrophysical Journal, 554, 778

Lippok, N., Launhardt, R., Semenov, D., et al. 2013, Astronomy & Astrophysics, 560, A41

Lowe, J. P., & Peterson, K. 2011, Quantum chemistry (Academic Press)

Manico, G., Raguni, G., Pirronello, V., Roser, J., & Vidali, G. 2001, The Astrophysical Journal Letters, 548, L253

Maret, S., Bergin, E., & Tafalla, M. 2013, Astronomy & Astrophysics, 559, A53

Martins, Z., Botta, O., Fogel, M. L., et al. 2008, Earth and planetary science Letters, 270, 130

Masel, R. I. 1996, Principles of adsorption and reaction on solid surfaces, Vol. 3 (John Wiley & Sons)

Matthews, C. N., & Minard, R. D. 2006, Faraday discussions, 133, 393

McElroy, D., Walsh, C., Markwick, A., et al. 2013, Astronomy & Astrophysics, 550, A36

Miller, S. L., & Urey, H. C. 1959, Science, 130, 245

Mortensen, J. J., Hansen, L. B., & Jacobsen, K. W. 2005, Physical Review B, 71, 035109

Mostofi, A. A., Skylaris, C.-K., Haynes, P. D., & Payne, M. C. 2002, Computer physics communications, 147, 788

Müller, H., Thorwirth, S., Roth, D., & Winnewisser, G. 2001, Astronomy & Astrophysics, 370, L49, [Last date of access: 29 October 2018]

Müller, H. S., Schlöder, F., Stutzki, J., & Winnewisser, G. 2005, Journal of Molecular Structure, 742, 215, [Last date of access: 29 October 2018]

Navarro-Ruiz, J., Sodupe, M., Ugliengo, P., & Rimola, A. 2014, Physical Chemistry Chemical Physics, 16, 17447

Nomura, H., & Millar, T. 2004, Astronomy & Astrophysics, 414, 409

Nuevo, M., Milam, S. N., Sandford, S. A., Elsila, J. E., & Dworkin, J. P. 2009, Astrobiology, 9, 683

Nummelin, A., Bergman, P., Hjalmarson, Å., et al. 1998, The Astrophysical Journal Supplement Series, 117, 427

Osorio, M., Lizano, S., & D'Alessio, P. 1999, The Astrophysical Journal, 525, 808

Parneix, P., & Bréchignac, P. 1998, Astronomy & Astrophysics, 334, 363

Perdew, J. P., Burke, K., & Ernzerhof, M. 1996, Physical review letters, 77, 3865

Perdew, J. P., Ruzsinszky, A., Tao, J., et al. 2005, The Journal of chemical physics, 123, 062201

Perdew, J. P., & Wang, Y. 1992, Physical Review B, 45, 13244

Perdew, J. P., Ziesche, P., & Eschrig, H. 1991, Electronic structure of solids' 91, Vol. 11 (Akademie Verlag, Berlin)

Perets, H. B., Biham, O., Manicó, G., et al. 2005, The Astrophysical Journal, 627, 850

Perry, J. S., & Price, S. D. 2003, Astrophysics and space science, 285, 769

Petucci, J., LeBlond, C., Karimi, M., & Vidali, G. 2013, The Journal of chemical physics, 139, 044706

Pirronello, V., Biham, O., Liu, C., Shen, L., & Vidali, G. 1997a, The Astrophysical Journal Letters, 483, L131

Pirronello, V., Liu, C., Roser, J. E., & Vidali, G. 1999, Astronomy & Astrophysics, 344, 681

Pirronello, V., Liu, C., Shen, L., & Vidali, G. 1997b, The Astrophysical Journal Letters, 475, L69

Plesničar, B. 2005, Acta Chim. Slov, 52, 1

Ponnamperuma, C., Lemmon, R. M., Mariner, R., & Calvin, M. 1963, Proceedings of the National Academy of Sciences, 49, 737

Prasad, S., & Huntress Jr, W. 1980a, The Astrophysical Journal Supplement Series, 43, 1

—. 1980b, The Astrophysical Journal, 239, 151

Puget, J., & Léger, A. 1989, Annual review of astronomy & astrophysics, 27, 161

Rangel, T., Caliste, D., Genovese, L., & Torrent, M. 2016, Computer Physics Communications, 208, 1

Ratsch, C., & Scheffler, M. 1998, Physical Review B, 58, 13163

Ricca, A., Bauschlicher, C. W., & Rosi, M. 2001, Chemical physics letters, 347, 473

Riley, K. E., Op't Holt, B. T., & Merz, K. M. 2007, Journal of chemical theory and computation, 3, 407

Rosen, N. 1931, Physical Review, 38, 2099

Roser, J., Manico, G., Pirronello, V., & Vidali, G. 2002, The Astrophysical Journal, 581, 276

Roser, J. E., Vidali, G., Manicò, G., & Pirronello, V. 2001, The Astrophysical Journal Letters, 555, L61

Rostgaard, C. 2009, arXiv preprint arXiv:0910.1921

Roy, D., Najafian, K., & von Ragué Schleyer, P. 2007, Proceedings of the National Academy of Sciences, 104, 17272

Sandford, S., Bernstein, M., & Allamandola, L. 2004, The Astrophysical Journal, 607, 346

Sandford, S. A., & Nuevo, M. 2014

Savee, J. D., Mann, J. E., & Continetti, R. E. 2009, The Journal of Physical Chemistry A, 113, 8834

Schneider, P., & Smith, J. 1968, American Institute of Chemical Engineers Journal, 14, 762

Scuseria, G. E., & Staroverov, V. N. 2005, Theory and applications of computational chemistry: the first, 40, 669

Slater, J. C. 1930, Physical Review, 35, 210

Smith, I., Talbi, D., & Herbst, E. 2001, Astronomy & Astrophysics, 369, 611

Sousa, S. F., Fernandes, P. A., & Ramos, M. J. 2007, The Journal of Physical Chemistry A, 111, 10439

Stephens, P., Devlin, F., Chabalowski, C., & Frisch, M. 1994, Chem. Phys. J. Phys. Chem, 98, 11623

Streater, R. F., & Wightman, A. S. 2000, PCT, spin and statistics, and all that (Princeton University Press)

Strömgren, B. 1939, The Astrophysical Journal, 89, 526

Szabo, A., & Ostlund, N. S. 1996, Modern quantum chemistry (Dover)

Tielens, A., & Hagen, W. 1982, Astronomy & Astrophysics, 114, 245

Van der Velden, W., & Schwartz, A. W. 1977, Geochimica et Cosmochimica Acta, 41, 961

Vazart, F., Latouche, C., Skouteris, D., Balucani, N., & Barone, V. 2015, The Astrophysical Journal, 810, 111

Vidali, G., Pirronello, V., Liu, C., & Shen, L. 1998, Astrophysical Letters and Communications, 35, 423

von Barth, U. 2004, Physica Scripta, 2004, 9

Wang, S. C. 1928, Physical Review, 31, 579

Wang, Y., & Perdew, J. P. 1991, Physical Review B, 43, 8911

Wang, Z., Zhang, J., Wu, J., & Cao, W. 2007, Journal of Molecular Structure: THEOCHEM, 806, 239

Watanabe, N., & Kouchi, A. 2008, Progress in surface Science, 83, 439

Watson, W., & Salpeter, E. E. 1972, The Astrophysical Journal, 174, 321

Woon, D. E. 2002, The Astrophysical Journal Letters, 571, L177

Xiong, X.-G., & Yanai, T. 2017, Journal of Chemical Theory and Computation

Yim, M. K., & Choe, J. C. 2012, Chemical Physics Letters, 538, 24

Zaleski, D. P., Seifert, N. A., Steber, A. L., et al. 2013, The Astrophysical Journal Letters, 765, L10

Zangwill, A. 1988, Physics at surfaces (Cambridge university press)

Zecho, T., Güttler, A., Sha, X., Jackson, B., & Küppers, J. 2002, The Journal of chemical physics, 117, 8486

Zhao, Y., & Truhlar, D. G. 2007, Journal of chemical theory and computation, 3, 289