Universidad de Concepción
Facultad de Ciencias Químicas
Doctorado en Ciencias con Mención en Química

QUÍMICA SOBRE MODELOS DE HIELO INTERESTELAR: DISTRIBUCIONES DE ENERGÍAS DE UNIÓN Y SÍNTESIS DE MOLÉCULAS PREBIÓTICAS

CHEMISTRY ON INTERSTELLAR ICES: BINDING ENERGY DISTRIBUTIONS AND SYNTHESIS OF PREBIOTIC MOLECULES

Tesis para optar al grado de Doctor en Ciencias con mención en Química /
A Dissertation Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Science with specialization in Chemistry

por / by Giulia Bovolenta

Profesor Guía / Advisor Professor: Dr. Stefan Vogt Geisse
Profesor Co-Guía / Co-Advisor Professor : Dr. Stefano Bovino

Noviembre 2023 / November 2023
Concepción, Chile

A mio padre

In a way, astrochemistry describes
a cosmic dance of the elements
in which atoms are constantly reshuffled
from one species to another

A.G.G.M. Tielens, The Physics and Chemistry of the Interstellar Medium

Agradecimientos

En primer lugar, agradezco a la Agencia Nacional de Investigación y Desarrollo (ANID) por respaldar mis estudios mediante la Beca de Doctorado Nacional (ANID-Subdirección de Capital Humano, Doctorado Nacional, Año 2020, Folio 21200180). Mi reconocimiento se extiende a la Universidad de Concepción, por el apoyo financiero brindado a través de la Beca de Internacionalización Proyecto UCO 1866, y en especial a la Facultad de Ciencias Química por entregarme las facilitades necesarias para desarrollar mis proyecto.

Quiero expresar mi gratitud a mi guía de tesis, Dr. Stefan Vogt Geisse, por su orientación y compromiso a lo largo de todo el proceso. Su conocimientos y su perspectiva han sido fundamentales para la realización de este trabajo, así como para dar forma a mi visión de la química cuántica. Agradezco a mi co-guía, Dr. Stefano Bovino por su apoyo y valiosas sugerencias, y por introducirme en el apasionante campo de la astroquímica. Agradezco a los miembros del grupo de Química Cuántica y Modelamiento Molecular (QCMM) por el intercambio de ideas y experiencias.

Por último, expreso mi gratitud a mi familia y a mis seres queridos, que siempre me han brindado motivación y un apoyo incondicional.

Contents

List of Figures

List of Tables

Resumen

Las capas de hielo que cubren los granos de polvo son promotores importantes de química de supercie en el medio interestelar. Esta tesis tuvo como objetivo establecer un marco integral para determinar con precisión parámetros como las energías de unión y las barreras de estado de transición, esenciales para comprender la formación de especies prebióticas en ese entorno. El enfoque en las energías de unión está motivado por su papel en dictar las razón de desorción de las superficies desde las capas de hielo, mientras que las barreras de estado de transición son esenciales para calcular las tasas de reacción entre moléculas en sitios adyacentes.

La investigación exploró el impacto de una superficie de agua sólida amorfa (ASW) sobre estos parámetros, construyendo modelos rigurosos de hielo ASW y empleando métodos de teoría de estructura electrónica de alto nivel. Se generaron varios modelos de hielo ASW, incluyendo un *set de clusters* de agua y una superficie periódica amorfa de gran tamaño, utilizando tecnicas de dinámica molecular *ab initio* y *Machine learning potentials*. Los modelos de hielo se caracterizaron en términos de átomos de hidrógeno colgantes (*dangling-H*), ya que esos constituyen sitios catalíticos importantes donde la adsorción y la reacción ocurren preferentemente.

Los métodos DFT utilizados para la evaluación de la energía de unión y los estudios de reactividad, se calibraron cuidadosamente utilizando pequeños *clusters* de agua en relación con un valor de referencia CCSD(T)/CBS, con el fin de obtener valores de energía altamente precisos para usarse con los modelos propuestos de ASW.

Para automatizar el cálculo de las energías de unión en modelos realistas de ASW, se construyó una Plataforma de Evaluación de Energía de Enlace (BEEP), basada en la idea de que la energía de unión de las especies adsorcionadas sobre superficies amorfas está mejor descrita mediante distribuciones. Este enfoque de tipo *multi-unión*, junto con la determinación rigurosa de las energías de unión, tuvo un impacto significativo en observables astrofísicos, tal como en la posición de los frente de sublimación (*snow-line*) en discos protoplanetarios. Además, el análisis de los modos de enlace de los reactivos en una síntesis de Strecker de glicina reveló que no todos los motivos de adsorción facilitan los encuentros reactivos. Por lo tanto, un enfoque *multi-unión* constituye un punto de partida más completo para el estudio de la reactividad en superficies de ASW.

El segundo objetivo principal de la tesis fue investigar el papel de ASW en reacciones asociadas a la síntesis de Strecker de glicina utilizando el enfoque *multi-unión* como punto de partida. En consecuencia, se exploró cada paso de reacción utilizando métodos DFT en un conjunto variado de sitios reactivos con diferentes características morfológicas, ofrecido por el modelo

más realista de hielo ASW. Los estados de transición de los diferentes pasos de reacción se optimizaron y caracterizaron junto con un análisis detallado de los mecanismos de reacción mediante la obtención de un perfil de Coordenada de Reacción Intrínseca (IRC). Los resultados indican que, dependiendo de la naturaleza del sitio catalítico, se producen variaciones significativas en los mecanismos de reacción. Notablemente, las reacciones de transferencia de protones asistidas por agua pasaron de ser concertadas a ser en varios pasos (*step-wise*), mostrando barreras de energía menores, lo que podría proporcionar rutas sintéticas viables a temperaturas extremadamente bajas (<10 K). La investigación concluyó que existe una ruta viable bajo condiciones interestelares para la primera etapa de la síntesis de Strecker, la adición nucleofílica de amoníaco y formaldehído dando el producto aminometanol. El resultado, junto con el análisis de los valores de energía de unión calculados para las especies, sugiere que el aminometanol esté presente en la superficie de las capas de hielo interestelares. Por lo tanto, el hecho de que la especie no se haya observado en el medio interestelar hasta ahora, podría estar relacionado con dificultades en su identificación espectroscópica.

Sin embargo, la segunda etapa de la síntesis, la deshidratación del aminometanol, presentó desafíos, afectando la validez de la síntesis de Strecker como una ruta viable para la formación de aminoácidos en regiones interestelares extremadamente frías.

Abstract

Ice mantles covering dust grains are important drivers of surface chemistry in the interstellar medium. This thesis aimed to establish a comprehensive framework for accurately determining parameters, such as binding energies and transition state barriers, essential for understanding the formation of prebiotic species in that environment. The focus on binding energies is motivated by their role in dictating desorption rates from ice mantle surfaces, while transition state barriers are essential for calculating reaction rates between adjacent molecules.

The research explored the impact of an amorphous solid water (ASW) surface on these parameters, constructing accurate ASW ice models and employing high level electronic structure theory methods. Various ASW ice models, including a set of amorphized water clusters and a large periodic ASW surface, were generated using *ab initio* molecular dynamics and Machine-Learning Potentials techniques, and characterized in terms of dangling-H atoms, as they constitute important catalytic sites, where adsorption and reaction preferentially occur.

The DFT model chemistries used for binding energy evaluation and reactivity studies, were carefully benchmarked using small water clusters with

respect to a CCSD(T)/CBS reference value, in order to obtain highly accurate energy values for the proposed ASW models.

To automate the computation of binding energies on realistic ASW models, a Binding Energy Evaluation Platform (BEEP) was build, driven by the idea that binding energies on amorphous surfaces are best described by distributions. This *multi-binding* approach paired with accurate determination of binding energies, resulted to have a significant impact on astrophysical observables, particularly on the position of snow-lines in protoplanetary disks. Additionally, the analysis of binding modes of reactants in a Strecker synthesis of glycine, revealed that not all adsorption motives facilitate the necessary reactive encounters. Therefore, a multi-binding approach constitutes a more comprehensive starting point for the study of reactivity on ASW surfaces.

The second main objective of the thesis was to investigate the role of ASW in reactions associated to the Strecker synthesis of glycine using the multi-binding approach as a starting point. Accordingly, each reaction pathway was explored using DFT methods in a varied set of reactive sites with different morphological characteristics, as offered by realistic ASW ice models. Transition states of the different reaction paths were optimized and characterized together with a detailed analysis of the reaction mechanisms by means of obtaining a Intrinsic Reaction Coordinate (IRC) profile. The results indicate that, depending on the nature of the catalytic site, significant variations of the reaction mechanisms occur. Notably, water-assisted proton transfer re-

actions shifted from concerted to step-wise, displaying lower energy barriers, which might provide viable synthetic routes at extremely low temperature ($<$10 K). The research concluded that a viable route exists under interstellar conditions for the first stage of Strecker's synthesis, the nucleophilic addition of ammonia and formaldehyde yielding the product aminomethanol. Such result, coupled with the analysis of the binding energy values calculated for the species, suggests that the aminomethanol is present on the surface of interstellar ice mantles. Therefore, the fact that the species has not been observed in the interstellar medium, might be related to difficulties in its spectroscopic identification.

However, the second stage of the synthesis, the dehydration of the aminomethanol, posed challenges, impacting the validity of the Strecker synthesis as a viable pathway for amino acid formation in extremely cold interstellar regions.

1. Introduction

In the cold, deep void among the stars, pristine material is rare, embedded in gas and covered in molecule-rich ices. This place is the *interstellar medium* (ISM) and contains the remnants of celestial objects, from which new generation of stars will emerge, in a sequence of recycle and transformation. The ISM is constantly changing, and its chemical evolution is controlled by the interplay between its components, namely gas and cosmic dust. Moreover, it is permeated by electromagnetic radiation, magnetic fields and pierced by energetic particles called cosmic rays[1,2]. The elemental composition of the ISM is dominated by hydrogen (found in ionized, neutral atomic, and molecular forms), amounts of helium, and only trace amounts of heavier elements. The state of the gas ranges from ionized plasma to molecular form, while the dust (about 1% by mass) is in the form of solid grains, which may be carbon or silicon-based, and in general show dimensions on the submicron scale. The life cycle of dust is tied to the evolution of stars, as it forms in the cool expanding environment around evolved stars, is injected in the ISM by stellar winds and then distributed throughout the interstellar regions by supernovae shock waves. The dust then cycles several times through the different regions until, ultimately, processes such as grain growth, grain-grain collisions and

sputtering, alter or destroy it.

The dust has several important functions in the ISM. Abundance studies revealed that many of the refractory elements (e.g. C, Si, Mg, Fe) are locked up in the dust. It also regulates the gas-phase abundances of the elements through accretion and destruction processes, as it acts as molecular sink and catalytic surface for the encounter and reaction of species. Moreover, the dust adsorbs and scatters the stellar radiation crossing the ISM, resulting in a collective process of radiation intensity loss, referred to as visual extinction, A_v. The different regions (or 'phases') of the ISM can be categorized according to their capacity of starlight dimming, in: dense clouds ($A_v > 5$ mag), translucent clouds (1 mag $< A_v <$ 5 mag), and diffuse clouds ($A_v \leq$ 1 mag). Those clouds are neither uniform nor dynamically quiescent on long timescales: they interact and cycle between phases, as the ISM is far from being in thermodynamic equilibrium[3] (a schematic representation of the classic interstellar matter cycle can be found in Figure 1.1).

Among the variety of interstellar environments, the *dense clouds* are maybe the regions of greater astrochemical interest. They show a large temperature gradient, from 10 K in the most dense regions, to 300 K approaching the forming protostar, and their gas density is about $10^4 - 10^8$ hydrogen atoms per cm^3. In these clouds, elevated concentration of gas, along with dust opacity, offer protection from the UV radiation, allowing the formation and sur-

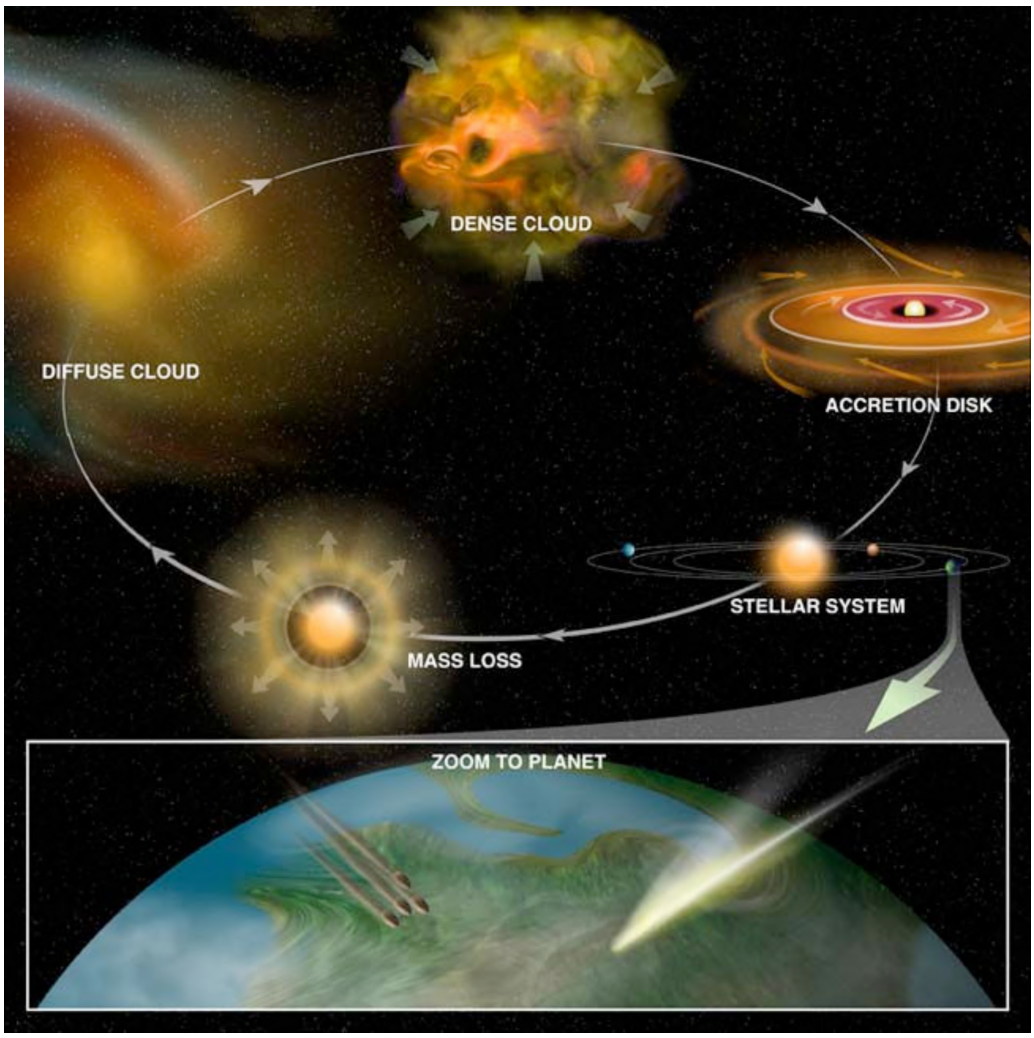


Figure 1.1: Schematic illustration of the classic interstellar matter life cycle. Adapted from Pontoppidan [4].

vival of 'large' interstellar species. Astronomical observations of the dense clouds' gas have been carried at infrared, radio, millimeter, and submillimeter frequencies revealing several organic species such as nitriles, aldehydes, alcohols, acids, ethers, ketones, amines, amides and long-chain hydrocarbons. Among them, a class of compounds which are thought to be involved in the first biochemical reactions leading to the origin of life, the so-called *prebiotic*

molecules. Small molecules such as hydrogen cyanide, urea or aminoacetonitrile, are considered prebiotic since they are precursors of life-necessary species, like amino acids, ribonucleotides and phospholipids, among others.

An important trait of the dense clouds environment is the fact that, in the regions where temperatures are below 150-180 K, the dust grains are covered by thick ice mantles ($\sim$100 monolayers). In terms of ice composition[5], water is known to be the primary component, along with traces of CO, NH_3, and CO_2, CH_3OH. Although precise information about the morphology is still lacking, water is assumed to be amorphous[6] and to contain a certain number of micropores[7] (even though energetic processes taking place in the ISM are expected to reduce the porosity[8]). Ices are an important reservoir of chemical elements, for example up to 60% of the oxygen not included in silicates is incorporated in known ice species. Therefore, addressing the origin and development of interstellar ices is relevant to understand the evolution of organic and volatile reservoir of such regions[9].

Large samples of ice sources spanning diverse dense cloud regions, evolutionary stages and luminosities have been necessary to determine ice abundances and to identify how ice processes depend on their environment[10]. In the modern accepted vision[11], interstellar ice formation can generally be divided into three stages, related to the cloud matter cycling between regions. The first stage of ice formation is driven by atomic hydrogenation reactions in cold clouds and results in H_2O-dominated ice. At later times, in star-forming

regions, CO depletion (freeze-out) stage takes place, leading to the formation of apolar ice layers on the grain. The last ice evolution stage is thermal and UV processing around the nascent protostar, resulting in CO desorption, ice segregation, and the formation of more complex organic molecules. A pictorial summary of the life-cycle of interstellar ice can be found in Figure A.1. Therefore, studying the composition and structure of ices, as well as the efficiency of the relevant chemical surface reactions as functions of the interstellar environments, appears to be pivotal for understanding ISM chemical evolution[10,12].

The dense clouds are also the birth sites of stars in the ISM. Their geometrical structure therefore is diverse, as it contains regions of relative quiescence (*dark clouds*) as well as warm and turbulent regions of current star formation (*star-forming regions*)[10]. The observed organic and prebiotic compounds display large spread and compositional variations between quiescent dark clouds and star-forming regions, so a description of the characteristics and chemical fauna of the different environments is required and will be provided in the following.

Dark clouds

In the coldest ($\leq$ 20 K) part of the dense clouds, prior to the onset of star formation, dust grains effectively shield molecules from interstellar UV photons; however, cosmic rays can penetrate throughout and cause a induced-

UV field through H_2 ionization and further excitation and decay processes. Chemistry in dark clouds, therefore, is driven by a gas-phase chemistry (typically ion-molecule and neutral-neutral reactions), supplemented by grain surface processes of radical coupling and atoms addiction/abstraction. Studies of interstellar absorption features are naturally dependent on the availability of background field stars to act as continuum sources. Recently, thanks to the extraordinary high sensitivity offered by James Webb Space Telescope (JWST), it has been possible to carry out the first observations[13] of pristine cloud ices at $A_v > 50$. The ice inventory suggested that the complex variety of organic species observed likely forms in water-rich ices, result of the aforementioned first ice formation stage (rather than after the second stage of catastrophic CO freeze-out, through purely hydrogenation pathways of the CO apolar layer, as it was previously assumed). The presence of prebiotic species in those regions, therefore, is compatible with non-energetic synthesis mechanisms that can be achieved before the formation of a protostar.

Star-forming regions

Star formation has a profound effect on the chemistry of the surrounding medium. Those regions are warmer[14], ranging from 25 to 40 K, since in the early phases of protostellar evolution, the surrounding gas and dust experience radiative heating and shock waves. Ground-based telescopes and space observatories[15–17] have provided abundance inventories along sight-

lines through the envelopes of nascent protostars. Star-forming regions are particularly rich in organic and prebiotic molecules, which are thought to derive from various possible origins. Already mentioned gas-phase and surface processes can be amplified at higher temperature. Moreover, massive accretion of molecules (e.g. CO, HCN and acetylene) onto dust during core collapse increases ice molecular diversity. Lastly, in the regions around the forming protostar (hot molecular cores), higher temperature (70 - 300 K) and density ($> 10^6$ cm^{-3}) is observed. The mobility of adsorbates on the dust is then enhanced, increasing the probability of species encounter. Several thermal reactive channels, that presented prohibitive activation barriers for the colder environments, are now enabled to take place on the grain surface, leading to the formation of more complex species, such as long hydrocarbon chains and sugars and amino acids precursors. Also, evaporation/sputtering of icy grain mantles can occur[18], which, in turn, facilitates warm gas-phase reactions, as ejected molecules from the mantles provide higher local concentration of species[19,20].

Protostellar disks and interstellar heritage in the solar system

The protostellar stage includes the formation of an accretion disk of solid material which ultimately become the formation site of planets. A schematic illustration of an accretion disk can be found in Figure A.2. The disks are observed to have cold (20–150 K) midplanes where ices should be abundant,

beyond the so-called disk snowline[21,22]. It is often difficult to study these disk ices directly because of confusion between disk ices and ices in foreground clouds[23]. Comets and other primitive bodies in our own solar system may, however, carry a record of the ice composition in such disk midplane.[11]

Comets are porous agglomerate of frozen gases, ices, and rocky debris[24], and are thought to form in a mildly warm[25] region (45 K) about the disk midplane, from materials that were not assembled into planets. There is no clear consensus on whether these materials survived the voyage from the outer space to the early Earth and to what extent their composition resembles the original icy-dust or it is the result of chemical processing within the star-forming cloud. However, analysis of the distribution of cometary ice abundances are consistent with the idea that most cometary ices have a protostellar origin[11], making them likely the most pristine remnant objects from the formation of the solar system. Recent measurements revealed that not all comets are fully covered in ice, some present rather warm average temperature ($\sim$200 K), compatible with a surface scenario composed of patches of ices and exposed refractory material[26].

Meteorites are carbonaceous (stony-meteorites or chondrites) or metallic (iron-meteorites) bodies that landed on Earth, derived from the fragmentation of larger interstellar objects such as asteroids, comets or moons. Before the advent of *in situ* cometary missions, meteorites were the only known extraterrestrial material. The importance of "daughter" objects such as comets and

meteorites lies in the surprising amount of prebiotic compounds discovered in those solar system primitive bodies, including a large number of amino acids[27,28]. The samples also present amino acids that are unknown in terrestrial biochemistry or present very low abundancies, suggesting an interstellar - or at least extraterrestrial - origin. Furthermore, the enantiomeric excesses (7–9%) measured in several alien amino acids opens questions about the role of solar system primitive bodies in relation to the developments of life, hinting to an extraterrestrial provenance for the homochirality found in protein amino acids, i.e. to a primordial L/D excess further amplified at a later stage[29]. Therefore, the early Earth and perhaps other primitive planets may have obtained most of its prebiotic organic material from the arrival of such primitive 'seeding' bodies.

A detailed and quantitative understanding of the relevant chemical alterations that the organic matter undergoes in dense clouds, comets, and meteorites during the ISM evolution, and their common link, is necessary to provide constraints for the processes that lead to the origin and distribution of life in the Galaxy.

1.1 Theoretical background

1.1.1 Interstellar ices modeling

Interstellar ices are mainly composed of water, in the so-called amorphous solid water (ASW) form. Experimental ASW, grown by vapor deposition at low-temperature has been studied extensively[30–32], in order to gain insights about processes occurring on ice mantels. However, vapor-deposited ASW may not necessarily represent realistic ice mantles: the formation of H_2O-ice on dust grains does not primarily occur through deposition of gaseous water, but rather by sequential surface-hydrogenation reactions of O and O_2, or through the surface reaction of OH with H_2 (See ref. Hama and Watanabe [33] and citations therein).

ASW has distinct properties from crystalline-ice e.g. lower thermal conductivity, larger surface area and higher porosity. Even though its structures is not precisely known, ASW has no long-range order, resulting in great surface heterogeneity. An implication is that more potential adsorption sites exist on the grains surface, as it has been confirmed by experimental[34,35] and theoretical[36] works.

Figure 1.2 reports some of the ice models that have been used during the past years for simulating desorption of molecules and reaction processes under dense clouds interstellar conditions. The simplest approach is approx-

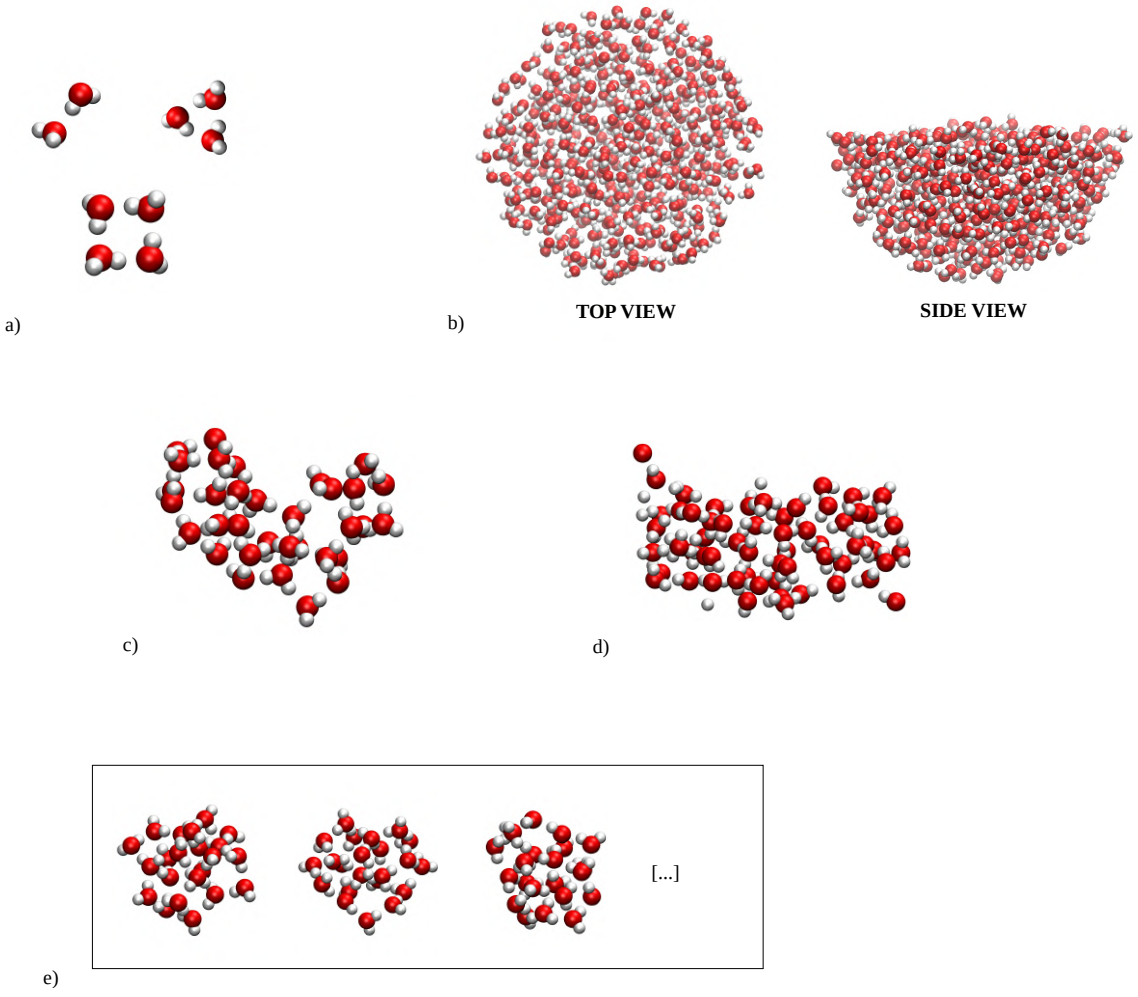


Figure 1.2: a) Small water cluster used for benchmark studies[36] and catalogues of binding energy of interstellar species[37,38]; b) Large periodic model of 683 water molecules computed with QM/MM techniques used for H_2 adsorption study[39]: 57 surface molecules at the center of the hemisphere have been treated at the quantum level, while the rest at MM level; c,d) Single medium size (33 and 63 water molecules) clusters used for reactivity studies[40] and binding energy calculation[41,42] respectively; e) Set of medium size clusters (20-22 water molecules) used for binding energy evaluation of interstellar atoms and molecules[36,43].

imating the surface using various water clusters up to the pentamer[37,44] (Figure 1.2a). This model has the computational advantage that it allows to use high level *ab initio* techniques that yield very accurate results. However, considering the small scale of the clusters, such model failed in reproducing the statistical variety of the adsorption sites present on a actual amorphous ice. Regarding the simulation of more realistic surfaces, there are mainly two strategies: using a large-scale system and applying a *embedding technique*, i.e. treating the molecules close to the adsorption site with a higher level of theory and the rest of the molecules with a lower level (Figure 1.2b), or preferring a medium-sized ASW clusters. This can be done on a single cluster by exploring different adsorption sites (Figure 1.2c,d), or using the so called *set of clusters* approach[36,43], where a larger ice surface is spanned by a series of many amorphized clusters (Figure 1.2e).

1.1.2 Gas-grain astrochemical models

There is a variety of processes that can lead to the formation of molecules in the ISM, but these can be separated into two broad classes: reactions that occur in the gas-phase and reactions that occur on the surface of dust grains[3]. Among the gas-phase reactions, only ion-molecule interaction, dissociative recombinations and radical associations are viable in the dense clouds environment, since they are exothermic and typically lack activation barriers. However, when it comes to the synthesis of saturated organic species such as

large prebiotics, gas-phase chemistry alone is not an efficient method. Suitable reaction pathways typically would necessitate high temperatures and/or three-body reactions, conditions that are not commonly encountered within the ISM. Moreover, some abundant small interstellar species (such as H_2, H_2O, CO_2, CH_3OH) require surface chemistry to attain the observed abundances[45,46].

As mentioned above, the surface serves several roles. First, molecules and atoms from the gas-phase adsorb on dust grains, due to van der Waals interactions between them and the ice mantles. The dust also acts as local encounter point for surface reactions, and eventually absorbs and dissipates some of the excess energy generated by association reactions, stabilizing the resulting products[47]. Since the flux of species coming from the gas-phase to the surface is extremely low, surface reactions are thought to occur through a sequence of elementary processes. In the most important mechanism, the Langmuir-Hinshelwood, the elementary steps include adsorption, diffusion across the ice mantle, prior to chemical encounter and reactions. The product is thermalized on the ice, or can be further processed and/or eventually desorb back to the gas-phase[2,33]. Surface chemistry is considered to be in the *diffusion limit*, where the reaction rate is essentially limited by the rate at which species can migrate on the surface. Diffusion is thought to be very slow under cold dark clouds conditions, especially for polar molecules, that interact strongly with the ice mantles[48,49].

Kinetic models, meant to predict the abundances of species over time in a certain interstellar environment, have historically treated gas-phase and grain surface chemistries separately. However, nowadays it is commonly accepted that a complete and realistic astrochemical kinetic model should account for both type of processes, within a paradigm known as *gas-grain* chemistry[50–52]. Gas-grain astrochemical models typically adopt the rate-equation approach, where the complex interplay of phenomena that interests a specific interstellar species M is accounted for in a set of coupled differential equations:

$$\begin{cases} \dfrac{d}{dt}n_g(M,t) = & f_{des,M} + \displaystyle\sum_i f_{react,g}(i \rightarrow M) + \sum_{i,j} f_{react,g}(i+j \rightarrow M) \\ & - \displaystyle\sum_i f_{react,g}(M+i \rightarrow) - f_{diss,M} - f_{ads,M} \\ \dfrac{d}{dt}n_s(M,t) = & f_{ads,M} + \displaystyle\sum_i f_{react,s}(i \rightarrow M) + \sum_{i,j} f_{react,s}(i+j \rightarrow M) \\ & - \displaystyle\sum_i f_{react,s}(M+i \rightarrow) - f_{diss,M} - f_{des,M} \end{cases} \tag{1.1}$$

(adapted from Tielens [3] and Cuppen et al. [53], to which reference is made for the functional form adopted for each term). The two equations represent gas-phase and surface processes, respectively; $n_{g,s(M)}$ are the number densities of species M in gas-phase and on the dust surface. For each equation there are positive terms that account for species formation and negative terms for destruction. A certain phenomenon, like adsorption on M onto the sur-

face, results in the loss of the species from the gas-phase ($-f_{ads,M}$ in the first equation) but also in the increasing of its surface density ($+f_{ads,M}$ in the second equation), and *vice versa* for desorption processes ($f_{des,M}$) of the species back to the gas-phase. The surface adsorption of M depends on its sticking coefficient as well as surface composition and properties. The reaction terms ($f_{react,g,s}$) account for the species' gain and loss due to uni- and bi-molecular grain-surface and gas-phase reactions between other generic species i, j. The bimolecular terms depend on the activation barrier of the chemical reactions and also on the diffusion rate of the species, as they scan the ice surface in order to meet. Examples of unimolecular reactions in the dense clouds are photodissociation or photoionization due to cosmic-ray-UV induced field. Large reaction networks treat up to a few hundred different species, and the generated set of differential equations is numerically solved using a multi-step integrator[53–55].

1.1.3 Estimation of Binding Energy

The desorption of a species from the ice mantle (f_{des} terms in Equation 1.1) consists of acquiring sufficient energy to overcome the energy barrier to return to the gas-phase. It can occur thermally (due to cloud temperature fluctuations) or non-thermally (due to external energy sources such as UV radiation, impact of cosmic rays or ice warming-up, thanks to the energy released

during *in situ* exothermic reactions). In a process of thermal desorption, the residence time of a species on the surface, τ(s) is defined as the inverse of its desorption rate, k_{des}, i.e. it is the characteristic time scale required to acquire sufficient energy to desorb;

$$\tau = \frac{1}{k_{des}} = \nu^{-1} \exp\left(\frac{E_{b,M}}{k_B T}\right) \tag{1.2}$$

where the rate constant is described by an Arrhenius equation where $E_{b,M}$ is the binding energy (or adsorption energy) of the adsorbate M, T is the dust temperature, and ν is bond-vibration frequency between the adsorbed species and the substrate[56]. Therefore, the binding energy magnitude depends on both the physical nature of the molecule (e.g. polarity, polarizability, etc) and of the grain surface (morphological parameters, coverage, etc)[57]. Normally, the diffusion rates that appear in the second term of Equation 1.1, depicting the hopping of species between different ice mantles' sites, are also described by Arrhenius equations assuming diffusion barrier as a fraction (0.3-0.5) of the binding energy[53]. Thus, the binding energy of a species is a quite important parameter that controls - directly and indirectly - the chemical behavior of interstellar species.

Unfortunately, one of the main problems in the kinetic modelling description of reaction networks is the lack of appropriate laboratory data of binding energies and desorption efficiencies on different interstellar ices. Binding

energies can be experimentally measured using Temperature Programmed Desorption (TPD). This kind of experiments are performed under ultra-high-vacuum conditions ($\sim 10^{-9}$ mbar): depositing a certain amount of species on a substrate at very low temperature and increasing the temperature in a linear manner until desorption happens. Monitor of the desorption behavior and analysis of the volatiles allow to estimate the residency time and derive the binding energy. However, current facilities are affected by sensitivity problems in the identification of volatiles with the same mass and by the incapability to study radical species[58]. It is, then, necessary to complement experimental studies with theoretical computations. It is worth pointing out that the accuracy of theoretical calculations is paramount since the desorption rate constant, k_{des}, depends exponentially on the binding energy, therefore uncertainties on its value might have a dramatic effect on the models[59].

The efforts to obtain a binding energy extensive catalogue for small molecules on ice surfaces have been so far limited to Density Functional Theory (DFT) calculations on small water clusters (up to 6 molecules,[37,44], Figure 1.2a) or interaction with water monomer by linear semi-empirical models[38]. However, as previously said, a more reasonable choice for theoretical description of interstellar grain is ASW, and the aforementioned works did not capture the complete statistical nature of the adsorption of molecules. Recently, there have been efforts to consider more sophisticated surface models

presenting a variety of sites available for the adsorbate to bind to the surface. Among them, Shimonishi et al. [43] proposed to use a a medium size amorphous surface model (Figure 1.2e), in order to provide a sets of values for C, N and O atoms. In addition, several reactivity studies have been carried out by using a QM/MM ASW simulated surface, including binding sites and binding energies analysis[39,60,61] (Figure 1.2b).

The most complete study thus far[41], provided binding energy of 21 molecules and atoms, using a slab of ASW with periodic boundary conditions (Figure 1.2d), considering up to eight binding sites per molecule. More recently, a ONIOM QM/QM hybrid method has been used to compute a set of 8-10 binding energies of several small species, including radicals[62,63].

The few available experiments on ASW surfaces show that the molecules interact with grains in different ways depending on the available type of sites[57,64]. Some sites are more suitable for strong interactions and are usually the first to be populated, while "peak" sites (as opposed to "valley" sites) produce weaker interactions. Thus, the capability of molecules to remain bound onto the surface is determined by the distribution of their binding energies. According to experimental[34,35] and theoretical[36] findings, these binding energy distributions resemble a Gaussian function, suggesting the need to reconsider the treatment of binding energy values in astrochemical models as well as in interstellar surface reactivity studies.

In a new so-called *multi-binding* approach, the binding energy average value for a certain species, normally included in kinetic models, is substituted by a binding energy Gaussian distribution. The possible impact of the new multi-binding approach in terms of kinetic study, has been recently presented by Grassi et al. [65]. Their results shown that considering a larger variety of binding sites with corresponding binding energy, allowed to more accurately estimate the amount of desorbed species. This is due to the fact that the new approach enables a (weighted) incorporation of sites where molecules stay on the ice for longer times, due to their greater binding energy, which were commonly excluded in the average binding energy value approach.

The multi-binding approach also represents a novel paradigm for the study of surface reactions. Taking into account a variegate set of sites on the ice where molecules are tightly or loosely bound corresponds to a more realistic census of possible reactive configurations. Strongly bound molecules with higher residence time onto the grain surface, - that usually corresponds to poor or no reactivity at all - would become available for reacting with loosely bound molecules that can diffuse more easily, even at dark cloud temperatures. In summary, surface reactivity is expected to be determined by the two extreme situations represented by the wings of the binding energy distribution: molecules very energetically bound and with low mobility, and vice versa, hence the pivotal importance of accurately predicting the 'entire' shape of the binding energy distribution, in order to constrain the uncertainties. As

mentioned above, a database of accurate binding energies distributions of important small and medium size interstellar species is, so far, lacking.

Lastly, the new multi-binding approach is meant to disclose the effect of the ice morphology on a certain reaction mechanism, clarifying the efficiency of the catalytic role exerts by the surface in a varied set of reactive sites with different characteristics, as offered by realistic ASW ice models.

1.1.4 Synthesis of glycine under interstellar conditions

The origin of amino acids within the prebiotic chemistry of early Earth is a subject of long-standing fascination. While a plethora of amino acid precursors has been discovered within the dense regions of the ISM, the quest for the amino acids themselves has yielded no success so far. Glycine is the simplest achiral amino acid, starting from it, numerous chemical reactions lead to more complex amino acids[66]. As mentioned above, glycine has been discovered in the solar system most pristine bodies, such as meteorites that landed on Earth[27,67] and also by ground based and *in situ* survey of comets[68]. The presence of glycine in the interstellar medium is still controversial[69–72] and to date its search has been the object of the largest number of investigations ever made for any prebiotic molecule[73].

Whether glycine can be formed under the harsh conditions of dense clouds, is also still to answer. As mentioned above, such complex molecules are difficult to produce and, additionally, they are challenging to identify unambi-

guously[1,74]. The majority of interstellar complex molecules have been detected by radio astronomical means, and by infrared, visible and ultraviolet spectroscopy in minor part, in particular for molecules with no rotational spectrum. A purportedly successful detection of glycine[69], which was later refuted[70], focused on three hot-core sources. Those regions present large abundance of complex molecules, which is a challenge for the detection, as such compounds have many rotational transitions, resulting in frequent overlap of spectral features. Besides, complex molecules are generally less abundant than smaller ones and often have greater moments of inertia, resulting in a higher number of rotational energy levels populated at a given temperature. Even if a complex molecule were as abundant as a simpler one, it would still be harder to detect, due to its less favorable partition function. In fact, the line intensity would be spread over more lines and across a narrower frequency range, causing a significant drop in maximum intensity, even at the same number densities. Lastly, for species adsorbed on dust grains, spectroscopic detection is much more difficult than in the gas-phase because the spectral features in solid-phase are broader and not very distinctive[2,74]. Additional possible obstacles to glycine detection is the fact that it is the least stable of the three common isomers of the same generic formula, therefore, in light of the 'minimum energy principle'[73], it might also be the least abundant one. However, ISM chemistry is not under thermodynamic control[3], and noticeable exceptions exists to such principle.

Regarding the synthesis of interstellar glycine, several gas-phase formation routes have been explored theoretically. Those studies involved mainly radical-radical coupling[75–77] and ion-molecule associations[78–81], and revealed high activation barriers, reinforcing the hypothesis that glycine, as many large interstellar species, is more likely attained on the icy-surface of dust grains. Among the surface reaction pathways, the most exploited is the Strecker synthesis, for historical reasons. It was inadvertently discovered in 1850[82] and represented the first successful demonstration of the laboratory synthesis of amino acids. It is assumed to have taken place in solution in the oceans of the primordial Earth, as well as in "aqueous alterations" of meteorites precursor. This last hypothesis embraces the interstellar-parent body scenario that views meteorites as deriving from objects like asteroids, comets or moons. Such bodies might have underwent physico-chemical transformations resulting in the melting of part of the water ice coating. That would allow liquid-phase organic reactions, like the ones associated to the Strecker synthesis, to take place on the lower layers of the remaining ice mantle of the parent body or in its interior, in case of porous materials such as comets[83]. The generated amino acids therefore might have stayed inside of the newly formed meteorite, shielded by radiation, until landing on Earth.

The entire Strecker synthesis of glycine has been successfully carried out in laboratory under simulated interstellar conditions, in various experimental settings[84–87]. Such experiments are typically performed on interstellar ice

analogs: the reactants are deposited at very low temperature on a substrate and the reaction takes place *via* UV photolysis followed by thermal processing of the irradiated ice. Final hydrolysis of the residue obtained yields glycine along with other amino acids.

The synthesis has also been studied theoretically in an extensive way (literature sources are provided in Table1.1, along with details about the synthesis mechanism). However, the synthesis is composed of a rather long sequence of neutral-neutral reactions, which typically have large activation barrier both in gas and solid phase, limiting its feasibility to some specific ISM environments.

Alternative surface reaction pathways based on neutral-neutral interactions of small molecules such as CO, HCN and $NHCH_2$ have been studied theoretically[88–91], although they resulted to be not viable at low temperature, as well. Recently, new non-energetic pathways expected to be suitable to efficiently synthesize glycine in cold cloud environments have been proposed. Krasnokutski et al. [92] carried out laboratory experiments based. on a mechanism initiated by addition of C atoms to NH_3, followed by reaction with CO. Carrying the reaction in liquid helium droplets at 10 K yielded glycine[92], while species co-deposition on a cold substrate followed by precursors condensation yielded a glycine polimer[93,94]. Although the novel hypothesis of producing various length peptides bypassing the stage of amino acids synthesis, is fairly intriguing, the experiment required external source

of energy. The system has been subjected to temperature increase, and the peptide bond formation has been estimated to take place reaching 100-120 K. Furthermore, theoretical studies of such C-addition reaction mechanisms have shown that, despite the chemisorption of excited state C atoms onto NH_3-[92] or H_2O-rich ices[95] is effectively barrierless, the following steps of glycine and poly-glycine synthesis are not. The authors suggested that in the dense cloud environments, the required energy surplus might have been provided non-thermally, thanks to the exothermicity of the previous steps, or by alterations in the conditions of the reactive site due to cosmic rays impact.

A last and, perhaps, unique actual non-energetic synthesis has been proposed by Ioppolo et al. [96]. They carried out laboratory experiments based on substrate co-deposit of simple precursor species ($CO:NH_2CH_3:O_2:H$ = 3:1:2:40) at 13 K. The reaction proceeds *via* a chain process of *in situ* radical formations and subsequent hydrogen abstractions, and glycine is ultimately obtained by radical coupling reactions. This promising mechanism is probably hitherto the most viable route to glycine in the ISM cold regions.

1.1.4.1 The Strecker Synthesis

The overall general reaction sequence takes into account a carbonyl compound, ammonia (NH_3) and hydrogen cyanide (HCN):

$$R_1C(O)R_2 + NH_3 + HCN + H_2O \rightarrow R_1CR_2(NH_2)COOH + NH_3 \quad (1.3)$$

Where R_1 and R_2 may be either a proton or various substituted or non-substituted alkyl chains. In the case of glycine, the aldehyde that starts the synthesis is formaldehyde (H_2CO). The synthesis is comprised of 4 steps, summarized in Table 1.1 along with details of the reaction mechanism.

Table 1.1: Steps of the Strecker synthesis of glycine. The reaction starts with the nucleophilic addition (1.) of ammonia to a carbonyl compound yielding the amino alcohol, followed by its dehydration (2.) to give a imine. Nucleophilic addition (3.) of the imine to a cyanide species provides the nitrile. Glycine is obtained through two consecutive nitrile hydrolysis (4.). Column 2 and 3 reports theoretical studies carried out for each step.

# Step	Literature studies	
	gas-phase	ice model
1. $H_2CO + NH_3 \longrightarrow NH_2CH_2OH$	97–101	97,98,100,102,103
2. $NH_2CH_2OH \longrightarrow NHCH_2 + H_2O$	98,100	100
3. $NHCH_2 + HCN \longrightarrow NH_2CH_2CN$	98,104	98,104,105
4. $NH_2CH_2CN + 2\,H_2O \longrightarrow NH_2CH_2COOH + NH_3$	98	98

One of the reasons that made the Strecker synthesis so appealing for the astrochemical community over the years, is the fact that all the participant species have been detected in the dense regions of the ISM[106–110], ranging from dark clouds to warmer environments showing large abundances[111,112] (density of the order of 10^{-4} cm^{-3}), which underlines their importance to interstellar chemistry[113]. Most of the intermediates compounds in the syn-

thesis have been detected as well[111,114,115], apart from the first step intermediate,aminomethanol. The presence of aminomethanol in the interstellar medium is object of debate[102,116–118]. This species belongs to the class of molecular compounds called hemiaminals, which contain a amino and a hydroxyl groups bound to the same carbon. Hemiaminals obtained from primary amines are generally considered unstable[99] and the amino alcohol in the Strecker synthesis is normally treated as a short lived intermediate subject to water elimination in order to provide the imine product. However, the aminomethanol has been successfully formed in a seminal work[119] using ice-analogs with a 10 K mixture of H_2O:NH_3:H_2CO, co-deposited on a cold surface, subsequently warmed up, but the identification of the reaction product was considered ambiguous. The experiment has been repeated by Bossa et al. [116], confirming the result and estimating the aminomethanol formation to take place between 80 and 100 K. Recently, the aminomethanol has been prepared also starting from different precursors (CH_3NH_2 and O_2), upon exposure to energetic electrons[117]. All the mentioned experimental works confirmed the hypothesis that aminomethanol can be formed in dense cloud conditions under the influence of some energy source, hinting that the major reasons behind the lack of detection are identification difficulties. In fact, the aminomethanol has strong absorption features that, in astronomical spectra, are blended with water and silicate bands. The remaining bands fall in regions where several other organic compounds adsorb. However, an important

aspect of the aforementioned work[116], has been attempting to constrain the contribution of aminomethanol to the astronomical spectra of different dense regions. They took into account the astronomical ice surveys of a dark cloud and two star forming regions and compared them to the experimental spectra provided in their work. The experimental aminomethanol ice displayed major similarities with the highly-processed ice of active star forming regions rather than the more pristine ice of the quiescent dark cloud. The result is in agreement with a proposed purely thermal synthesis, hinting that the aminomethanol should be formed and detectable in warmer environments, such as hot cores of star forming regions. The work[116] also provided the first experimental result for the desorption energy of the aminomethanol. The value (of about 14 kcal/mol) is larger than previous estimations[37,38] for the other reactants, suggesting that the aminomethanol might reside on the interstellar ice surface. In light of these considerations and unlike previous studies[120], stage 1. and 2. of the Strecker synthesis of glycine will be henceforth assumed to be distinct and non necessarily consequential.

Each stage of the synthesis has been simulated in gas-phase and also adopting ice surface models of various kind; previous studies are listed in Table1.1. The surface models vary from minimal water clusters ranging from dimer to pentamer[97,100,102,103,105] to medium size clusters of 9-12 molecules[103] to the most complex 18-water molecules crystalline model[98]. Some studies

also included a isodensity-surface polarizable continuum model (PCM) to include solvation effects[97,103,104]. Overall the presence of the ice surface lowers the activation barrier of all the processes, although the second and forth stages present large activation energies (more than 30 kcal/mol).

Catalytic effects due to external agents such as water molecules is well known[121,122]. The effect of the surface water molecules is twofold: on the reaction site geometry and on the electrostatic properties of the reactants. Having the reactants coordinated to a third body (the surface) facilitates the encounter and the structural changes that precede the transition state electronic rearrangement (i.e. helping in reaching the activated complex), thus diminishing the activation energy. Moreover, in all the Strecker synthesis simulation on the ice surface, the steps that involve proton transfer processes are assumed to take place *via* a water-assisted concerted mechanism called *proton-relay*. Such catalytic arrangement, previously proposed theoretically in astrophysical context[123], considers the explicit participation of water molecules as mediators in the proton transfer, through the establishment of a ring structure that connects the reactants. The result is that the icy-surface allows for the easier proton-transfer between the adsorbed molecules.

The variability of the substrate ice models is reflected in considerable uncertainty in the barrier heights, suggesting the need for a comprehensive approach to estimate the catalytic role of water in a ASW surface. Besides, the use of a more elaborate amorphous surface model would provide a better

description of the reactive sites.

In conclusion, the Strecker synthesis of glycine appears to be a suitable case of study to apply the *multi-binding* approach, in order to explore the consequences of the new paradigm in terms of adsorption properties and reaction mechanisms under interstellar conditions.

1.2 Hypothesis and Objectives

Hypothesis

The surface morphology of amorphous solid water (ASW) models, constructed using *ab initio* molecular dynamics simulations and quantum chemistry methods, affects the physicochemical processes that take place on interstellar ice mantles, such as desorption of small molecules and chemical reactions.

General Objective

Calculate binding energy distributions of astro-relevant molecules, including prebiotic precursors, adsorbed on ASW models. Use the distributions as a starting point in the study of reactions associated with a Strecker-type synthesis of glycine, within a *multi-binding* approach, taking into account surfaces with different morphological characteristics and compositions.

Specific objectives

1. Modelling homogeneous and porous surfaces of ASW using *ab initio* molecular dynamics techniques. The appropriate porous size will allow the study of the stabilizing effect of water molecules on the reactions of interest

2. Identification of the binding sites of the molecules involved in each step of the Strecker synthesis of glycine on the ice model, and evaluate the binding energies. Building of a binding energy distributions database of small molecules present in the ISM

3. Calculation of each stage of the Strecker synthesis with high-level wave function based methods, using a 2-water molecules cluster, to benchmark stationary point energies and geometry of various DFT methods

4. Identification of the most suitable binding sites on the ice model, taking into account several initial situations for the adsorbed species and calculation of the reaction pathways using the surface model. For each step, the stationary points along the reaction will be characterized in terms of geometry, energy and zero-point vibrational energy (ZPVE)

5. Assess the effect of using a ice mixture (H_2O:NH_3 = 2.5:1) surface on the reactions of interest

2. Methods

The chapter illustrates the theoretical and computational methods used in this thesis.

2.1 Electronic Structure Methods

The postulates and theorems of quantum mechanics serve as the foundational framework for predicting observable chemical properties from first principles[124–128]. The state (wavefunction) of a quantum mechanical system is represented by a vector ($|\Psi\rangle$) in the Hilbert space and every observable - quantity of interest - attributed to it is described by an hermitian linear operator. The only possible result of the measurement of a certain observable A on a state is one of the eigenvalues of the corresponding operator ($\hat{A}$), the probability of which is given by the Born rule, being $|\langle a_n|\Psi\rangle|^2$, in the simplest case, for an eigenstate a_n. Nevertheless, immediately after the measurement yielded a_n, the state of the system is postulated to "collapse" to the normalized eigenstate $|a_n\rangle$, hence the probabilistic nature of quantum mechanics.

The states evolve in time according to:

$$\hat{H}\Psi(x,t) = i\hbar\frac{\partial\Psi(x,t)}{\partial t} \tag{2.1}$$

where H is the observable Hamiltonian and corresponds to the total energy of the system. For most systems, H does not contain time explicitly, therefore a separation of variables is applicable, yielding the time-independent not relativistic Schrödinger equation:

$$\hat{H}(\mathbf{r})\Psi(\mathbf{r}) = E(\mathbf{r})\Psi(\mathbf{r}) \tag{2.2}$$

The probability density $|\Psi(\mathbf{r})|^2$ and the energy of states that are solution of Equation 2.2 are independent of time, thus they are denominated stationary-states wave functions and are of central importance in chemistry. The Hamiltonian operator for a system of M nuclei and N electrons is given by a sum of kinetic and potential energy operators:

$$\hat{H} = -\sum_{i=2}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2.3}$$

where summations over i and j run over electrons, and summations over A and B run over nuclei; r_{ij}, r_{iA}, and R_{AB} denote electron-electron, electron-nuclei, and internuclear distances; M_A is the ratio of the mass of nucleus A to the mass of an electron, and Z_A is the atomic number of A.

In order to find approximate solutions of Equation 2.2, quantum chemistry resorts to some approximation, the central one is the Born-Oppenheimer approximation[129], that stems from the assumption that the nuclei are stationary from the electronic point of view, due to the considerable difference in mass between the particles. That corresponds to neglecting the second term of Equation 2.3 and consider the last a constant. The remaining terms are called the electronic Hamiltonian, $\hat{H}_e$, so that the resulting electronic wave function, E_e, describes the motions of electrons and depends only parametrically on the nuclear coordinates:

$$\begin{aligned} \Psi_e &= \Psi_e(\{\mathbf{r}_i\};\{\mathbf{R}_A\}) \\ E_e &= E_e(\{\mathbf{R}_A\}) \end{aligned} \tag{2.4}$$

Solving the electronic problem for a large set of nuclear coordinates would provide the total energy, E_{tot}, (that also includes the constant nuclear repulsion), in the form of a parametric hypersurface - potential energy surface (PES) - upon which the nuclei move.:

$$E_{tot} = E_e(\{\mathbf{R}_A\}) + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{2.5}$$

If such solutions are generated without reference to experimental data, the methods are usually called *ab initio* (latin for “from the beginning”).

2.1.1 Independent-particle models

The dynamics of a many-electron system is very complex. However, a significant simplification, conceptually and computationally, can be achieved by employing independent-particle models, where the motion of one electron is treated as if it is unaffected by the presence of others. In the *Hartree-Fock* (HF) approximation, this is realized by averaging out the interactions between the particles. The resulting one-particle solutions pave the way for more advanced computational methods.

In the HF model, each electron is described by an orbital χ_i, and the total wave function is a product of orbitals. As electrons are indistinguishable fermions, the wave function must be antisymmetric. This is achieved by arranging the orbitals in a Slater determinant:

$$|\Psi_0 = |\chi_1 \chi_2 \cdots \chi_N\rangle \tag{2.6}$$

which refers to the ground state of a N-electron system. The best set of orbitals is determined according to the variational principle, i.e. it corresponds to the wave function that gives the lowest energy:

$$E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \tag{2.7}$$

Minimizing E_0 with respect to the choice of orbitals provides one-electrons

HF equations:

$$f(i)\chi(\mathbf{x}_i) = \varepsilon\chi(\mathbf{x}_i) \tag{2.8}$$

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{1A}} + v^{HF}(i) = h(i) + v^{HF}(i) \tag{2.9}$$

where $f(i)$ is the one-electron Fock operator that incorporates the average electron– electron repulsion potential $v^{HF}(i)$. As the HF potential experienced by a certain electron, is created by other electrons, each described by their orbital, the HF equations depends on their own solutions and the model requires an iterative solution procedure known as the self-consistent field (SCF). When spin orbitals are expanded using basis functions, the HF equations become a matrix eigenvalue problem. A larger and more compete basis set allows for greater flexibility in expanding spin orbitals and lower energy expectation values, approaching the *Hartree-Fock limit*.

The HF model serves as a starting point that can be improved by adding further approximations (leading to semi-empirical methods), or including more determinants in the wavefunction description (multi-determinant methods). The latter takes into account electron correlation and, while computationally more expensive than HF, can produce results that progressively converge towards the exact solution of the Schrödinger equation. The next section illustrates the primary approaches for incorporating electron correlation into the wavefunction.

2.1.2 Post Hartree-Fock methods

There are three widely used approaches to the accurate inclusion of electron correlation in molecular applications[128,130–133]. These are configuration interaction[134], finite order many-body perturbation theory[135] and coupled-cluster methods[131,132,136].

In the configuration interaction (CI) ansatz, the N-electron system wavefunction may be written as:

$$\Psi_{\mathrm{CI}} = \Psi_0 + \sum_n \hat{C}_n |\Psi_0\rangle \tag{2.10}$$

where C_n is an excitation operator that introduces n-times excited configurations with respect to Ψ_0 (Hartree-Fock ground state reference) into the CI wavefunction, hence the name *configuration interaction*. The coefficients of N-electron expansion basis are determined variationally in order to give:

$$E_{\mathrm{CI}} = \langle \Psi_{\mathrm{CI}} | \hat{H} | \Psi_{\mathrm{CI}} \rangle \tag{2.11}$$

The lowest eigenvalue of the Hamiltonian matrix represents the exact non-relativistic ground state energy of the system, within the Born-Oppenheimer approximation. The correlation energy therefore is the difference between this exact energy and the Hartree-Fock energy for a given basis set.

When all excitation levels, up to the number of electrons, are included, the

result is denominated *full CI*. The full CI has several attractive properties. It offers the best possible upper bound to the exact energy in a chosen basis set; it is invariant to any transformation among the orbitals; and for separated molecules A and B it has the size-extensive property that $E_{\mathrm{CI}}(AB) = E_{\mathrm{fullCI}}(A) + E_{\mathrm{fullCI}}(B)$ for a non-interacting system. A full CI is infeasible for any but the simplest problems; in practice the expansion is truncated to a certain number of excited configurations, typically including only singly and doubly excited determinants (CISD). This method constitutes an improvement over the HF result as - for medium-size molecules and basis sets - it typically recovers 80–90% of the available correlation energy.

An alternative approach is offered by finite order *many-body perturbation theory* (MBPT). The idea in MBPT is partitioning the total Hamiltonian of the system into two pieces: a reference ($\hat{H}_0$), which has known eigenfuntions and eigenvalues, and a perturbation ($\hat{H}'$). The exact energy is then expressed as a sum of contributions of increasing complexity. E_{MBPT} is size-extensive to any-order perturbation theory but it is not variational. Furthermore, it has the deficiency of being of finite order, as all types of corrections to the reference wavefunction are added up to a given order. Infinite-order methods, on the other hand, have advantages in stability and invariance properties. The most popular MBPT method is *Møller–Plesset*[137] perturbation at the second order (MP2), where the reference Hamiltonian is taken as sum of Fock op-

erators (Equation 2.9). The inability of truncated CI to be size extensive, is closely associated with the poor convergence of the method toward the full CI limit. CI calculations demonstrated that, even though the effect of double excitations is always dominant, the second most important contribution is usually the quadruple excitations (rather than triple ones). This reflects the fact[138] that two types of quadruple excitation contributions are possible: the simultaneous interaction of four electrons (referred to as "connected") and two simultaneous interactions of two electrons (or "disconnected"). The latter is largely the most frequent (physically, this means that two simultaneous two-body interactions occur far more frequently than a true four-body interaction) and, thus, more important to the energy. Since these disconnected terms are fundamentally simple, their contribution simply being related to the double-excitation CI coefficients, this part of the quadruple effect can be included in a direct manner. By including these types of terms in a correlation model, convergence to the full CI limit should be better than in the truncated CI model. This procedure provides the basis for various pair theories, particularly within the framework of the *coupled cluster* (CC) approach, and allows to include parts of higher-excitations without the difficulties that are present in CI.

The CC Ansatz includes disconnected pair correlation by means of a exponential operator of the sum of connected pair excitations (**T**) such that the

wavefunction can be expressed as $e^{\mathbf{T}}|\Psi_0\rangle$ with the product terms in the expansion of the exponential operator representing the disconnected part. Hence, an alternative to the order-by-order expansion in $\hat{H}'$ of MBPT is to employ the exponential ansatz for the wavefunction:

$$|\Psi_{\mathrm{CC}}\rangle = e^{\mathbf{T}}|\Psi_0\rangle \tag{2.12}$$

building in approximations by restricting $\mathbf{T}$ to categories of electron cluster operators, $\mathbf{T}_n$. In a way, CC theory can be considered a more general infinite-order extension of MBPT, since it includes all corrections of a given type up to infinite order. Once $|\Psi_{CC}\rangle$ is known, the energy may be obtained from:

$$E_{\mathrm{CC}} = \langle\Psi_{\mathrm{CC}}|\hat{H}|\Psi_{\mathrm{CC}}\rangle \tag{2.13}$$

Analogously to the excitation operator $\hat{C}$, the cluster operator $\mathbf{T}$ can be decomposed into its one-electron, two-electron, and higher order cluster terms:

$$\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2 + \cdots + \mathbf{T}_n \tag{2.14}$$

The amplitudes (coefficients) of the $\mathbf{T}_n$ operator can be related to those of C_n in CI, where necessarily:

$$\Psi_{\mathrm{fullCI}} = \Psi_{\mathrm{CC}} \tag{2.15}$$

This implies a very important feature of coupled-cluster theory. Namely,

categories of CI excitations are further decomposed into clusters and their disconnected products. Since it is far easier to evaluate contributions from products of lower-order cluster operators that sum to n, CC theory offers a way to evaluate most of the contributions of higher-excitations while explicitly considering only relatively low n-values. The other primary advantage is that CC methods are size-extensive, as required by the exponential form of the wavefunction. Typically, both single and double excitations are included (CCSD), while the inclusion of connected triple excitations is computationally not affordable for most of the systems. A common approach is the so-called CCSD(T) method, which estimates the effect of the triples using perturbation theory, and includes a singles/triples coupling term[139,140]. The CCSD(T) model's remarkable effectiveness in the majority of cases has established it as the *de facto* gold standard for single-reference calculations.

The convergence of the electron correlation energy with respect to the increase of basis set size is extremely slow. A way to improve it is by including terms into the wave function that depend explicitly on the interelectronic distances (*explicitly correlated* methods or F12 theory[141]). The general Ansatz for closed-shell CCSD(T)-F12 wave functions employs excitation operators that include, apart from the conventional singles and doubles amplitudes, additional amplitudes relative to a complete orbital basis set. The additional

amplitudes include a correlation factor $\hat{F}_{12}$, taken to be a Slater function:

$$F(r_{12}) = exp(-\beta r_{12}) \tag{2.16}$$

where r_{12} is the interelectronic distance and β is a coefficient to be optimized. The optimum value of β varies for different molecules and basis sets: for the aug-cc-pVTZ basis, it ranges between $1.1 a_0^{-1}$ for pure hydrocarbons and $1.5 a_0^{-1}$ for molecules containing mainly oxygen atoms.

CCSD(T)-F12 has shown remarkably good performance for geometries of non-covalent interaction systems[142], which makes it especially suitable to optimize the model systems in study, namely, small molecules bound to minimal water clusters.

2.1.3 Density Functional Methods

Density Functional Theory (DFT) in the Kohn–Sham version can be considered as an improvement on HF theory, as it incorporates the many-body effects of electron correlation by means of a function that depends on the electron density[143,144].

The first Hohenberg-Kohn (HK) theorem[145] states that the electron density, $\rho(\mathbf{r})$, determines the external (i.e., due to the nuclei) potential, $V_{\text{ext}}(\mathbf{r})$ as well as it determines N, the total number of electrons, via its normalization. Since $V_{\text{ext}}(\mathbf{r})$ (and N) determines the molecular Hamiltonian (Equation 2.3), ρ(r) ultimately determines the system's energy, E. Energy - as all other system's electronic properties - can be expressed as functional of the density: $E = E_V[\rho]$. The theorem states in fact that ρ(r) corresponds to a specific and unique system (or electrons and nuclei). The second HK theorem provides a variational ansatz for obtaining $\rho(\mathbf{r})$, as the density that minimizes $E_V[\rho]$. The following is the DFT analogue of Schrödinger's time-independent equation:

$$V_{\text{ext}}(\mathbf{r}) + \frac{\delta F_{\text{HK}}}{\delta \rho(\mathbf{r})} = \Lambda \tag{2.17}$$

where Λ is the Lagrangian multiplier and F_{HK} is the *Hohenberg-Kohn functional*:

$$E_V[\rho] = \int \rho(\mathbf{r}) V_{\text{ext}}(\mathbf{r}) d\mathbf{r} + F_{\text{HK}}[\rho] \tag{2.18}$$

Where the first term is the potential energy due to the nuclei-electron attraction, and depend on the actual system, and F_{HK} which is universal (i.e. independent of N, and nuclear coordinate or charge) and contains the electronic kinetic energy functional, $T[\rho]$ and the electron-electron interaction functional, $V_{\mathrm{ee}}[\rho]$:

$$F_{\mathrm{HK}}[\rho] = T[\rho] + V_{\mathrm{ee}}[\rho] \tag{2.19}$$

The practical treatment of Equation 2.17 was provided by Kohn and Sham[146] (KS), who in analogy with HF theory, introduced the concept of orbitals. Their starting point was a N-electron non-interacting reference system with a certain Hamiltonian:

$$\hat{H}_{\mathrm{ref}} = \sum_{i}^{N} h_{\mathrm{ref},i} \tag{2.20}$$

and corresponding orbitals Ψ_i, eigenfunctions of the one-electron operator in Equation 2.20. Assuming that the reference system shares the same electron density as the real (interacting) system, such density can be written as:

$$\rho_{\mathrm{ref}} = \sum_{i}^{N} |\Psi_i|^2 \tag{2.21}$$

therefore, HK functional in Equation 2.19 is further factorized as:

$$F_{\mathrm{HK}}[\rho] = T_{\mathrm{ref}}[\rho] + J[\rho] + E_{\mathrm{XC}}[\rho] \tag{2.22}$$

where $T_{\mathrm{ref}}[\rho]$ represents the kinetic energy functional of the reference system,

$J[\rho]$ representing the classical Coulombic interaction energy and a last term $E_{\mathrm{XC}}[\rho]$ (the *exchange correlation energy*, containing the difference between the exact kinetic energy and $T_{\mathrm{ref}}[\rho]$, the nonclassical part of V_{ee}, and the self-interaction correction of J). Finally, the Euler equation 2.17 can be rewritten as:

$$\Lambda = V_{\mathrm{eff}}(\mathbf{r}) + \frac{\delta T_{\mathrm{ref}}}{\delta \rho} \tag{2.23}$$

which refers to a non-interacting N-electron system, with electrons being subjected to an external effective potential, $V_{\mathrm{eff}}(\mathbf{r})$. Thus, for a given $V_{\mathrm{eff}}(\mathbf{r})$, the molecular orbitals Ψ_i should satisfy the KS equations:

$$(-\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r}))\Psi_i = \varepsilon_i \Psi_i \tag{2.24}$$

as well as the orthonormality condition. Moreover, a variational choice of the orbitals yield the lowest energy for a given density.

The energy of the interacting system, written in terms of the separation described previously and highlighting the dependence on the orbitals, therefore

is:

$$\begin{aligned} E_V[\rho] &= \int \rho(\mathbf{r}) V_{\text{ext}}(\mathbf{r}) d\mathbf{r} + T_{\text{ref}}[\rho] + J[\rho] + E_{\text{XC}}[\rho] \\ &= \sum_i^N \int \sum_A^M \frac{Z_A}{\mathbf{r}_{1A}} |\Psi_i(\mathbf{r}_1)|^2 d\mathbf{r}_1 - \frac{1}{2} \sum_i^N \langle \Psi_i | \nabla^2 | \Psi_i \rangle \\ &+ \frac{1}{2} \sum_i^N \sum_j^N \int \int |\Psi_i(\mathbf{r}_1)|^2 \frac{1}{\mathbf{r}_{12}} |\Psi_i(\mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2 \\ &+ E_{\text{XC}}[\rho] \end{aligned} \tag{2.25}$$

The KS equations are one-electron equations, just as the HF equations, and have to be solved iteratively. But unlike the Hartree-Fock model, where the approximation is introduced right from the start (the wave function is assumed to be a single Slater determinant, which can never deliver the true solution) the KS approach is in principle exact. The electron correlation is accounted for in the effective potential, namely in a term called exchange correlation potential, $V_{\text{XC}}(\mathbf{r})$, which is defined as:

$$V_{\text{XC}} = \frac{\delta E_{\text{XC}}}{\delta \rho(\mathbf{r})} \tag{2.26}$$

The approximation only enters KS theory when we have to decide on an explicit form of the unknown functional for the exchange-correlation energy E_{XC} and the corresponding potential V_{XC}. The main disadvantage of DFT is that there is no systematic approach to improving the results towards the exact solution. Therefore, the quality of DFT calculation rests on the qual-

ity of the approximation used for E_{XC}. The simplest approach is the *Local Density Approximation* (LDA)[147], where the density is locally treated as a uniform electron gas - that largely overestimates the electron correlation. Significant improvements are given by *Generalized Gradient Approximation* (GGA)[148,149] methods, that consider the exchange and correlation energies to be dependent not only on the density but also on derivatives of the density. Very popular examples of this class are LYP (Lee-Yang-Parr)[150], PW86 (Perdew–Wang 1986)[151], PW91 (Perdew–Wang 1991)[152] and PBE (Perdew–Burke–Ernzerhof)[149]. The logical extension of GGA methods is to allow the exchange and correlation functionals to depend on higher order derivatives of the electron density, or to the orbital kinetic energy density, leading to to the so-called *Meta-GGA* functionals[153]. Among them, TPSS (Tao–Perdew–Staroverov–Scuseria)[154] can be considered as the next improvement over the PBE functional. Another class of XC-functionals are the *Hybrid-GGA* methods, that include a fraction of the exact exchange given by HF theory. Those methods also include parameters determined by fitting to experimental data; the most common are Becke 3-parameter functional methods[155], especially B3LYP and B3PW91 (both augmented with 20 % of exact exchange). Popular examples also are PBE0[156](the improved version of PBE, augmented with 25% of exact exchange) and TPSSh[157](TPSS augmented with 10% exact exchange). One last important class is the *long-range corrected* (LC) hybrid density functionals[158]. Those methods resolve qual-

itatively the self-interaction errors employing 100% Hartree–Fock exchange for long-range electron–electron interactions. LC methods are obtained *via* a systematic optimization procedure using a certain number of parameters. One example is ωB97X functional[159], where a extra parameter is added, corresponding to an adjustable fraction of short-range exact exchange.

2.1.4 *Ab initio* molecular dynamics

Molecular dynamics (MD)[160] is a widely-used method for studying chemical processes in condensed phases. It involves numerically solving classical Newtonian equations of motion for a system from a predefined initial state while considering specific boundary conditions. The accuracy of MD calculations heavily relies on the method used to specify forces acting on the system. In many cases, forces are determined using an empirical model or "force field," which has been highly successful in various applications, spanning from simple liquids and solids to complex systems like polymers and biological substances.

However, most force fields lack consideration of electronic polarization effects[161,162] and have limited capacity to address chemical reactivity[163], necessitating the use of ab initio molecular dynamics (AIMD) methods[162,164–166]. In an AIMD calculation, finite-temperature dynamical trajectories are generated by using forces obtained directly from electronic structure calculations performed “on the fly” as the simulation proceeds. Thus, AIMD permits

chemical bond breaking and forming events to occur and accounts for electronic polarization effect[164]. In its most ideal form - also called Born–Oppenheimer MD (BOMD) - a AIMD calculation for a system composed of M nuclei and N electrons, relies on three principles: that the Born–Oppenheimer and the adiabatic approximations are valid, and that the dynamics of the nuclei can be treated semi-classically on the ground-state electronic surface (Ψ_0). Among them, the Born–Oppenheimer approximation has been mentioned above (Section 2.1), while the adiabatic approximation[167], assumes that the electronic wave functions adapt quasi-instantaneously to a variation of the nuclear configuration and it is justified by the fact that the electrons are much lighter than the nucleus. The last approximation implies that the nuclei follow a Newton equation whose potential is determined by the Ehrenfest theorem:

$$M_A \ddot{\mathbf{R}}_A(t) = -\nabla_A \langle \hat{H}_e \rangle_{\Psi_e} \tag{2.27}$$

where it appears the aforementioned electronic Hamiltonian and wave function (Equation 2.4). BOMD is characterised by a Lagrangian which is defined as the difference between the kinetic and the potential nuclear energy (first and second terms, respectively):

$$L^{BO} = \frac{1}{2}\sum_A M_A \dot{\mathbf{R}}_A^2 - \langle \Psi_0 | \hat{H}_e | \Psi_0 \rangle + \sum_{i,j} \Lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij}) \tag{2.28}$$

the last term is a Lagrange function which ensures the orthogonality condition of the orbitals, as is not preserved by the stationary Schrödinger equation over time. The equations of motion associated with Equation 2.28 are obtained from the Euler–Lagrange equations, separately for the nuclear and electronic degrees of freedom:

$$\begin{aligned} M_A\ddot{\mathbf{R}}_A(t) &= -\nabla_A \min_{\Psi} \langle \Psi_0 | \hat{H}_e | \Psi_0 \rangle \\ \hat{H}_e \psi_i(\mathbf{r}) &= \sum_j \Lambda_{ij} \psi_i(\mathbf{r}) \end{aligned} \tag{2.29}$$

Kohn–Sham formulation (Kohn–Sham MD) introduces a further approximation in order to reduce the complexity of the electronic Hamiltonian, replacing Ehrenfest term in the nuclear part of Equation 2.29 by the KS energy:

$$\min_{\Psi_0} \langle \Psi_0 | \hat{H}_e | \Psi_0 \rangle = \min_{\Psi} E^{KS}[\rho] \tag{2.30}$$

So basically, Equation 2.29 is integrated numerically by using a symplectic integrator, and, at each timestep, the forces are computed by minimizing the KS energy functional (Equation 2.25) at the present nuclear configuration.

2.2 Gaussian Moments Neural Network

Machine learning techniques can provide a efficient approximation of the PES of a system as they allow a direct mapping of atomic positions and nuclear charges with almost *ab initio* accuracy and the computational efficiency of empirical potentials. Moreover, machine learning tools can be applied to large scale systems, when molecular dynamics simulations are currently unfeasible. In this thesis, we adopted a recently proposed Machine Learning Potential based on physically inspired molecular descriptors, denominated Gaussian Moments (GM)[168,169]. As the used Machine Learning technique belongs to the class of feed-forward neural networks (NN), the Potential is referred to as GM-NN. The choice of molecular descriptor to be used as input to the NN algorithm is particularly important, since the NN methodology does not exploit any information about the physics of the problem. Advantage of the GM is that, in addition to the structural description, it encodes the information about the atomic species in the molecular representation. Therefore, a single NN is trained and used for all atomic species, in contrast to using an individual NN for each species as frequently necessary previously[170]. GM-NN model is based on tensor contractions, which fulfills physical constraints and are inspired by the molecular wave function. The main idea is splitting the descriptor for the whole chemical system into functions which describe

the environment of each atom individually:

$$\phi_{s,l_x,l_y,l_z}(\mathbf{r}) = \frac{x^{l_x} y^{l_y} z^{l_z}}{r^L} \Phi_s(r) \tag{2.31}$$

with **r** =(x, y, z) being position of a central atom. The prefactor $x^{l_x} y^{l_y} z^{l_z}/r^L$ covers the angular dependence of the function ($L = l_x + l_y + l_z$, angular momentum), while the radial part is:

$$\Phi_s(r) = \left(\frac{2N_{Gauss}^2}{\pi R_{max}^2} \right)^{1/4} e^{-N_{Gauss}/R_{max}^2 (r-\gamma_s)^2} f_{cut}(r) \tag{2.32}$$

and depends on the total number N_{Gauss} of gaussian functions (each one is centered at γ_s, s ranging from 1 to N_{Gauss}). and incorporates a cutoff function $f_{cut}(r)$, which restricts the descriptor to the local neighborhood of the atom and decays smoothly to zero at the cutoff radius R_{max}. Those atomic functions can be linearly combined so that the total "wave" function reads:

$$\Psi_{i,L,s} = \sum_{j \neq i}^{N_{atom}} \beta_{Z_i,Z_j,s} \phi_{s,l_x,l_y,l_z}(\mathbf{r}_{ij}) = \sum_{j \neq i}^{N_{atom}} \beta_{Z_i,Z_j,s} T_{i_1,\cdots,i_L} \frac{1}{L} \Phi_s(r_{ij}) \tag{2.33}$$

where a Cartesian tensor $T_{i_1,\cdots,i_L}$ is introduced in order to fulfill the invariances requirements of the descriptor with respect to translation, rotation, and permutation of like atoms.The dependence on the central atom species is ensured *via* the coefficients $\beta_{Z_i,Z_j,s}$, which are optimized during training. The invariant scalars obtained by contracting $\Psi_{i,L,s}$ corresponds to the GM repre-

sentation $\rho_{i,\dots}$. This results in the expression for the total energy:

$$E = \sum_{i=1}^{N_{atom}} NN(\mathbf{x}_{in}^{i} = \{\rho_i, \dots\}) \tag{2.34}$$

The model is trained on energies and forces. Incorporating forces into the training is pivotal, since accurate force prediction is required for geometry optimizations and MD simulations, furthermore, obtaining forces for all atoms is relatively cheap, as they are calculated analytically from the total energy, by taking the partial derivative with respect to atomic positions. Details about the training procedure and NN quality tests are reported in Appendix Section B.

2.3 Transition state search and characterization, minimum energy path

Finding and characterizing minima on the $3N_{atom}$-dimensional PES of a system is a complex problem[128]. The Born-Oppenheimer approximation of the total energy (Equation 2.5) is a continuous function of the $3N_{atom}$ nuclear coordinates $E(\mathbf{X})$ where $\mathbf{X} = (X_1, \ldots, X_{3N_{atom}})$. For a given stationary point of geometry $\mathbf{X}_0$, the energy can be approximated as as a second-order Taylor expansion around a $\mathbf{X}_0$:

$$E(\Delta\mathbf{X}) = \frac{1}{2}(\Delta\mathbf{X})^t\mathbf{H}(\Delta\mathbf{X}) \tag{2.35}$$

since the first derivative is zero and $E(\mathbf{X}_0)$ has been chosen as zero as well. $\mathbf{H}$ is the $3N_{atom} \times 3N_{atom}$ *Hessian* matrix containing the second derivatives of the energy with respect to the coordinates. Equation 2.35 is substituted in the Schrödinger nuclear equation, that is then conveniently transformed to mass-dependent coordinates by a $\mathbf{G}$ matrix containing the inverse square root of atomic masses. A unitary transformation is then introduced to diagonalize the entry wise product $\mathbf{H} \cdot \mathbf{G}$ matrix, yielding eigenvalues ε_i and eigenvectors

$\mathbf{q}_i$, in the *vibrational normal coordinates* ($\mathbf{q}-$coordinate) system:

$$\left[\sum_{i=1}^{3N_{atom}}\left(\frac{1}{2}\frac{\partial^2}{\partial q_i^2}+\frac{1}{2}\varepsilon_i q_i^2\right)\right]\Psi_{nucl}=E\Psi_{nucl} \tag{2.36}$$

ε_i are related to the vibrational frequencies as:

$$\nu_i=\frac{1}{2\pi}\sqrt{\varepsilon_i} \tag{2.37}$$

If the stationary point is a minimum on the PES (reactants, products and intermediates), the ε_i are all positive. If, however, the stationary point is a transition state (TS), one of the eigenvalues is negative. This situation corresponds to the energy being a maximum in one direction and a minimum in all others. The frequency for the vibration along the eigenvector with a negative eigenvalue will formally be imaginary, and such eigenvector corresponds to the direction leading downhill from the TS towards the reactant and product, i.e. the reaction coordinate.

There are several approaches to find TS structures, the ones adopted in this thesis fall into the category of interpolation methods, such that the reactant and product geometries are known, and that the TS is located somewhere "between" these two end-points. The simplest of those method is usually called *coordinate driving* and consists in selecting one or a few internal "reaction" coordinates, that are suitable to describe the changes happening during

the reaction (e.g. a bond distance for a bond breaking/forming reaction). The selected coordinate is fixed at certain values, while the remaining variables are optimized, thereby adiabatically mapping the energy as a function of the reaction variable. If the variables choice is good, the constrained optimized geometry with the smallest residual gradient is a good approximation to the transition state.

Methods that operate with multiple structures or 'images' connecting the reactant and product, are often called *chain-of-state methods*. Relaxation of the images will not only lead to the saddle point, but also to an approximation of the whole reaction path. Among them, one of the most popular is the Nudged Elastic Band (NEB)[171] method. It consists of defining a target function ("elastic band") as the sum of energies of all images and adding a penalty term having the purpose of distributing the points along the path:

$$T_{\mathrm{NEB}}(\mathbf{R}, \mathbf{x}_1, \ldots, \mathbf{x}_M, \mathbf{P}) = \sum_{i=1}^{M} E(\mathbf{x}_i) + \sum_{i=1}^{M-1} \frac{1}{2} k (\mathbf{x}_{i+1} - \mathbf{x}_i)^2 \tag{2.38}$$

where **R** and **P** represent reactants and product geometries respectively and the spring constant k distributes the images evenly along the path.

Once the TS has been found, there are several methods to locate the reaction path[172–174]. The most widespread one is tracing the *Intrinsic Reac-*

tion Coordinate (IRC)[175], which corresponds to the minimum energy path in mass-weighted coordinates, from the TS to the reactant and product, and it will allow to verify if it indeed connects the desired minima. The IRC path is defined by the following differential equation:

$$\frac{\mathrm{d}\mathbf{x}(s)}{\mathrm{d}s} = -\frac{\mathbf{g}}{|\mathbf{g}|} = \mathbf{t} \tag{2.39}$$

where $\mathbf{x}$ are the mass-weighted coordinates, s is the path length and $\mathbf{t}$ is the negative normalized gradient. Determining the IRC requires solving eq. 2.39, starting from a geometry slightly displaced from the TS, along the normal coordinate for the imaginary frequency.

2.4 Reaction force and partition of the reaction barrier

Useful information about the mechanism of a chemical process can be obtained from the reaction force profile. For a certain process, the potential energy $E(\xi)$ of the system along the intrinsic reaction coordinate ξ has an associated reaction force $F(\xi)$, defined by:

$$F(\xi) = -\frac{dE}{d\xi} \tag{2.40}$$

It has been shown[176] that the critical points of $F(\xi)$ define regions along ξ in which different reactive events take place. The reactive events are identified as inflection points of the reaction force profile. Within each reaction event i it is possible to identify three regions limited within the critical points present in $F(\xi)$, namely the local minimum, at $\xi_{min,i}$, and local maximum, at $\xi_{max,i}$. The pre- and post- event region ($\xi \leq \xi_{min}$, $\xi_{max} \leq \xi$) are characterized by structural preparation or relaxation of the participating species. On the other hand, the event region itself is governed by changes in the electron density associated to bond formation and dissociation processes.

In general, a single kinetic step of a chemical reaction can be composed of N_{events} different reaction events such that the total reaction energy corre-

sponds to the sum of the energy of the individual reactive events:

$$\Delta E^{o} = \sum_{i=1}^{N_{events}} \Delta E_i \tag{2.41}$$

The event energy ΔE_i will be positive if the event takes place before the TS and negative if takes places passed the TS. Conversely, the TS energy barrier ($\Delta E^{\ddagger}$) is partitioned among the events leading up to the TS of a given reaction

$$\Delta E^{\ddagger} = \int_{\xi_0}^{\xi_K} F(\xi) d\xi = \sum_{i=1}^{K} \Delta E_i^{\ddagger} \tag{2.42}$$

where K is the number of events that take place before reaching the TS, ξ_0 and ξ_K corresponds to the reaction coordinate of the reactants and TS respectively.

The energy required for a certain event i can in turn be obtained by integrating the reaction force profile:

$$\Delta E_i^{\ddagger} = \int_{\xi_{i-1}}^{\xi_i} F(\xi) d\xi \tag{2.43}$$

Events leading up to the TS are denominated hidden transition states (h-TS)[177], which are associated to hidden intermediates (h-I). Such intermediates are not observable along the reaction energy profile and hence can only be estimated to lie between the h-TS and TS.

Such integration of the force for the reaction events intervals, allows to esti-

mate the energy expenditure associated to each phase of the chemical process and proved to be a valuable partition for isolating the chemical changes.

2.5 Binding energy

The majority of the small interstellar species physisorb on icy-grains, such non-covalent interaction is accounted for by the binding energy (BE). The BE of a species (i) physisorbed on a surface (ice) is defined as:

$$BE_i = E_{sup} - (E_{ice} + E_i) \tag{2.44}$$

where E_{sup} stands for the energy of the supermolecule formed by the adsorbate bound to the surface, E_{ice} is referred to the surface energy and E_i is the energy of the adsorbate. The BE is assumed to be a positive quantity, according to convention.

2.5.1 Binding energy calculation

The BE has been calculated as:

$$\Delta E_b = \Delta E_{CP} + \Delta_{ZPVE} \tag{2.45}$$

with ΔE_{CP} being the binding electronic energy corrected for the basis set superposition error (BSSE) and ΔE_{ZPVE} the zero-point vibrational energy (ZPVE) correction for the BE. In the following, the electronic energy of a molecule M in the geometry G computed with the basis γ, is defined as

$E_M^G(\gamma)$. Considering this notation, the BE of a species i with a basis set χ on a surface *ice* with a basis set ω can be calculated as:

$$\Delta E_e = E_{sup}^{sup}(\chi \cup \omega) - (E_i^i(\chi) + E_{ice}^{ice}(\omega)) \tag{2.46}$$

where *sup* represents the geometry of the supermolecule as in Equation 2.44. However, this expression does not consider that the basis function centered at *ice* assists in lowering the energy of fragment i and vice versa, resulting in a lower electronic energy of the supermolecule ($E_{sup}^{sup}(\chi \cup \omega)$) and hence an overestimation of the BE. This effect is commonly known as basis set superposition error. A way to correct for the error is the so-called counterpoise method (CP)[178], that considers the energy of the fragments in the geometry of the supermolecule with the basis of the respective partner. Thus the correction is calculated as:

$$\Delta_{CP} = E_i^{sup}(\chi \cup \omega) - E_i^{sup}(\chi) + E_{ice}^{sup}(\chi \cup \omega) - E_{ice}^{sup}(\omega) \tag{2.47}$$

Such that the resulting BE is:

$$\Delta E_{CP} = \Delta E_e - \Delta_{CP} \tag{2.48}$$

It is important to notice that at the complete basis set(CBS) limit, the correction term is zero since, χ, ω and $\chi \cup \omega$ are the same.

The ZPVE (Δ_{ZPVE}) contribution of the BE of i is obtained by computing the Hessian matrix at the equilibrium geometry at the same level of theory as the geometry optimization.

$$\Delta_{ZPVE} = ZPVE_{sup} - (ZPVE_i + ZPVE_{ice}) \tag{2.49}$$

Due to computational cost, the correction has been calculated for a selected group of binding sites of each species and a linear model has been used to correct ΔE_{CP}:

$$\Delta E_{CP} + \Delta_{ZPVE} = m\Delta E_{CP} + b \tag{2.50}$$

with m and b being the ZPVE correction factors. Finally, the factors are applied to the complete set of computed BEs for each species in order to derive the ZPVE corrected BE distribution.

2.6 Non-covalent Interaction Energy

BE of a species i can be further decomposed into physically meaningful terms:

$$BE_i = DE_i - IE_i \tag{2.51}$$

where DE_i and IE_i represents its *deformation energy* (DE) and *interaction energy* (IE), respectively. Usually the geometry of each interacting fragment is distorted in the environment of its neighbours. In the case of interstellar physisorption, the deformation energy quantitatively characterises the structural changes of the adsorbate and the ice surface with respect to the isolated fragments and is defined as:

$$DE_i = E_i^{sup} - E_i^0 \tag{2.52}$$

where E_i^{sup} the energy of the adsorbate in the supermolecular (distorted) configuration and E_i^0 is the energy of the isolated molecule.
Conversely, the interaction energy is a measure of the strenght of the non-covalent intermolecular interaction between the fragments. Thus, knowledge of the DE allows to calculate the IE indirectly through Equation 2.51.

IE is dominated by intermolecular forces that arise from physical phenomena responsible for attraction and repulsion between molecules and often

involve many-body effects. Symmetry-adapted perturbation theory (SAPT)[179] provides a means of directly computing the non-covalent interaction between two fragments, M_1 and M_2 (also called 'monomers'), that is, the interaction energy is determined without computing the total energy of the monomers and the dimer (i.e. complex formed by the two fragments together). In addition, SAPT provides a decomposition of the interaction energy into physically meaningful components. In SAPT, the Hamiltonian of the dimer is partitioned into contributions from each monomer and the interaction between them.

$$\hat{H} = (f_{M_1} + W_{M_1}) + (f_{M_2} + W_{M_2}) + V_{M_1 M_2} \tag{2.53}$$

Here, the Hamiltonian is written as a sum of the usual monomer Fock operators, F, the fluctuation potential of each monomer, W, and the interaction potential, V. The monomer Fock operators, are treated as the zeroth-order Hamiltonian and the interaction energy is evaluated through a perturbative expansion of W_{M_1}, W_{M_2} and $V_{M_1 M_2}$. Several truncations of the closed-shell SAPT expansion are available, the simplest is denoted SAPT0[180]. The performance of SAPT0 relies entirely on error cancellation, which seems to be optimal with a truncated aug-cc-pVDZ basis, namely, jun-cc-pVDZ, while it is not advised to use SAPT0 with larger basis sets.

2.7 Natural Atomic Orbital Population Analysis and Bond Order

The idea in the Natural Atomic Orbital (NAO) analysis[181], is to use the one-electron density matrix for defining the shape of the atomic orbitals in the molecular environment, and to derive molecular bonds from electron density between atoms. The starting point of the procedure is obtaining a density matrix written in terms of blocks of basis functions belonging to a specific centre. That can be achieved by arranging the basis functions such that all orbitals located on the specific centre are in sequence.

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}^{AA} & \mathbf{P}^{AB} & \vdots \\ \mathbf{P}^{AB} & \mathbf{P}^{BB} & \vdots \\ \cdots & \cdots & \ddots \end{pmatrix} \tag{2.54}$$

Therefore, the NAOs for atom A in the molecular environment may be defined as those that diagonalize the corresponding block of the density matrix ($\mathbf{P}^{AA}$), an so on for each atoms. Once orthogonalized, the final set of orbitals are simply denoted NAOs, and the diagonal elements of the density matrix in this basis are the orbital populations. Summing all contributions from orbitals belonging to a specific centre produces the NAO atomic charge,

while bonds between atoms may be identified from the off-diagonal blocks. The bond order is defined from the Wiberg index as the sum of squares of off-diagonal density matrix elements between atoms:

$$W_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} P_{\mu\nu}^2 \tag{2.55}$$

where summation is over atomic orbitals μ on atom A and atomic orbitals ν on atom B, and $P_{\mu\nu}$ is the corresponding density matrix element. W_{AB} corresponds to the number of covalent bonds formed by that atom, corrected for the ionic character in each bond. Advantages of using the NAOs are given by the fact that they account for 99+% of the electron density, that the electron occupation is guaranteed to be between 0 and 2, and that they converge to well-defined values as the size of the basis set is increased, which is not the case for other population analysis. A drawback is that the NAOs may extend significantly beyond the atom from which they originate, potentially leading to inaccuracies in describing the electron density of a particular centre.

2.8 Maps of Electrostatic Potential

A way to predict the portions of molecules which are relatively electron rich (i.e., negative) and others that are correspondingly electron poor (i.e., positive), is using electronegativity scales and atomic charges. There are several schemes to assign charges to individual atoms (See Murray and Politzer [182] and references therein), but they have the problems, however, that there is no rigorous physical basis for them and they cannot be measured experimentally. In order to avoid the arbitrariness it is possible to focus on the system's electrostatic potential $V(\mathbf{r})$ instead. This is a three-dimensional local property that arises from the Coulomb's law. $V(\mathbf{r})$ can be evaluated at any or all points in the space of a system, it has a rigorous physical definition, and can be determined experimentally and computationally. The following expression, in atomic units of energy, indicates the electrostatic potential at a point $\mathbf{r}$ created by a system of nuclei and electrons:

$$V(\mathbf{r}) = \sum \frac{\mathrm{Z}}{|\mathbf{R}-\mathbf{r}|} - \int \frac{\rho(\mathbf{r}_i)\mathrm{d}\mathbf{r}_i}{|\mathbf{r}_i-\mathbf{r}|} \tag{2.56}$$

The nuclei are treated as stationary point charges and the electronic density function, $\rho(\mathbf{r})$, expresses the average number of electrons per volume element. Z is the charge on a certain nucleus, located at $\mathbf{R}$; $|\mathbf{R}-\mathbf{r}|$ represents its distance from $\mathbf{r}$, just as $|\mathbf{r}_i-\mathbf{r}|$ is the distance of each electronic charge

increment $\rho(\mathbf{r}_i)\mathrm{d}\mathbf{r}_i$ from $\mathbf{r}$.

$V(\mathbf{r})$ is normally computed and shown on a closely spaced grid covering a 3D outer surface of the molecule (a map of the electrostatic potential, MEP). Choosing the surface to be an outer contour of its electronic density $\rho(\mathbf{r})$ - typically 0.001 au - has the advantage that it reflects the specific features of the particular molecule, e.g., lone pairs etc. The MEP is a reflection of how the molecule is 'seen' by an approaching reactant, and thus is a useful guide to the molecule's reactive behavior, especially in non-covalent interactions.

2.9 Computational Details

QCArchive Framework

Quantum chemistry data has been traditionally generated through user-defined individual input files, which are processed by a specific software that stores the results of the computation in output files. These outputs are then parsed either by hand or using custom scripts. This approach has serious limitations when attempting to compute a large volume of data as it is error-prone and the results are difficult to reproduce, since parsing scripts and output files are usually not available. To overcome these limitations, the Binding Energy Evaluation platform (BEEP) is built within the Python-based QCArchive framework. The details about the different components of the QCArchive infrastructure have been described elsewhere[183]. The core component of BEEP is a central server to which computation results are added in the form of JSON (JavaScript Object Notation) objects that contain the same level of information as a traditional output file. The access to this database, where the user can query existing data and submit additional computations, is controlled by a standard username/password system. Moreover, several data objects can be defined to generate and sort the data. These collections (called *Datasets*) make it possible to extend a procedure, such as a geometry optimization or a BE computation, to a large number of objects in a single operation. Finally,

the generated values can be easily accessed from the stored collections.

Software packages

Ices modelling: "amorphization" AIMD simulations on the cluster models are performed at BLYP/def2-SVP level, adding D3 Grimme[184] correction for dispersion interactions, as implemented in TERACHEM.[185,186]. *BE evaluation*: High level CCSD(T)-F12 optimizations are performed using MOLPRO[187]. Binding sites optimization and binding energy computations are performed using PSI4[188]. *NN training and results*: To run MD simulations on the periodic systems, the GM-NN program[168] was interfaced to the ASE package[189]. ORCA[190] software is used to compute DFT energies and gradients that constitute NN training set. *Reactivity study*: minimum energy structures have been identified and optimized using ORCA. IRC calculations are performed with GAUSSIAN[191]. Analysis of the IRC profiles have been carried out using Kudi[192]. NAO population analysis use the NBO software version[193] implemented in GAUSSIAN and ORCA.

3. Results and Discussion

The chapter illustrates the most relevant results obtained in this thesis. The first section (Section 3.1) illustrates the modelling strategies to generate different ice surfaces. Section 3.2 presents the DFT geometry and energy benchmark results. Afterwards, the estimations of binding energies of small molecules adsorbed on interstellar ices are shown (Section 3.3). Sections 3.4, 3.5, are dedicated to the results of applying the novel multi-binding approach to the study of reactions associated to the Strecker synthesis of glycine. The final Section 3.5.5 of the chapter encompasses the general discussion.

3.1 Interstellar ices modeling

3.1.1 Set of homogeneous amorphous clusters

The ice surfaces used for desorption and reactivity studies are modelled according to the *cluster approach*. The cluster size (22 water molecules) is selected in order to guarantee a reasonable number of available binding sites, while at the same time being able to use high-level model chemistry.

The steps to obtain the homogeneous amorphous water models are the fol-

lowing:

1. Generation of the initial water cluster, by MD tecniques, using the TIP3P[194] model.

2. Amorphization of the surface, performing an high temperature (300 K) AIMD simulation on the cluster, at BLYP/def2-SVP[148,150,195,196] level of theory. AIMD has been preferred over molecular dynamics for ice simulations, since the former generates interaction potentials using quantum-chemical methods, while common empirical water force fields are parameterized using properties of liquid water[197,198]. The simulation length of 100 ps assured correct system equilibration, while the choice of time-step of 1 fs accounted for the high frequency of the O-H vibrations of the water substrate. The temperature has been controlled using a Langevin thermostat[199]. Finally, spherical periodic boundary conditions are applied in order to avoid evaporation effects.

3. Extraction of 100 independent structures from the resulting trajectory. For the selection a Kendall's $\tau_{correlation} \simeq 1ps$ between the trajectory frames is used. The structures underwent temperature annealing of 10 ps to reach the target interstellar conditions of $\sim$10 K.

4. Selection of the 20 most representative structures, according to geometrical criteria. The similarity threshold has been assigned using the root-mean-square deviation of atomic positions (RMSD), using the Kabsch[200]

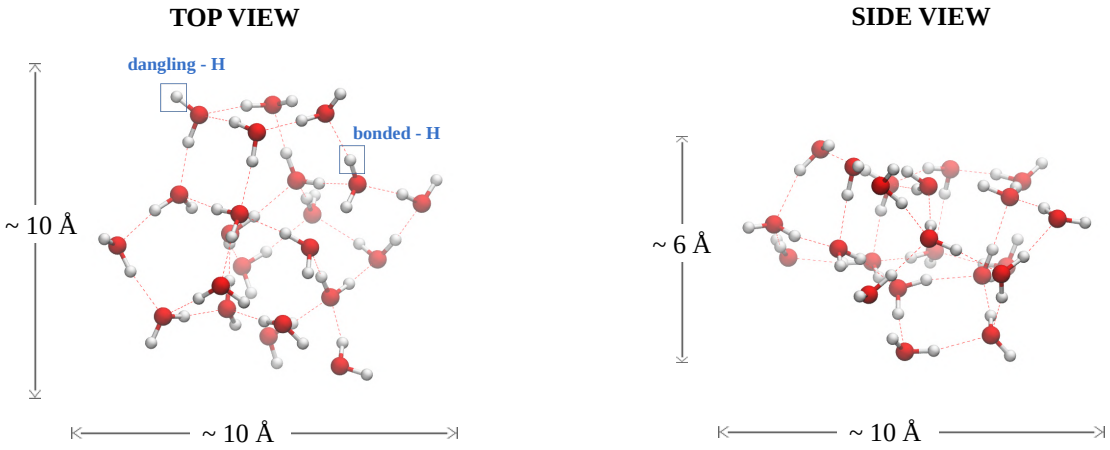


Figure 3.1: Top (left) and side (right) view of one of the homogeneous 22-water molecules amorphous clusters used in this thesis for binding energy evaluation and reactivity studies. After modelling through AIMD technique, the structures undergo geometry optimization. The surface molecules are separated in two classes: the ones that are coordinated with four water molecules, hence, each H atom is engaged in HB with a neighbor ("bonded-H"), and surface molecules that lack a HB-coordination, generating the so-called "dangling-Hs", that are not participating in any HB.

algorithm. Hence, assuming a similarity threshold of RMSD $\leq$ 0.4 Å, the annealed structures are grouped and the 20 most populated clusters - i.e. the structures that present the high number of copies - are selected.

5. Geometry optimization of the set of clusters, to be used in the successive steps of the work.

One of the clusters is displayed in Figure 3.1. The surface spanned by these 20 clusters represents the amorphous model, denominated ASW_{22} or simply ASW.

3.1.2 Homogeneous models of various sizes: $ASW_{12,37,60}$ clusters

Analogous procedure has been followed using different initial cluster sizes, generally denoted as ASW_S with S = 12, 37 , 60 number of water molecules. For ASW_{60} a shorter simulation length of 50 ps is used, due to the demanding computational time. Figure 3.2 reports an example taken from each set of clusters.

3.1.3 Amorphous ice mixture: $H_2O : NH_3 = 2.5 : 1$

In order to study a substrate other than pure water ice, a homogeneous "ice mixture" cluster is modelled. NH_3-rich ices are reported to exist in outer Solar System objects such as the moons of Saturn, Uranus and Pluto, dwarf planets and in comets[201–203]. Interstellar H_2O:NH_3 ice analogs have been extensively studied[116,119,204–210]. In this thesis, such ice mixture is used to simulate the first step of the Strecker synthesis. Experimental works[116,119] indeed reported that aminomethanol is the dominant product of the reaction when NH_3 is in excess with respect to H_2CO, hence the importance to include such kind of ices in the analysis.

The literature[116] indicates a optimal ratio of H_2O:NH_3 = 2:1. In order to avoid agglomerations of NH_3 molecules on the surface, and to maintain a system

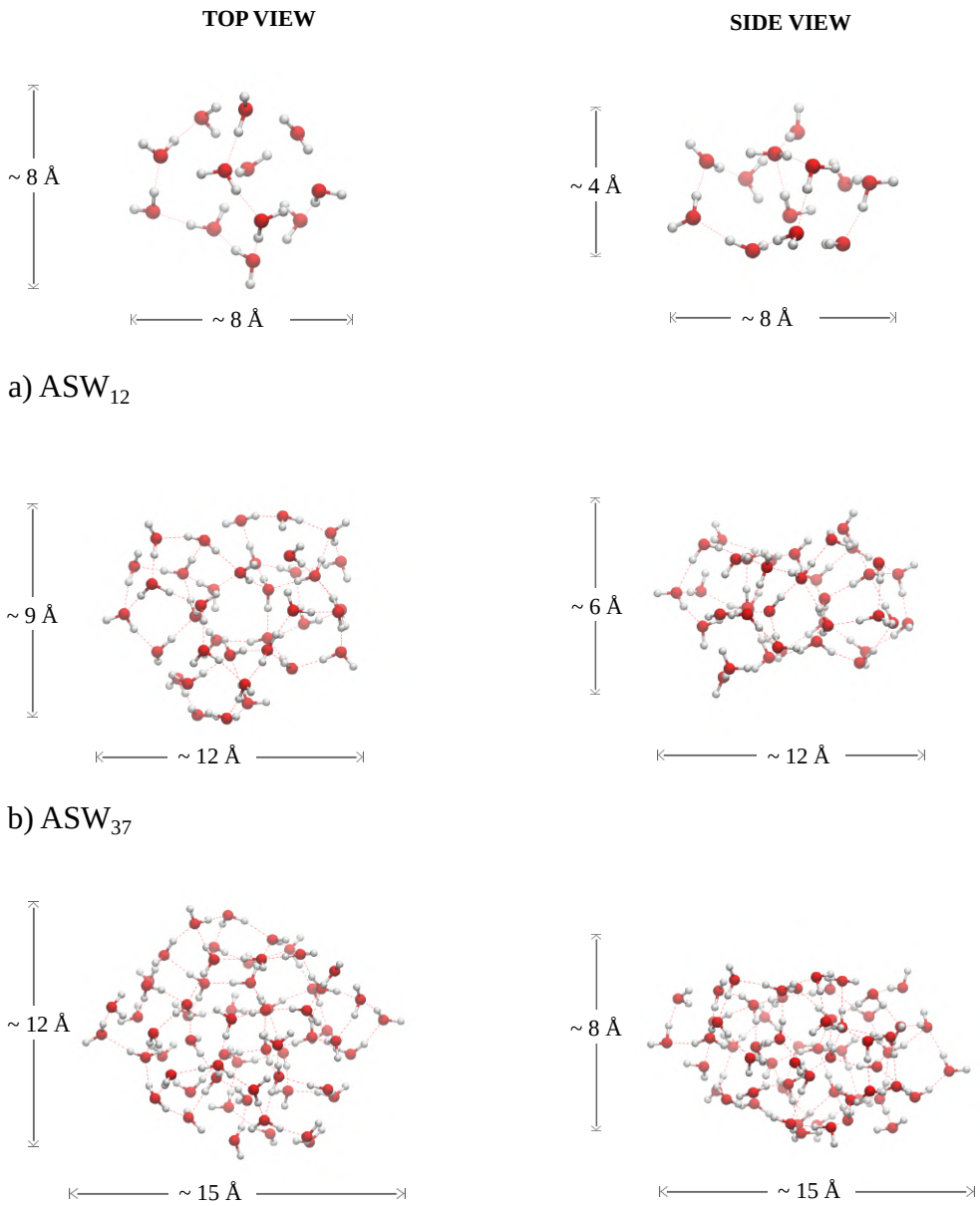


Figure 3.2: Top (left) and side (right) view of one structures from each set of amorphous clusters, denominated ASW_S where S is the number of water molecules ·

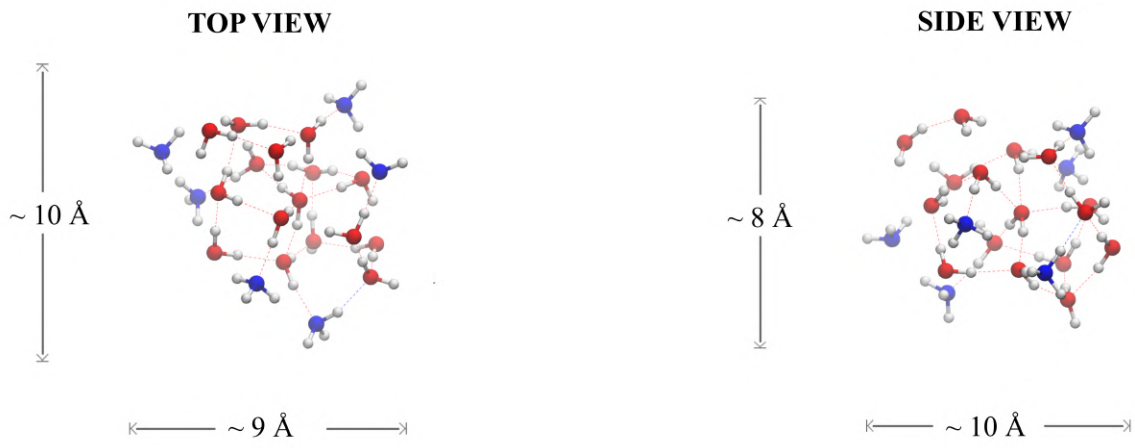


Figure 3.3: Top (left) and side (right) view of the ice mixture model. Composition ratio is reported in the figure. Color code is: red for O, blue for N and white for H.

size similar to the homogeneous pure water clusters, the ratio has been altered to H_2O:NH_3 = 2.5:1, so that the final surface composition is: $(H_2O)_{15}$: $(NH_3)_6$. The model is generated according to the computational procedure in Section 3.1.1, selecting only the structure corresponding to the last point of the dynamics trajectory. The optimized structure can be found in Figure 3.3.

3.1.4 Periodic ice: ASW_{500}

A *ad hoc* trained Machine Learned Potential (MLP) (see Section B.1) is used to built a set of 5 periodic ice models composed of 500-water molecule each. The choice of such model size allows to account for the diverse mor-

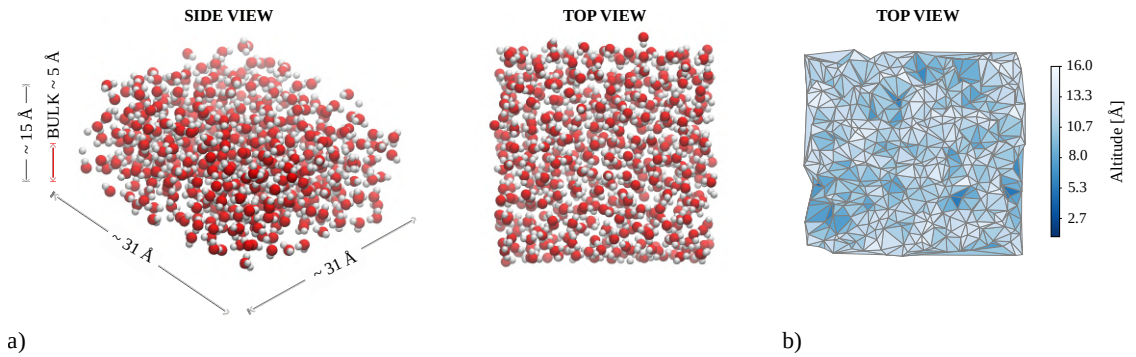


Figure 3.4: a) Side and top view of one of the 500-molecules periodic surfaces generated in this work. A constraint is applied to the position of the molecules at the bottom of the surface, representing the *frozen bulk* of the model. b) Tri-Surface altitude plot relative to the top view of the same surface. Darker regions represent "valleys" while "crests" are in lighter color.

phological and energetic characteristics of the binding sites on a more realistic ice. The initial 3D cell has volume (X $\times$ X $\times$ X/2) and cells dimensions (X) according to the molecular density value: 0.9982 g cm^{-3}. After minimization, the system was equilibrated in conditions of constant number of molecules, volume and temperature (NVT), for 100 ps at 300 K. All the simulations are carried out using a Langevin thermostat. 5 structures are extracted from the resulting trajectories ($\tau_{correlation} \simeq 20$ ps), and underwent temperature annealing of 10 ps to reach $\sim$ 10 K. Periodic boundary conditions are applied in two directions along the surface and constraints are imposed to the position of the molecules in proximity of the cell's lower face (covering 1/3 of the height), as to reproduce the *bulk* of the ice models. The applied constraint ensures that the specific structural diversity of the surfaces is pre-

served. The frozen bulk condition is maintained through all the following calculations. One of the periodic surfaces is reported in Figure 3.4a. In order to visually appreciate the presence of concave regions ('valleys'), porous and 'crests' on the surface, Figure 3.4b includes a altitude map (see also Section D.2 for details about Tri-Surface plots generation).

3.1.5 Nano-porous model

ASW in interstellar environments is thought to be partly porous[33]. Such porous sites could facilitate chemical encounters of species adsorbed on the ice and enhance surface reactivity. In order to generate a nano-porous, a periodic surface with lower initial density (0.8 g cm^{-3}) has been used as starting point. The periodic slab ASW_{500} has been generated as in Section 3.1.4 and it is reported in Figure 3.5a (top view). It can be evinced from the altitude map that the initial surface presented several suitable porous sites of nanometric size (in darker). A spherical portion of one of the sites with the necessary characteristics has been extracted (highlighted in Figure 3.5a) consisting of 64 water molecules. A constraint has been applied to the atoms belonging to the outer sphere, in order to preserve the porous during geometry optimization. Figure 3.5b, shows the nano-porous after the optimization: the atoms that are kept frozen are represented as sticks. The constrained is maintained throughout each phase of the work.

3.1.6 Surfaces characterization

Cluster models

In terms of shape, the medium-size clusters (ASW composed of 12 and 22-water molecules and the ice mixture of 21 molecules in total) are mostly spherical, while the larger clusters (ASW of 37 and 60 molecules) are oval-like shaped. The different dimensions are indicated in Figure 3.1, 3.2, 3.3. As the clusters get bigger, the amount of molecules belonging to the bulk of the system increases, i.e. inner layers are generated.

Binding sites on each surface can be described in terms of water orientation. Unlike the bulk molecules, which typically form four hydrogen bonds (HB) with neighbors, surface molecules can have two possible orientations. One orientation is given by molecules having both OH groups pointing obliquely down, engaged in HB with surface neighbors or with molecules belonging to inner layers. The hydrogens of such OH groups are labelled as "bonded-Hs". Other surface molecules have one of their OH groups pointing upwards away from the surface, not engaged in any HB. Such hydrogens are labelled “dangling-Hs”. Examples of dangling-H and bonded-H sites are highlighted in Figure 3.1. Dangling-H sites are a measure of the porosity of ice as they are observed to increase in experimental porous ASW with respect to non-porous ASW[33,211]. Moreover, dangling-Hs act as sites of preferential

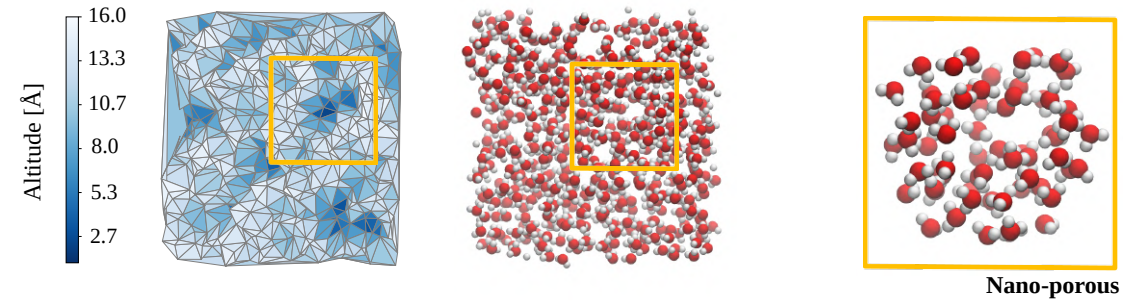
adsorption and reaction[212].

The ice models have been characterized in terms of dangling-Hs frequency, yielding the following results. The homogeneous clusters ASW_{12}, ASW_{22}, ASW_{37}, ASW_{60}, present the average number of 5, 8, 10, 11 dangling-Hs per cluster, respectively. Larger clusters have less dangling-Hs per unit surface, although, a proper analysis would require to evaluate how many unique binding sites exist for each model.

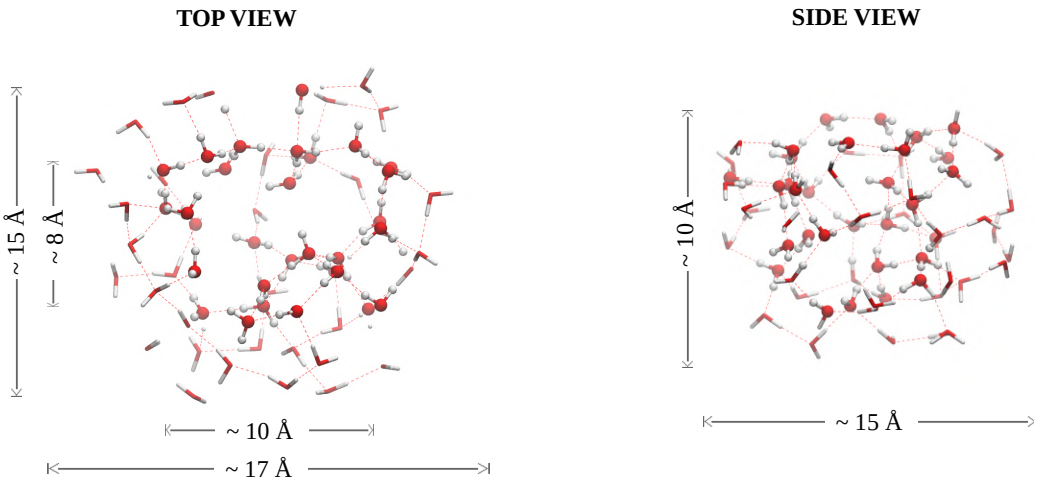
Periodic slab models

The surface of the periodic slabs is constituted by molecules at the top layer, where the target species can adsorb. Assuming the adsorption site to have an area of approximately 3×3 Å, each surface might accommodate $\sim$100 potentially unique binding sites.

Providing a proper estimation of the percentage of dangling-Hs of the periodic systems is more complex with respect to the cluster models. The estimation took into account only the surface layer. Edges molecules have been excluded in order to avoid border-effects. Considering the aforementioned unique surface sites per system, 25% of total binding sites are estimated to harbour dangling-Hs, in average over the 5 models. It is worth mentioning that the number includes dangling-Hs that are located into concavities or porous, as those regions constitute part of the periodic surface.



a)



b)

Figure 3.5: Computational steps to generate the nono-porous model. a) Left: top view and altitute map of a periodic surface of initial density 0.8 g cm^{-3}. One of the nano-porous suitable sites is highlighted. Right: the spherical site as extracted from the model. b) Top (left) and side (right) view of the nano-porous after geometry optimization. Molecules that are kept constrained at each step of the work are represented as sticks. Inner and outer sphere dimensions are included in the figure.

3.2 Geometry and Energy Benchmark

The performance of DFT relies heavily on the exchange-correlation density functionals used. Some functionals are designed to be suitable for specific systems or to reproduce geometries or energy barriers. Besides, they vary greatly in terms of computational cost. The usual practice is testing (*benchmarking*) a set of DFT functionals and basis on minimal surface models in order to find the most appropriate method to be applied to larger systems.

3.2.1 Binding energies

In order to obtain the best possible equilibrium geometry at a reasonable computational cost, geometry benchmark is performed on the $W_{2-3}-X$ systems, with X being the target molecule, W_n the water cluster of different size and n is the number or water molecules. The benchmark has been conducted for 13 selected molecules. A DF-CCSD(T)-F12/cc-pVDZ-F12 geometry was used as a reference; 24 GGA exchange-correlation density functionals are probed, including GGA, Hybrid-GGA, Double Hybrid-GGA, Meta-GGA, Meat-Hybrid-GGA and Long Range Corrected; paired with a def2-TZVP basis. The energy benchmark is also conducted, using the W_4-X system to compare BSSE corrected DFT BE values to a CCSD(T)/CBS reference

energy. Table C.1 reports geometry benchmark results. Generally, the Meta-Hybrid-GGA methods have a very good performance. The most dependable functionals are B3LYP[150,155] for structures bonded by dispersion interaction ('Group D') and PWB6K[213] for structures bonded by electrostatic interaction ('Group H'), as both show an average RMSD value below 0.1 Å with respect to the reference geometry. Parametrized HF-3c/MINIX[214] and PBEh-3c/def2-mSVP[215] levels of theory are also probed. The results are reported in the third column of Table C.1 and show an average RMSD that is below 0.2 Å for both groups, which is in line with the RMSD values of Hybrid and Meta-Hybrid functionals. This makes HF-3c a cost-effective alternative to the computationally more expensive DFT methods.

Regarding the energy benchmark, Table C.1, for both groups the best DFT functional is the ω-PBE[158,216] with BSSE and D3BJ dispersion corrections, coupled with def2-TZVP basis set. The average mean absolute error (MAE) is 37 and 160 K for Group D and H respectively. Full benchmark results can be found in GitHub[217].

3.2.2 Reaction and Transition State Energies

The Strecker synthesis of glycine is comprised of 4 stages. It starts by nucleophilic addition of NH_3 to H_2CO to give aminomethanol (NH_2CH_2OH) (S1). The intermediate is then converted to methanimine ($NHCH_2$) by water

elimination (S2):

(S1) $$H_2CO + NH_3 \rightarrow NH_2CH_2OH \tag{3.1}$$

(S2) $$NH_2CH_2OH \rightarrow NHCH_2 + H_2O \tag{3.2}$$

The third stage (S3) is the nucleophilic addition of HNC to $NHCH_2$ to give aminoacetonitrile (NH_2CH_2CN)

(S3) $$NHCH_2 + HNC \rightarrow NH_2CH_2CN \tag{3.3}$$

The final stage is the hydrolysis of NH_2CH_2CN. It involves the nucleophilic attack (S4a) of a water molecules on the C atom of NH_2CH_2CN, to yield a hydroxy imine intermediate, which tautomerizes (S4b) to the more stable glycine amide form ($NH_2CH_2C(O)NH_2$), followed by a second nucleophilic addition of water (S4c) to give glycine.

(S4a) $$NH_2CH_2CN + H_2O \rightarrow NH_2CH_2COHNH \tag{3.4}$$

(S4b) $$NH_2CH_2COHNH \rightarrow NH_2CH_2C(O)NH_2 \tag{3.5}$$

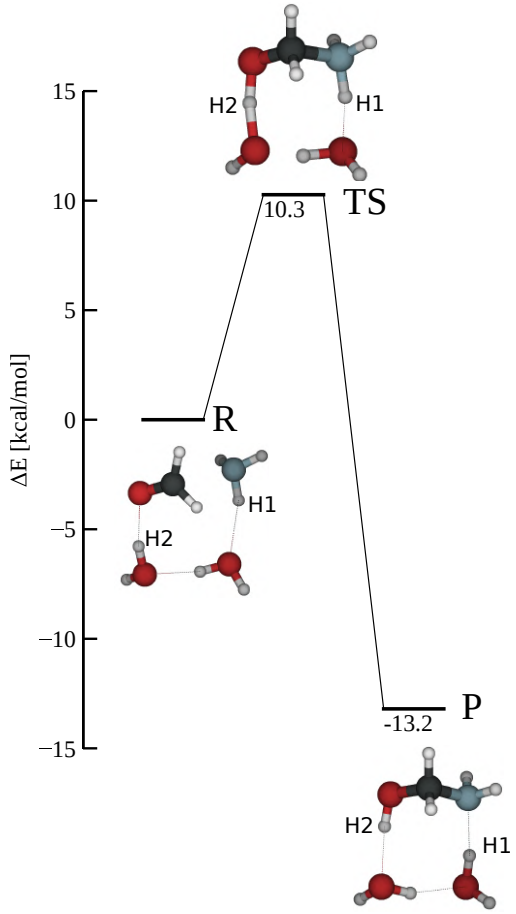


Figure 3.6: Energy diagram for S1, reference system (NH_3 + H_2CO + W_2), using DF-CCSD(T)-F12/cc-pVDZ-F12 geometries. Energies have been computed at CCSD(T)/CBS level of theory. Minimum energy structures - reactants R, transition state (TS) and product (P) - are reported as well. The color scheme for the atoms is red for O, black for C, blue for N and white for H.

(S4c) $$NH_2CH_2C(O)NH_2 + H_2O \rightarrow NH_2CH_2COOH + NH_3 \quad (3.6)$$

Stage 1

Stage 1 (S1) is the addition of NH_3 to H_2CO to give NH_2CH_2OH. The mechanism involves the nucleophilic attack of the N atom towards the C-carbonyl atom and the transfer of a proton from NH_3 to the O-carbonyl. NH_3 +

$H_2CO + W_2$, is considered as reference system. The water molecules are acting as proton transfer intermediaries, in an arrangement commonly named as *proton relay*. The reference system geometry is DF-CCSD(T)-F12/cc-pVDZ-F12, which has been shown to provide excellent geometries[142], while the reaction energies have been computed at CCSD(T)/CBS extrapolated basis set.

The energy differences are reported in Figure 3.6. The reaction is composed of a single exothermic step. The main bond distances involved in the bond breaking/forming processes have been reported in Table 3.1.

Table 3.1: Structural parameters involved in S1 main bond breaking/forming processes for the reference system. Bond distances (B) are in angstrom (Å), bond angles (A) are in degrees (^{o}).

Strucure	B(C-N)	B(N-H1)	B(O-H2)	A(O-C-N)
R	2.79	1.01	1.81	112.3
TS	1.55	1.13	1.09	108.7
P	1.48	1.84	0.98	110.2

Around 53 DFT functionals are taken into account for the geometry benchmark belonging to different classes, and two different basis sets: def2/SVP and def2/TZVP. This is due to the fact that a double ζ basis is used to study the reaction on many binding sites on the larger surfaces, hence the need to assess the consistency of a specific DFT functional with the two tiers of method and basis. Dispersion effects are treated using D3BJ and D4 correction factors. Figure C.1 shows the RMSD between the reference and the DFT level methods. Geometries of reactant (R), transition state (TS)

and product (P) are taken into account. The plot displays only the methods that successfully managed to optimize all the minima. The best method is BHANDHLYP-D4/def2-SVP with a average RMSD error of 0.10 Å. It also provided the best performance using the larger basis set (in average 0.05 Å) tied with MPWB1K-D4/def2-TZVP, ω-B97-D4/def2-TZVP and PW6B95-D4/def2-TZVP. In terms of geometry, R structure has been the most challenging to properly optimize for the majority of the functionals. This is very likely due to the fact that R equilibrium geometry is dominated by non-covalent interactions which are more sensitive to the method used to describe them. Regarding the dispersion treatment, the error is comparable employing D4 or D3BJ correction, with slightly better results with D4, for most of the methods.

Regarding the energy benchmark, 258 DFT functionals are considered. Since a high level of accuracy is required for the energy computation, only def2-TZVP basis set has been tested. Figure C.4 shows the mean absolute error (MAE) between the methods, considering both the energy barriers (ΔE_{TS}) and the reaction energies (ΔE), and the reference geometries system. The plots report only the best 20 Meta-Hybrid-GGA and Hybrid-GGA methods that provided the best performance, ordered according to the lower ΔE_{TS} energy. On average, the reaction energies displayed higher MAE error compared to the TS energies. In order to check the consistency of a specific tier of geometry and energy DFT methods in reproducing the reference energy, ΔE_{TS} and ΔE are computed on 7 of the best DFT geometries using

the best 4 DFT energies selected from the energy benchmark. The results are reported in Figure C.7. The most consistent are the geometries obtained with BHANDHLYP-D4 and MPWB1K-D4, as expected from the geometry benchmark results. The choice ultimately fell on BHANDHLYP because MPWB1K belong to the Meta Hybrid-GGA class and, therefore, requires larger computational time. Regarding the energy, the performance of the 4 energy methods is very similar, the best ones being BMK/def2-TZVP and ω-B97M/def2-TZVP with MAE error much below 1 kcal/mol on both double and triple-ζ geometries.

To summarize, the benchmark allows to identify two suitable tiers of methods for stage S1: ω-B97M/def2-TZVP // BHANDHLYP-D4/def2-SVP and BMK/def2-TZVP // BHANDHLYP-D3BJ/def2-SVP. The former is the best one and is to be used with the ORCA software, while the latter is meant to be used with GAUSSIAN software, due to the unavailability of the first tier.

Stage 2

Stage 2 (S2) is the dehydration of the aminomethanol to give methanimine. The water molecule is formed thanks to the proton transferring from the amino to the alcoholic group of the aminoalcohol. $NH_2CH_2OH + 2H_2O$ ($S2-W_2$), is taken as model system, where the reaction is mediated by two water molecules, establishing a *proton relay*.

The energy barrier for the process is shown in Figure 3.7. The reaction is concerted and endothermic. The structural parameters involved in the bond breaking/forming processes have been reported in table 3.2.

Table 3.2: Structural parameters involved in S2 main bond breaking/forming processes for the reference system. Bond distances (B) are in angstrom (Å), bond angles (A) are in degrees (^{o}).

System $S2-W_2$	B(O-C)	B(N-H1)	B(O-H2)	A(O-C-N)
R	1.46	1.01	1.81	116.4
TS	2.20	1.07	1.06	113.8
P	2.93	1.84	0.97	123.7

The benchmark results are illustrated in Figure C.2 and Figure C.5. Regarding the geometry, BHANDHLYP-D4/def2-SVP and MPWB1K-D4/def2-SVP are the best level of theory with a average RMSD error of 0.21 Å. The product structure has been most difficult to reproduce for the majority of the methods, as expected, since it is a two-fragments interacting complex. Regarding the energy, PWPB95/def2-TZVP (with and without dispersion cor-

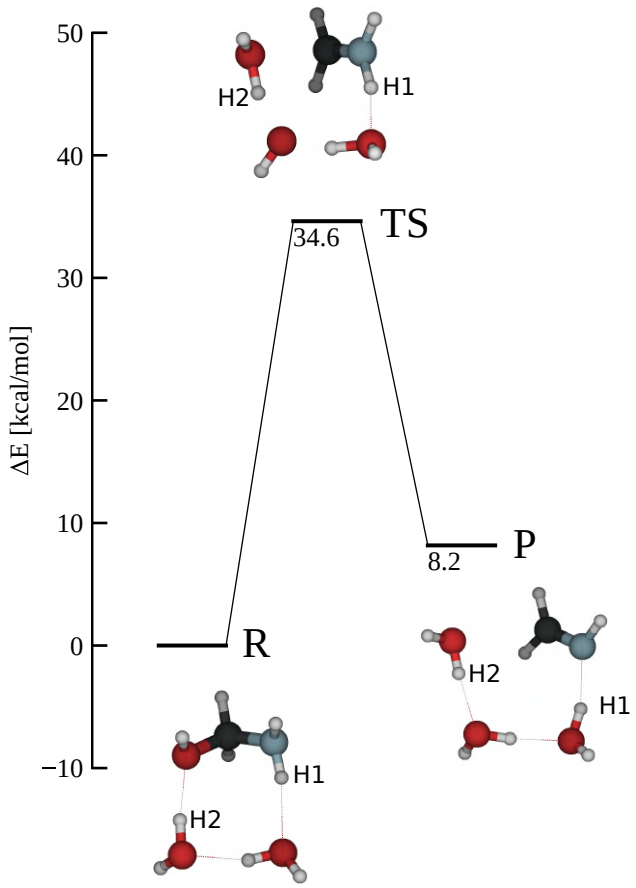


Figure 3.7: Energy diagram for S2, reference system (NH_2CH_2OH + W_2), using DF-CCSD(T)-F12/cc-pVDZ-F12 geometries. Energies have been computed at CCSD(T)/CBS level of theory. Minimum energy structures - reactants R, transition state (TS) and product (P) - are reported as well.

rection) and ω-B97M-D3BJ/def2-TZVP provided the best energies, with a MAE below 1 kcal/mol. The performance is better for the TS energies compared to the reaction energies, probably due to the challenging product configuration. In summary, the tier ω-B97M-D3BJ/def2-TZVP // BHANDHLYP-D4/def2-SVP, used to study S1, is confirmed to be valid also for S2.

Stage 3

The third stage (S3) is the aminoacetonitrile formation. The model system is $NHCH_2$ + HNC + $2\,H_2O$ ($S3-W_2$), meaning that the nucleophylic addition is assisted by two water molecules. S3 is concerted and exothermic. The energy diagram is reported in Figure 3.8. The structural parameters involved in the bond breaking/forming processes can be found in Table 3.3.

Table 3.3: Structural parameters involved in S3 main bond breaking/forming processes for the reference system. Bond distances (B) are in angstrom (Å), bond angles (A) are in degrees (^{o}).

System $S3-W_2$	B(C-C)	B(N-H1)	B(N-H2)	$A(N_{HNC}-C-C)$
R	3.34	1.03	1.82	79.1
TS	2.76	1.46	1.31	93.6
P	1.49	2.01	1.02	177.0

The benchmark results are illustrated in Appendix Figure C.3 and Figure C.6. Regarding the geometry, many DFT functional failed to reproduce TS structure, due to its highly zwitterionic character (*vide infra*), including the tier of method used for S1-2. The best method is PW6B95-D4/def2-SVP

with a average RMSD error of 0.10 Å. Regarding the energy, BMK/def2-TZVP provides nearly exact TS energies, along with M05/def2-TZVP, that is the best overall method, with MAE error less than 0.5 kcal/mol for both reaction energy and energy barrier.

Stage 4

Due to the size of the system, it was not possible to carry out a benchmark for this stage, therefore the benchmark results of the previous section were taken into account for geometries and energies.

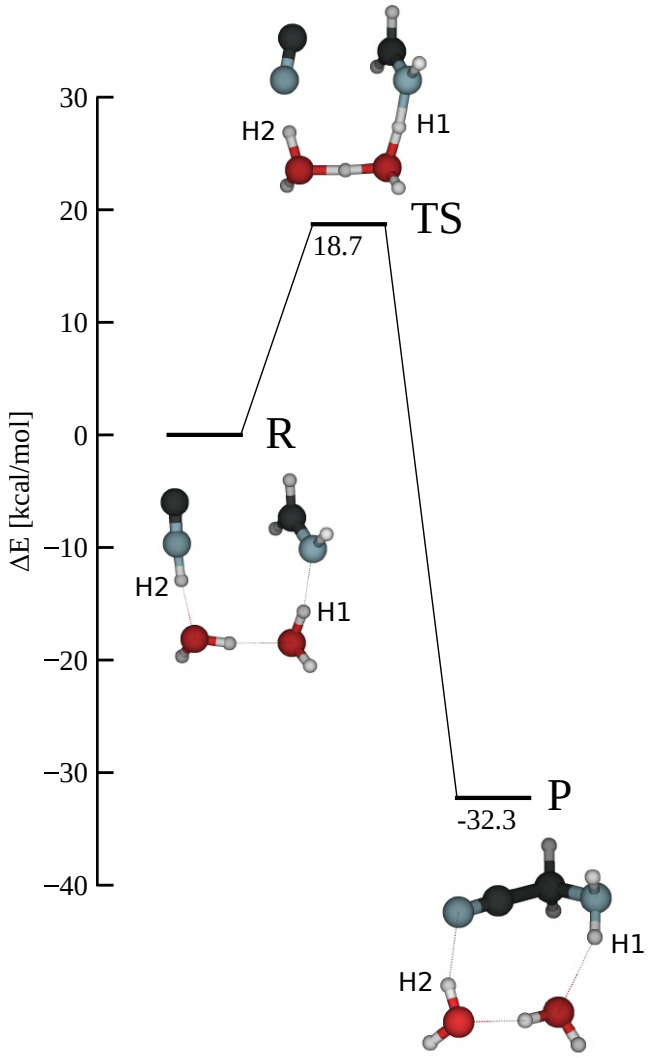


Figure 3.8: Energy diagram for S3, reference system ($NHCH_2$ + HNC + W_2), using DF-CCSD(T)-F12/cc-pVDZ-F12 geometries. Energies have been computed at CCSD(T)/CBS level of theory. Minimum energy structures - reactants R, transition state (TS) and product (P) - are reported as well.

3.3 Binding energy distributions of small species

The binding energy (BE) is a crucial parameter when modeling interstellar gas-grain chemistry. According to the *multi-binding* approach (see Section 1.1.3), using a BE distribution of values reflects a more realistic desorption behaviour for molecules adsorbed on ASW ice. Notwithstanding, computing a large set of BEs requires a significant amount of computational resources and data management. In the following section, a new computational protocol to compute BE distributions in a highly automated fashion is presented (Section 3.3.1) and validated (Section 3.3.2). In Section 3.3.3 the characteristics of the novel BE evaluation platform are presented. The first BE results obtained are discussed (Section 3.3.4), along with possible astrophysical implications of the new framework. Finally, the effects of the size, morphology and topology of the ice model on the BE is addressed (Sections 3.3.5, 3.3.6).

It is reminded that interstellar BEs are typically expressed in units of Kelvin (K), because temperature is a convenient unit of measurement in astrophysics, as it allows to easily compare the BE with the thermal energy available in the system. The convention is followed in this thesis, as well. However, when the BE is used in the context of reactivity processes, such as to define the strength of a certain binding motif of a species, in relation to its ability to engage in a chemical reaction, the BE is expressed in units of kcal

mol^{-1}.

3.3.1 Computational protocol

In order to easily produce BE distributions of species adsorbed on interstellar ices, in the context of this thesis a novel Binding Energy Evaluation Platform (BEEP) was build. BEEP implements the capability to compute binding energies on surface cluster models in the form of a computational protocol. It also contains a database that allows one to query the results produced by the protocol. BEEP consists of three highly automated steps, showed in Figure 3.9: target molecule sampling procedure (1), geometry optimization of the binding sites (2), and binding energies computation (3), by means of DFT methods.

The computational platform is implemented within the QCArchive[183] framework. QCArchive data structure and terminology has been introduced in Section 2.9.

Sampling procedure

The first operation is sampling the ice model with the species of interest. Cluster surface models of different sizes and composition (e.g., different ice mixtures) can be easily added to the platform environment. In order to perform the sampling procedure (Figure 3.9, blue panel labelled “sampling") within the QCArchive environment, both the ASW clusters and the target

molecules have to be stored in collection objects (ASW Dataset and Species Dataset). The sampling procedure is carried out at BLYP/def2-SVP level of theory, and consists of extracting one ASW structure at a time from the ASW Dataset and sample it with the target molecule X. The sampling algorithm places the center of mass of both species on the origin of the system coordinates, and displaces the species X around the surface randomly, within a range of distances which maximizes the chance of finding a binding site on the surface. Starting with ice cluster ASW_a, several groups of 10 $\mathrm{ASW}_a - \mathrm{X}$ binding site candidates (BSC) are generated. These are optimized (*opt1*) and filtered according to geometrical criteria, such that only the structures of RMSD $\geq$ 0.40 Å with respect to previously found BSC are stored, until 25 BSC are reached or no more new BSC are found. This procedure is repeated on a second cluster ASW_b until reaching a total of at least 225 $\mathrm{ASW} - \mathrm{X}$ equilibrium structures, distributed among 12-15 ASW clusters.

Geometry Optimization

In this step (Figure 3.9, yellow panel labelled “geometry optimization"), the BSCs previously obtained, are further optimized at a more accurate level of theory, such as a Hybrid or Meta-Hybrid GGA functional with a triple-ζ basis set, chosen according to geometry benchmark results (see Section 3.2). In the next step, BEs are calculated carrying out single point energy computations on the BSCs.

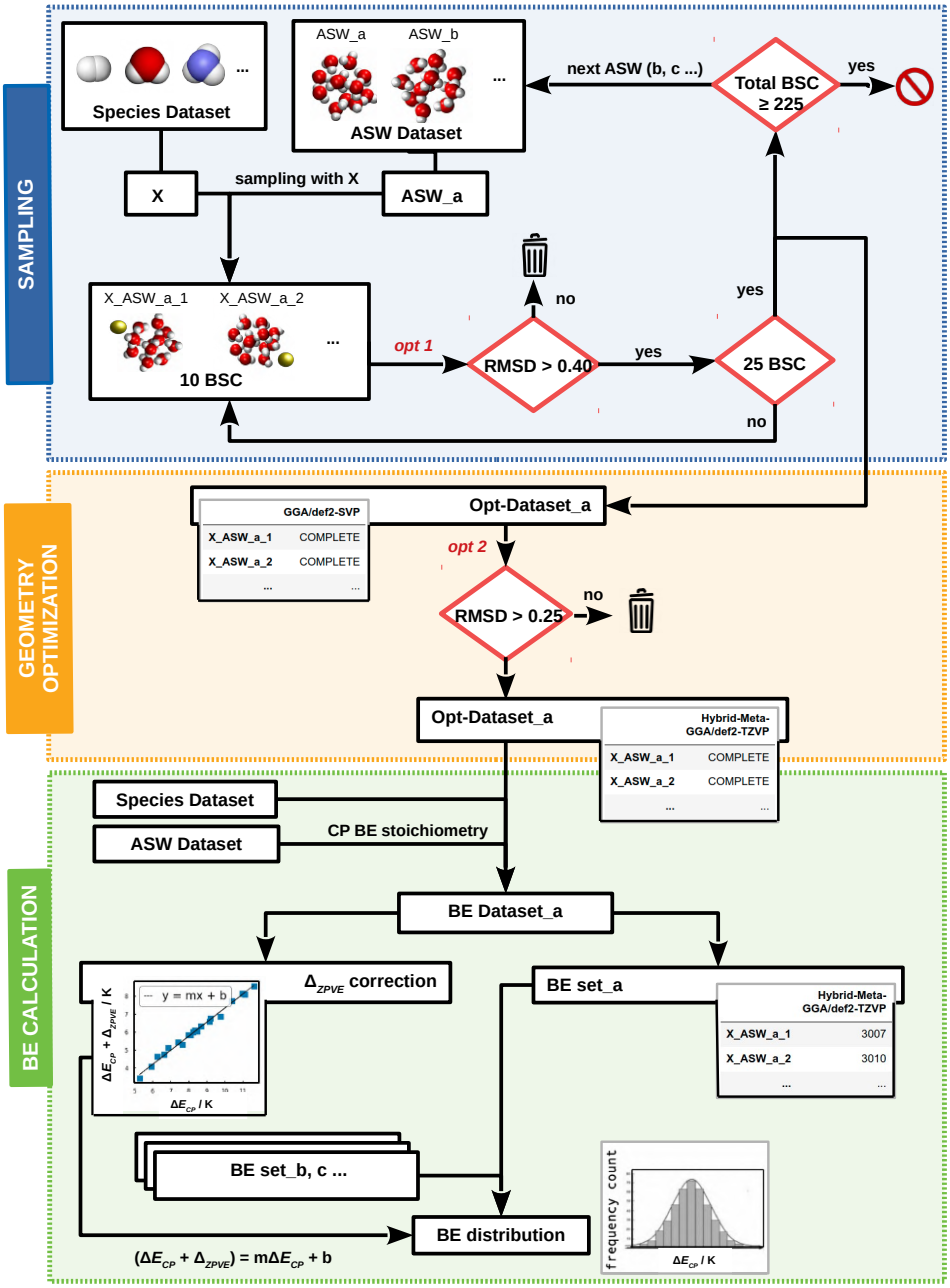


Figure 3.9: Three-step computational procedure used for building a binding energy distribution. BSC stands for binding site candidate; *opt 1* stands for optimization at gradient generalized approximation (GGA) exchange-correlation DFT functional and *opt 2* for optimization at a higher level of theory that further refines the geometry. The color scheme for the atoms is red for O, white for H, blue for N and yellow for the generic target atom X.

Binding energy calculation

The final part of the procedure (Figure 3.9, green panel labelled “BE calculation") is the computation of BE values and the assembly of a ZPVE corrected BE distribution. To do so, first, the optimized structures are filtered with geometry criteria (RMSD $\geq$ 0.40 Å) to make sure that all binding sites on the ASW cluster are unique. The resulting equilibrium structures are included into a BE Dataset collection, together with the optimized target molecule and water cluster to create the BE stoichiometry including the counterpoise correction for the BSSE error (Section 2.5.1). Once a BE Dataset for $\mathrm{ASW}_a-\mathrm{X}$ is generated, it contains all the fragments necessary to compute the BSSE corrected BE values on ASW_a (BE set_a). Analogously, a set of BEs for each of the sampled clusters is collected. Assuming the clusters share common morphological characteristics, as they originate from a single AIMD trajectory and are annealed in the same way, the BEs collected are considered as a single BE distribution of the target molecule on the ice mantle model. The values are finally corrected by adding Δ_{ZPVE} to the BE. Due to computational cost, the Hessian is computed only for the elements of a single BE Dataset (e.g., $\mathrm{ASW}_a-\mathrm{X}$), and a linear model is used to correlate ΔE_{CP} and $\Delta E_{CP}+\Delta_{ZPVE}$. Finally, the correction factors are applied to all the computed BEs to obtain a ZPVE corrected BE distribution. The source code of the BEEP protocol and scripts to generate the data can be found in GitHub[217].

3.3.2 Validation of the protocol

The procedure has been first showcased on a simple but astrophysically relevant species, hydrogen fluoride (HF)[36]. HF is reliable tracer of the molecular gas in galaxies[218], thanks to the direct connection between HF absorption depth and H_2 column density. The observational data, however, suggest that in cold clouds (< 20 K), HF may condense onto dust grains. Although freeze-out effects might reduce the effectiveness of HF as a diagnostic tool, the density and temperature conditions needed for its adsorption and desorption, have been studied in astrophysical context only qualitatively[218,219], hence the importance to contrain such parameters.

As a surface model, the ASW set of cluster of 22-water molecules are employed. The computational protocol provided the BE distribution reported in Figure 3.10, left. A Gaussian fit of the histogram data allows to locate the center of the distribution at 5313 ± 74 K (corresponding to 10.6 ± 0.1 kcal mol^{-1}) and to calculate a dispersion is of 921 ± 155 K (1.8 ± 0.2 kcal mol^{-1}). In total 255 unique structures have been found for the system, corresponding to 3 different binding modes, an example of any is reported in Figure 3.10, right. The binding mode that exhibits the highest energy values, corresponds to HF bound to the ASW surface through two different hydrogen bonds (Figure 3.10a). The ZPVE has been calculated to reduce the BE value by about 1000 K for each binding mode.

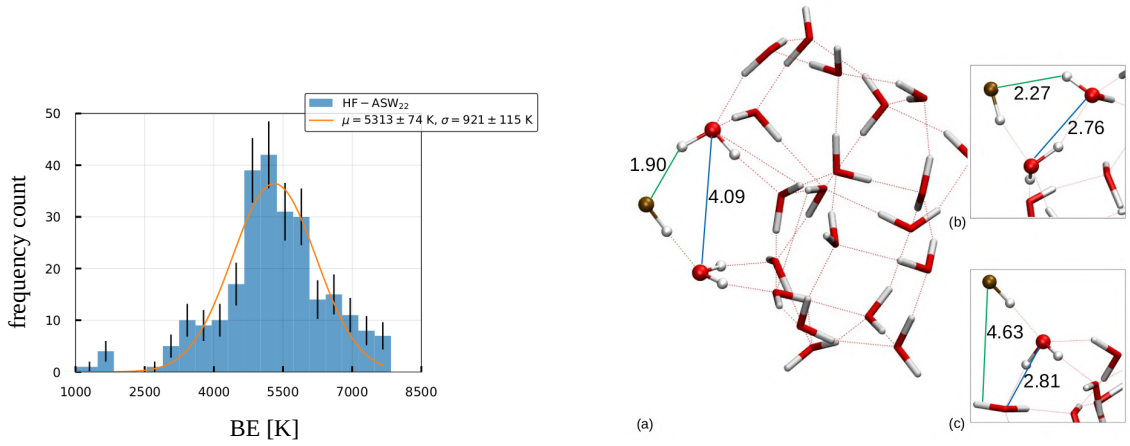


Figure 3.10: Left: histogram of the binding energy distribution obtained for the HF – ASW_{22} system. Mean (μ) and standard deviation (σ) of the Gaussian fit are reported in the legend. Right: equilibrium geometries calculated at revPBE0/def2-TZVP level. The structures highest in BE for each the three binding modes identified (a, b, c) is reported. The atoms within 3 Årange from F atom are represented as ball and sticks. The color scheme for the atoms is yellow for F, red for O, and white for H.

The existence of such distinct adsorption modes, even for a diatomic species, suggests that the surface properties of the ASW in proximity of the binding site exert significant effects. In order to specifically gauge the influence of the water environment on the BE, the evolution of the binding energy with respect to the size of the cluster is followed for a selected structure of each binding mode. In a practical way, it is done by removing one water molecule at a time, starting from the optimized structure, and keeping the remaining molecules frozen. The process is repeated until all but the two nearest water molecules to the HF binding site are removed. The resulting ΔBE values, computed for each frozen water cluster, are plotted in Figure 3.11. The 0 K baseline corresponds to the original energy value for one of the HF –ASW_{22}

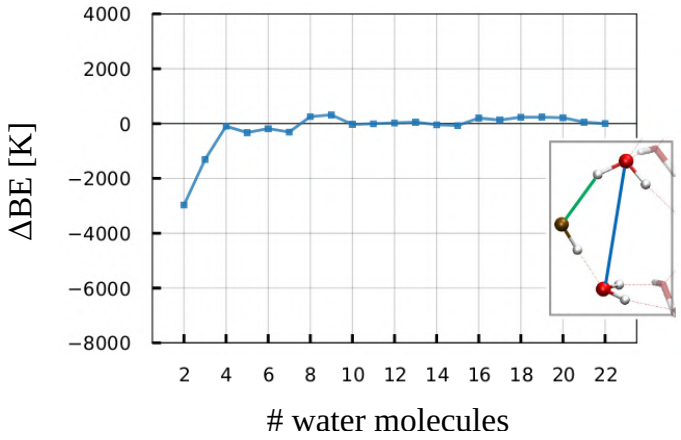


Figure 3.11: Changes in the binding energy (Δ BE) with respect to the numbers of water molecules in the frozen ASW_{22} cluster in the supermolecular configuration for one of the binding site. The 0 K baselines correspond to the original BE value.

structure, such that the plotted energies represent the deviation from that BE at each cluster size. The plot shows a rather fast convergence behavior to the final value; at a cluster size of 4 water molecules, ΔBE is already converged, indicating that only the local water environment has a significant effect on the BE. In conclusion, the absence of major variations in the BE passed the close proximity of the binding site (i.e. beyond the 4 water molecules nearest to the adsorbate), ensure that cluster models of 22-water molecules guarantee that the relevant part of the interaction is taken into account.

3.3.3 BEEP: Binding Energy Evaluation Platform

BEEP is a Binding Energy Evaluation Platform meant to offer a straightforward, highly automated and easy-to-use interface for the computation and processing of full BE distributions of molecules. The platform is imple-

mented within the QCArchive framework[183], which allows to transform the database in a fully open-source endeavour, from the data generation to the final user-query of the BE data. Due to the nature of QCArchive Databases, BEEP is extendable to an increasingly large number of molecules. Moreover, different cluster surface models of different sizes and composition (e.g., different ice mixtures) can be easily added to the platform environment and used to produce new BE distribution data. To make the access to the database a user-friendly experience, a Python module is also included, querying the data without having to know QCArchive syntax. The core of the Python module is the *BindingParadise* class that is initialized with the user's credentials and allows to set a species and obtain all the related BE data. BEEP has been created aimed by the idea of a collaborative endeavour in which researchers of different groups use the protocol to generate new BE data and store it in the open database, expanding it both in terms of ice models and interstellar species. A database of reproducible and accurate BEs is also a fundamental starting point to chemical reactivity studies and diffusion of molecules on the surface of interstellar ices, as having a potential energy map of neighbouring binding sites is paramount in finding diffusive transition states and computing diffusion energy barriers.

3.3.4 Binding energy results

To present the utility of BEEP, BE distributions of 21 astrophysically-relevant molecules have been computed.

The species are divided into two groups according to the nature of the interaction with the ice surface. 'Group D' accounts for interactions dominated by dispersion, while molecules in 'Group H' predominantly bind through hydrogen bonds. The BE distribution results are reported in Figure 3.12, 3.13. The equilibrium geometry is of HF-3c/MINIX quality, as it is proved to be a cost-effective alternative to the more expensive DFT methods. For CO species, the geometry is M05/def2-TZVP[213], as HF-3c failed to properly describe the binding sites. The ZPVE corrections have been computed at the HF-3c/MINIX level of theory for Group H, while for most of the molecules in Group D it was not possible to apply the linear model used to derive the correction factors, due to poor correlation. This could be attributed to the inadequacy of the harmonic approximation to correctly describe the potential energy well. Notwithstanding, the correction value for Group D molecule is small enough to fall within the accuracy of the method. The BE values were computed using the best performing DFT functional from the energy benchmark for each molecule (see previous section). If no benchmark value is present, the best performing functional for each group in employed.

A single BE value, representative of the entire distribution, is also pro-

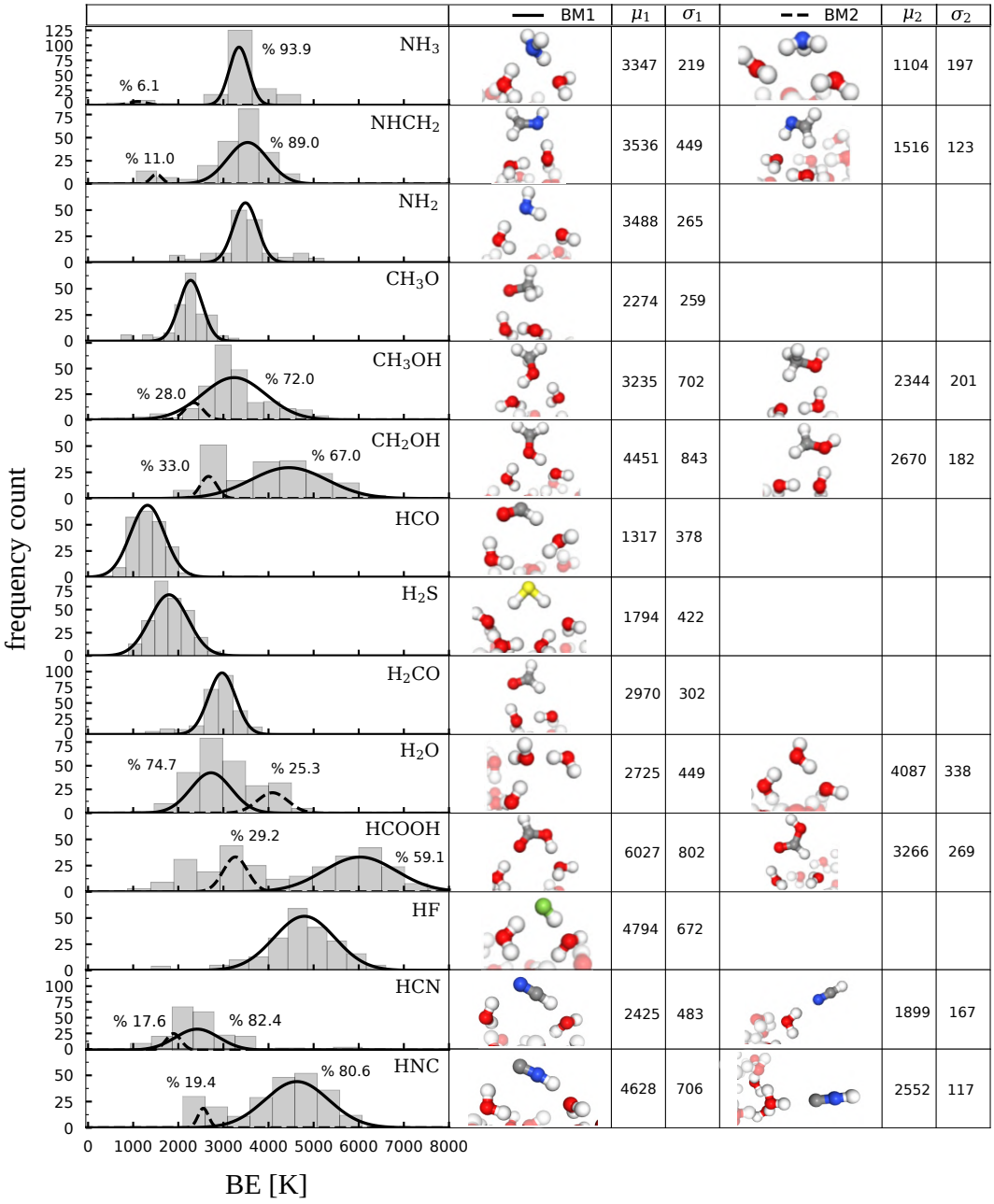


Figure 3.12: Binding energy distributions for Group H, ASW−X systems, using HF-3c/MINIX geometries and including ZPVE correction. According to the benchmark results, the energy has been computed at ω−PBE/def2-TZVP level of theory for all species except HNC (B97-2/def2-TZVP), H_2CO (CAM-B3LYP/def2-TZVP), CH_3OH (TPSSH/def2-TZVP), HF and HCN (MPWB1K/def2-TZVP). Mean (μ) and standard deviation (σ) of the Gaussian fit are reported for the major binding mode (BM1, solid line) and the minor binding mode (BM2, dashed line). The color scheme for the atoms is red for O, grey for C, white for H, blue for N, yellow for S and green for F.

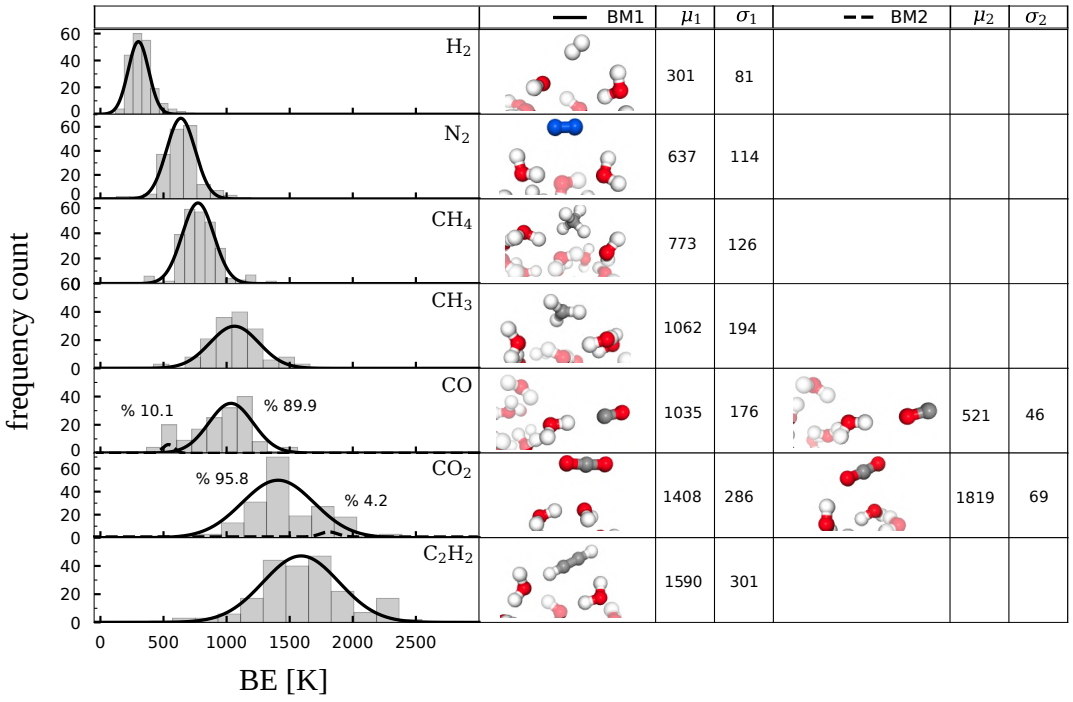


Figure 3.13: Binding energy distributions for Group D, ASW−X systems, using HF-3c/MINIX geometries except for CO (M05/def2-TZVP). ZPVE correction has been included only for C_2H_2, see Section 3.3.4. The energy has been computed at ω-PBE/def2-TZVP level of theory for all species except CH_4 (TPSSH/def2-TZVP). D3BJ dispersion correction has been applied to all DFT energies. See Figure 3.12 caption for further details.

vided to accommodate the usage of BEEP calculations in standard chemical models. For this purpose, mean BE (μ) and standard deviation (σ) are obtained by fitting a Gaussian function to the distribution using a bootstrap method (Section D.1). Binding mode analyses has been carried out for all species, in order to identify different binding motifs which are labelled in the figure, along with corresponding percentage and μ and σ values. An example of binding mode analysis procedure can be found in the following section for NH_3 and also in a Jupyter Notebook in the BEEP GitHub repository.

Hydrogen bonded structures

Figure 3.12 shows BE distributions of molecules in Group H. These molecules are mostly bound through electrostatic interactions in the form of hydrogen bonds (HB) and therefore present a strong interaction with the ASW surface. This is reflected in the BE values which are in the range of 1000 to 8000 K. It is worth noting that several species exhibit two distinct distributions.

For NH_3 and $NHCH_2$, there is a main binding mode ($\mu_1 \sim 3400$ K) where the molecule bridge two water molecules via a double HB. In the minor binding mode ($\mu_2 \sim 1400$ K), the surface water molecules act solely as HB acceptors, resulting in a lower mean BE. For CH_3OH and its radical species (CH_2OH), the main binding mode is the surface interaction via the OH moiety (μ_1 = 3235 and 4451 K, respectively); while in the minor binding mode the methyl end of the molecule also participates in the interaction. In the distribution of the CH_3O radical, a single binding mode was found, corresponding to the less energetically favorable interaction where both the oxygen and the methyl take part. This is consistent with the inability of this radical to form a donor-type HB. Due to its lack of symmetry, the formic acid (HCOOH) presents a rather complex BE distribution with two different components, spanning a range of almost 7000 K. The minor mode present dangling OH-bonds as in the case of the methanol species. Regarding the water molecule, a closer inspection of the binding modes shows a varied scenario where the

molecule establishes a single (μ_1 = 2725 K) or double (μ_2 = 4087 K) HB to the surface. Even though the former occurs more often during the sampling procedure, the majority of the water molecules that compose the ASW surface form two HBs, therefore water surface evaporation would mostly fall within the higher BE regime. The halogen (HF) and pseudo-halogen (HNC, HCN) molecules have a high standard deviation ($\sigma \sim 600$ K) which reflects a high capacity of insertion into the ASW environment. This is especially seen in the HF case, in which the molecule is easily inserted into the hydrogen bond network, forming strong HB with the water surface, as it was shown in the previous validation work[36]. Both HCN and HNC species exhibit two binding modes. In the main one (μ_1 = 2425 and 4628 K, respectively), the molecules establish a double HB interaction with the surface. Finally, it is worth noting that the ZPVE correction significantly reduces the BEs, in some cases up to 25% of the non-corrected value.

Dispersion bonded structures

Figure 3.13 shows the BE distributions of Group D. In order to identify the molecules that belong to this group, the BE distributions obtained with and without including D3BJ dispersion correction to the energy computation are compared. For molecules in Group D, the dispersion interaction is fundamental in order to achieve an attractive interaction with the water surface (see Appendix, D.3). They are mainly homonuclear or highly symmetric

molecules. The mean BE values range between 300 and 1800 K and are significantly lower than in the Group H molecules. Furthermore, the standard deviation is also less than in Group H molecules, which is consistent with a smaller capacity of the molecule to deform the binding site environment. Most molecules therefore present a single binding motif. An outlier is CO, since its BE distribution reveals two distinct binding modes: the majoritarian one comprises 89.9 % of the structures and involves the C-extremity of the molecule (μ_1 = 1035 K). The other molecule that presents more than one binding mode is CO_2. In the highest BE motif, CO_2 interacts with the surface through both C- and one of the O-atoms of the molecule (μ_2 = 1819 K).

3.3.4.1 Participants of the Strecker Synthesis

A closer analysis of the distinct binding motives that characterize the participants of the Strecker synthesis of glycine might allow to predict potential reactive situations. The procedure to separate the binding modes is carefully illustrated for NH_3, in the following. Similar methodology can be applied to the other species.

NH_3

NH_3 is one of the reactants of the first step of the synthesis. As showed in Figure 3.12, the BE distribution revealed a majoritarian peak and a low BE tail. A certain binding mode is given by the HB interactions that the system

features. Two water molecules (W) are taken to be hydrogen bonded[220,221] if their (H − acceptor O) distance is less than 2.4 Å; the same cut-off is employed for any kind of HB.

For the system NH_3 – ASW, two types (T) of HB are observed:

- T1: $N-H_W$ with H_2O as donor and NH_3 as acceptor
- T2: $H_{amino}-O_W$ with NH_3 as donor and H_2O as acceptor

where the atom sub-index indicate which fragment it belongs to: 'amino' stands for amino group (here used in the general sense of H-atoms bound to a N-atom) and W for one of the ice water molecules. The *modus operandi* to identify distinct binding modes is based on careful analysis of characteristic structural parameters (e.g. bond distances) of the adsorption site. First, the distance between N-atom and the water molecule nearest to the binding site ($B(N-O_{W1})$) is analyzed. The histogram of the distances is reported in Figure 3.14a, left, and shows a clear separation in two groups. The first group ($B(N-O_{W1}) < 3$Å) is characterized by structures where N-atom is oriented toward one water molecule, engaged in a T1 interaction. It corresponds to the binding mode labelled as BM1, that accounts for 93.9% of the binding sites. The second, BM2, ($B(N-O_{W1}) > 3$Å) is composed by rarer structures where NH_3 is acting solely as HB donor, in agreement to a T2 interaction. Among BM1 structures, it is possible to discern two arrangements. Analysis of the distance between the two O-atoms nearest to

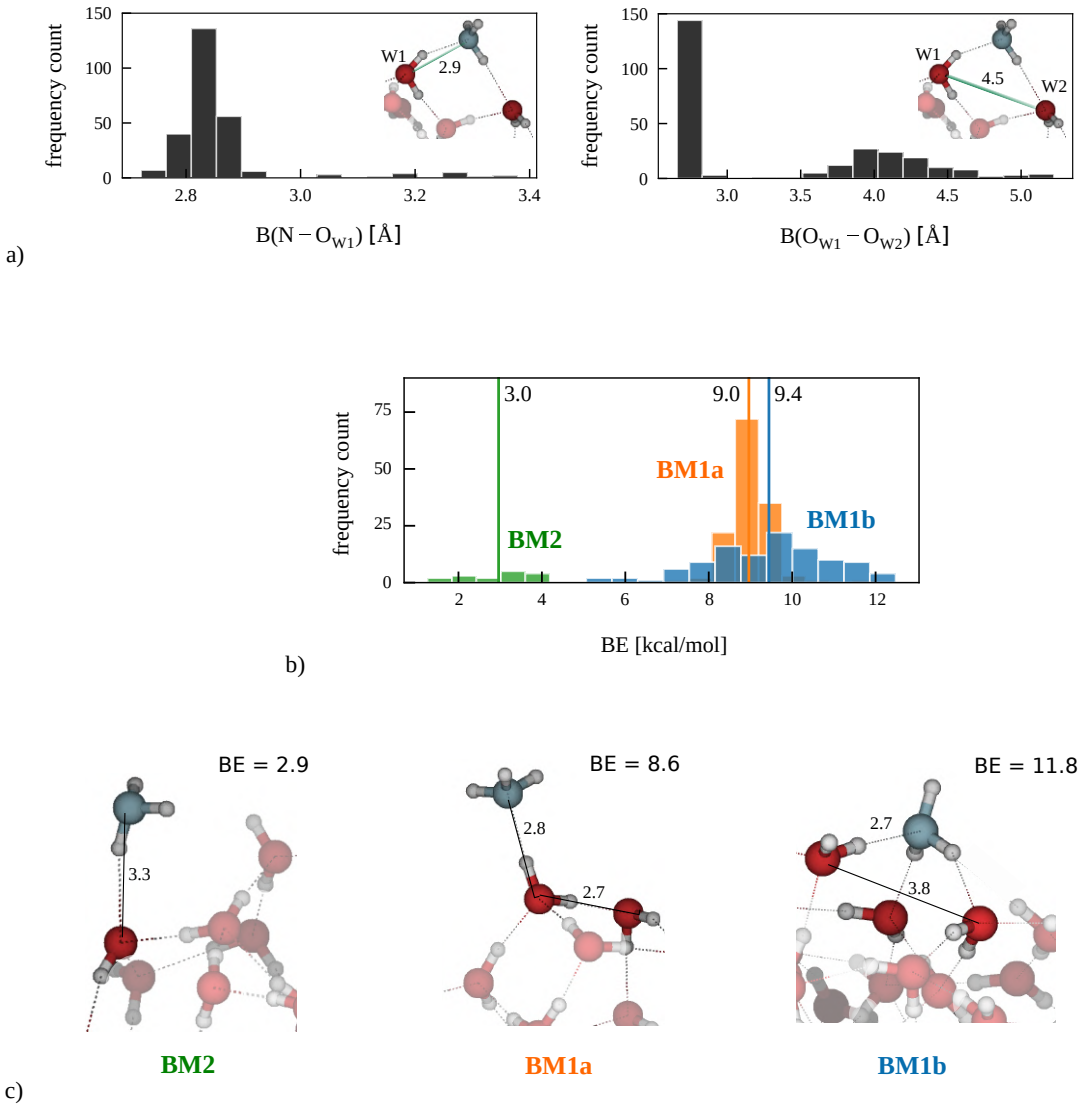


Figure 3.14: a) Histogram of two distances relevant for the binding modes attribution of the system NH_3 – ASW. Left: distance between N and the nearest O. Right: distance between the two O nearest to N. The distances are highlighted in green in the inset figure, which represents an example structure that belongs to BM1 (B(N – O$_W$) <3Å) and specifically to the subgroup BM1b (B(O$_{W1-W2}$) > 3.5Å). b) BE distributions for the system, calculated at ωPBE-D3BJ/def2-TZVP//HF-3c/MINIX level of theory, without including ZPVE correction, after being deconvoluted in the three binding modes contributions: green BM2, orange: BM1a, blue: BM1b. The average BE for each mode is also displayed as a solid line of matching color. c) Example structure for each binding mode.

the binding site ($B(O_{W1}-O_{W2})$), Figure 3.14a, right, reveals a first group of structures where only T1 is present, BM1a, ($B(O_{W1}-O_{W2})<3$) and one where both T1 and T2 are, BM1b, $B(O_{W1}-O_{W2})>3$. The latter requires the presence of a second water molecule to be close to N-atom, resulting in the two water molecules coordinated to NH_3 (W1 and W2) to be further apart, hence, the larger $B(O_{W1}-O_{W2})$ distance. In other words, in BM1b, NH_3 is more inserted in the HB network of the ice.

Energetically, the fact that BM1b structures establish more numerous HB interactions with the ice results in a larger range of BEs, involving the high BE tail of the distribution: ~ 12 kcal mol^{-1}. On the other hand, considering the lesser interactions established, BM1a structures are not significantly affected by the ice: the BE peak is narrower and centered around 9 kcal mol^{-1}. Deconvolution of the toal BE distribution in the 3 binding modes peaks is displayed in Figure 3.14,b. It appears that the presence of HB interaction where NH_3 is acceptor is pivotal to guarantee strong BE, and that the best energetic arrangement corresponds to the molecule being engaged in two interactions (BM1b).

The characteristics of the binding modes encountered for NH_3 and the following species are summarized in Table 3.4.

With respect to the role of NH_3 in the Strecker synthesis, the molecule is required to lose a proton in forming the aminoalcohol. When the reaction is water-assisted, the proton is delivered to the ice surface, and subsequently

Table 3.4: Categorization of the binding modes for Strecker synthesis species (first column). Second column reports the binding mode. The categorization is based on: total number of HBs established by the reactants with the ice (HBs, third column), HB type (forth column), as illustrated in the paragraph, and their average BE (fifth column, in kcal mol^{-1}). Last column indicates whether the binding mode is a potential reactive candidate for the Strecker synthesis (Y = yes; N = no).

Species	Mode	HBs	HB type	e.g. BE	reactive?
NH_3					
	BM1a	1	T1	8.6	N
	BM1b	3	T1,T2	11.8	Y
	BM2	1	T2	2.9	Y
H_2CO					
	BM1a	2	T3,T4	5.5	Y
	BM1b	3	T3×2,T4	9.5	Y
Aminomethanol					
	BM1a	3	T6,T1,T2	13.3	Y
	BM1b	2	T6,T2	9.6	Y
	BM2	2	T5,T1	7.4	N
$NHCH_2$					
	BM1a	2	T1,T4	11.8	Y
	BM1b	1	T1	7.3	Y
	BM2	2	T2,T4	3.8	N
HNC					
	BM1a	2	T2,T7	12.6	Y
	BM1b	1	T2	8.7	Y
	BM2	1	T7	5.1	N
Aminoacetonitrile					
	BM1a	4	T8,T1,T2,T4	12.6	Y
	BM1b	2	T8,T4	8.7	Y

to the carbonyl oxygen (proton relay mechanism). Such pathway is enabled by T2 interaction, which is present in BM1b and BM2, therefore only those binding modes are relevant for the synthesis.

H_2CO

H_2CO is also participating of the first step of the synthesis. According to the analysis in Figure 3.12, the BE distribution presents a single peak, usually synonymous to the existence of a single predominant binding mode. A summary of possible HB interactions that the system H_2CO-W establishes includes:

- T3: $H_W - O_{carbonyl}$ with H_2CO as acceptor and H_2O as donor
- T4: $H_{alkyl} - O_W$ with H_2O as acceptor and H_2CO as donor

where the sub-index 'alkyl' generally indicates H-atoms bound to a C-atom. In fact, the vast majority of the structures interacts with the ice through both T3 and T4 interactions, i.e. both the extremities of the molecules are engaged in HBs, binding mode denoted BM1 in Figure 3.12. Although, the variety in BE among the distribution is determined by additional factors, like steric impediment or multiple HBs. Among the latter, a rare and energetically favorable example is constituted by structures where the O-carbonyl is engaged in two HB-acceptor interactions (T3 $\times$ 2), with different water molecules.

Such mode corresponds to high BE tail of the distribution. An example can be found in Figure 3.15, labelled 'BM1a', while the average case structures where only a single T3 is established, are denoted BM1b. The ZPVE has not been included since the scope of the figure is to illustrate the different binding modes and corresponding BE ranges. ZPVE corrected BE values have been reported in the previous sections.

Regarding the Strecker synthesis, the role of H_2CO is to accept the proton from NH_3 side, mediated by the water, which corresponds to a T3 interaction, common to all the binding sites. Therefore, virtually, all the structures are potential reactive sites.

Aminomethanol

A large molecule such as the aminomethanol, product of the first step of the synthesis, present several binding modes. A complete census of them is out of the scopes of this thesis. Therefore, only the most relevant ones will be listed in the following. Among the possible HB interactions that the system NH_2CH_2OH-W establishes, there are:

- T1, T2, T4
- T5: $H_{alcoholic} - O_W$ with NH_2CH_2OH as donor and H_2O as acceptor
- T6: $H_W - O_{alcoholic}$ with NH_2CH_2OH as acceptor and H_2O as donor

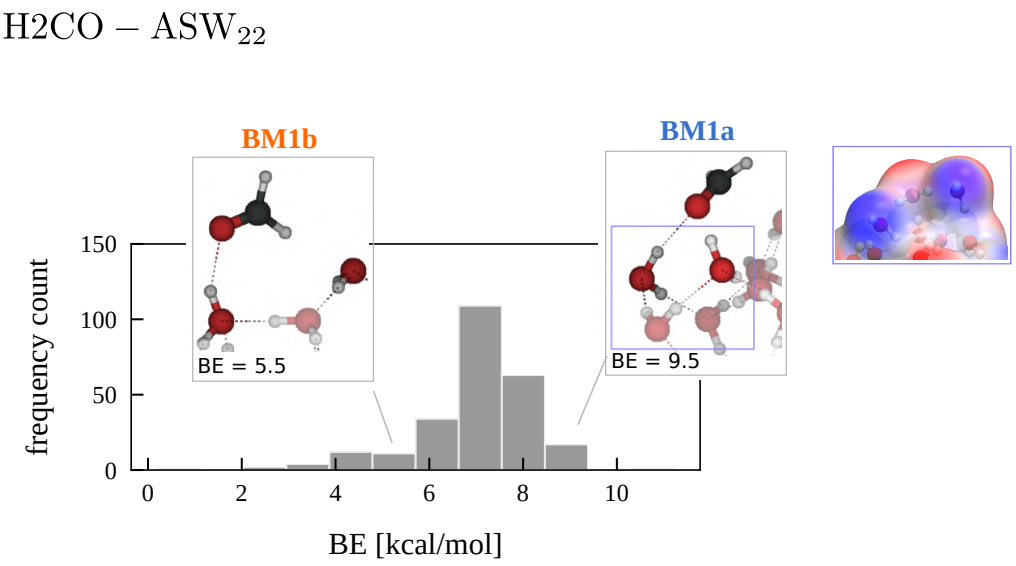


Figure 3.15: Histogram of the BE distribution of H_2CO, calculated at ωPBE-D3BJ/def2-TZVP//HF-3c/MINIX level of theory, without including ZPVE correction. An example of the two binding modes is reported, as selected from the distribution. The inset on the right represents a image of a map of electrostatic potential (MEP) of the ice region highlighted in BM1a: peculiar site that allows the double-HBs coordination of H_2CO. MEP color code ranging from blue (electron-poor regions, dangling-H sites) to red (electron-rich).

Some of the HB interactions have been previously encountered for other species (T1, T2, T4), while the other are characteristics of the aminomethanol. The principal two modes (BM1, BM2) (Figure 3.16) are defined by the arrangement of the alcoholic moiety of the molecule. The OH group can be pointing downward, engaged in donor HB with water (T5, which determines BM2 mode), or upward with respect to the cluster (T6, BM1). BM1 is more energetically favorable than the other. The orientation of the amino group (T1 or T2) is used to define subset of BM1 structures. In the average BM1 binding sites (at the center of the BE distribution, BE $\sim$ (kcal mol^{-1}), the amino group is usually establishing a donor interaction (T2): BM1b variation

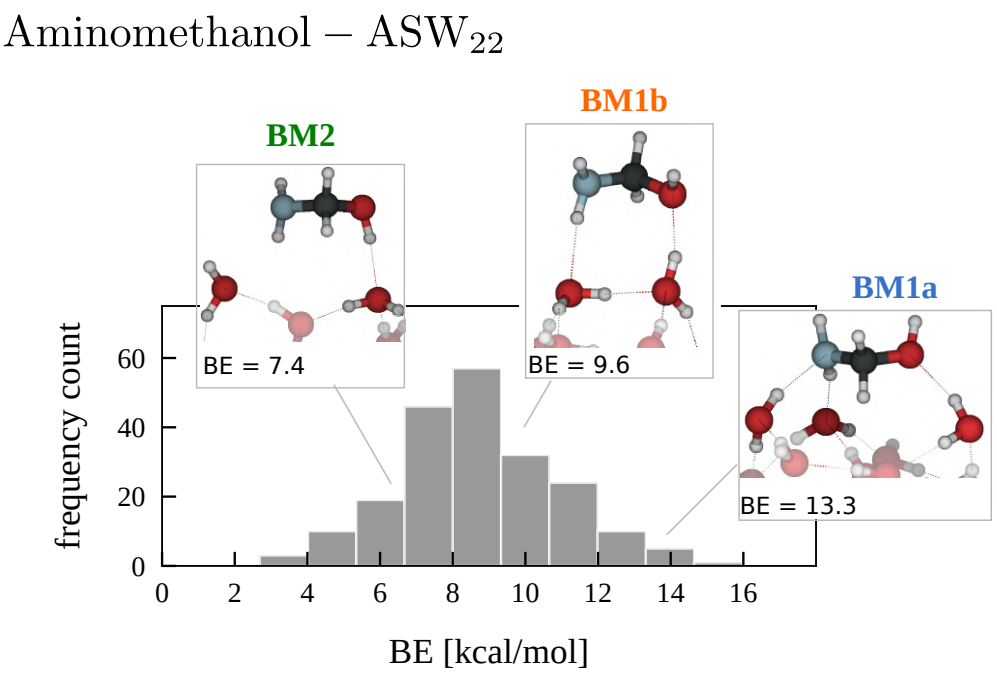


Figure 3.16: Histogram of the BE distribution of NH_2CH_2OH, calculated at ωPBE-D3BJ/def2-TZVP//BHANDHLYP/def2-SVP level of theory, without including ZPVE correction. An example of the three binding modes is reported, as selected from the distribution.

of BM1. The presence of both T1 and T2 (i.e. the amino group is highly inserted in the ice network) determines a very favorable situation labelled BM1a, corresponding to the high energy tail. Such hyper-coordination of the amino group is analog to the high BE mode found for NH_3.

In rarer cases, also $-CH_2$ group is coordinated to the surface (T4), but such weaker HB interaction is not determinant for the creation of new binding modes. Summary of the characteristics of the modes are listed in Table 3.4.

In order to undergo the second step of the synthesis - aminomethanol dehydration - the relevant interactions are: T6, that allows the protonation of

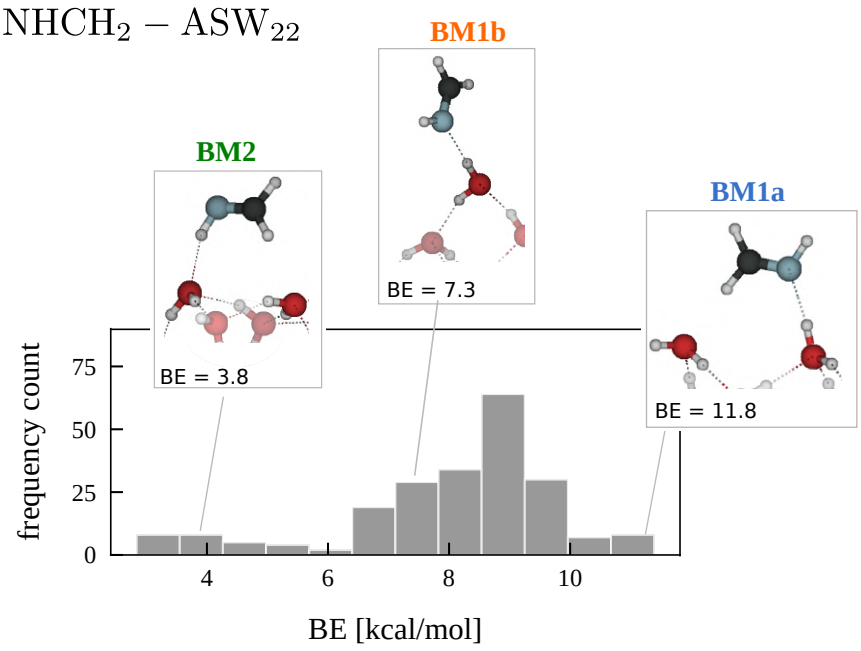


Figure 3.17: Histogram of the BE distribution of $NCCH_2NH_2$, calculated at ωPBE-D3BJ/def2-TZVP//HF-3c/MINIX level of theory, without including ZPVE correction. An example of the three binding modes is reported, as selected from the distribution.

the alcoholic group, and T2, that serves for the deprotonation of the amino group. Those two interactions are present solely in BM1 (in both variations), which makes BM1 more interesting for reactivity studies, while structures in the other binding mode, BM2, are loosely bound and are also not suitable to undergo dehydration.

$NHCH_2$

$NHCH_2$ is the product of the dehydration of the aminomethanol. The geometrical characteristic required to carry out the third step (the reaction with HNC), is for $NHCH_2$ to be oriented with -NH group pointing upward with

respect to the the surface, in order to receive the proton from the ice assisting water, i.e. establishing a T1 HB: This configuration correspond to the predominant binding mode between the ones we identified for $NHCH_2$ in Section 3.3.4.1, (BM1 89.0 %). It is also the binding mode where the species is more strongly bound, as opposed to the arrangement where -NH group is pointing downward (establishing a T2 interaction, binding mode BM2). Among BM1 structures, similarly to previous cases, the orientation of $-CH_2$ moiety defines subgroups of different BE range. The most favorable situation in terms of BE corresponds to both extremities of $NHCH_2$ engaged in HB (T1 + T4), as expected. An example of $NHCH_2$ binding modes is reported in Figure 3.17.

HNC

HNC is the other protagonist of the third step. The HB interactions that the system can establish are:

- T7: $C_{cyano} - H_W$ with HNC as acceptor and H_2O as donor
- T2

The analysis of HNC binding sites distribution, (see Binding Mode Analysis, Section 3.3.4.1) suggested that that are two main orientation of the molecule with respect to the ice: with the cyano group pointing upward

(BM1, predominant binding mode, 80.6 %), or downward (BM2). BM1 structures are bound *via* the hydrogen atom (T2 HB), which is also the only interaction required in order to undergo stage 3 reaction, as it guarantees HNC deprotonation. Among BM1 structures, only the strongly bound ones establish also T7, i.e. the molecule is bent on the ice, in a high BE configuration where both extremities are coordinated to the surface (BM1a). Example cases, as extracted from the distribution, are reported in Figure 3.18. In summary, BM1(a,b) structures constitute the high BE part of the distribution and are also suitable for the reaction, while BM2 cases are rare, therefore not representative, loosely bound (only T7 is established) and not likely fit to react with $NHCH_2$.

Aminoacetonitrile

As it was the case for the aminomethanol, large species present many binding modes, but the analysis will focus only on the major ones.

The system features a series of possible HB interactions:

- T1, T2, T4
- T8: $N_{cyano} - H_W$ with NH_2CH_2CN as acceptor and H_2O as donor

The only one necessary for stage 4 is T8, i.e. the HB that allows the protonation of the cyano group. The BE distribution (Figure 3.19) shows a single peak, hence a major binding mode (BM1). Inspection of the binding sites, in

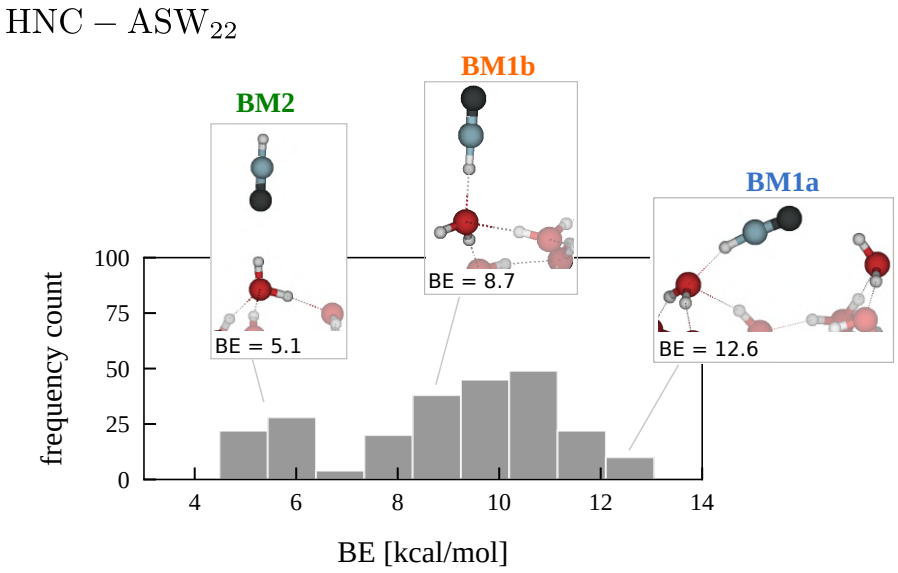


Figure 3.18: Histogram of the BE distribution of HNC, calculated at ωPBE-D3BJ/def2-TZVP//HF-3c/MINIX level of theory, without including ZPVE correction. An example of the three binding modes is reported, as selected from the distribution.

fact, disclosed that T8 is present in the vast majority of them. Whether the amino group is also coordinated to the surface, determines the high BE tail (BM1a) or the low BE tail (BM1b). However, almost all binding sites are potential reactive candidates.

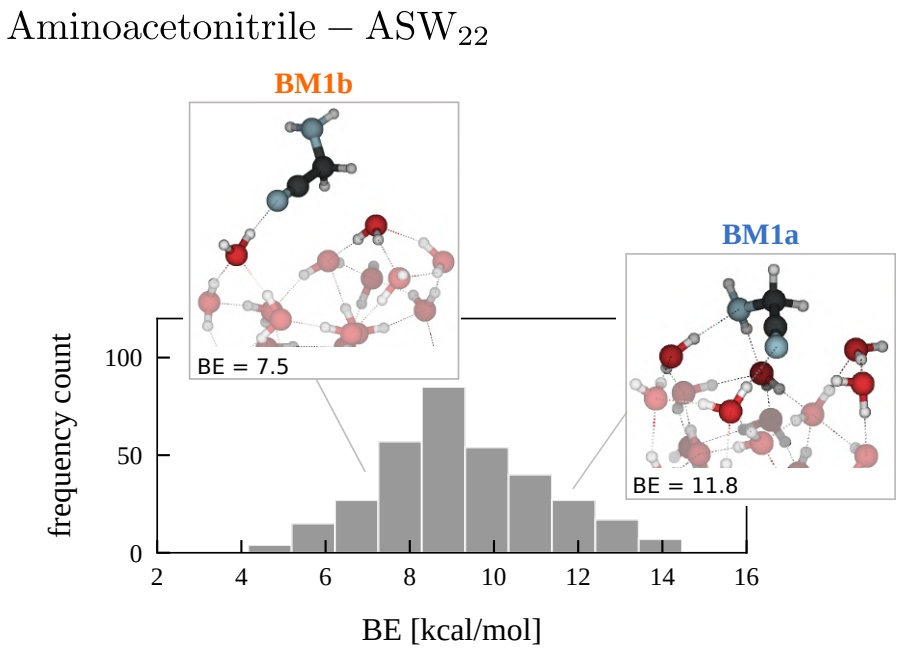


Figure 3.19: Histogram of the BE distribution of $NCCH_2NH_2$, calculated at ωPBE-D3BJ/def2-TZVP//HF-3c/MINIX level of theory, without including ZPVE correction. An example of the two binding modes is reported, as selected from the distribution.

3.3.5 Effect of cluster size on the Binding Energy

In order to evaluate the effect of the ice model size on the BE results, the computational procedure provided by BEEP, has been applied also to the set of clusters of different size at disposal ($ASW_{12,22,37,60}$). Details about cluster modelling and characteristics of the final structures can be found in Section 3.1.2. The analysis has been carried out for NH_3 species, since it presents several binding modes and a wide range of BE. In order to provide a meaningful comparison, all the binding sites have been obtained at the same level of theory, and due to the size of the larger systems ($ASW_{37,60}$), both geometries and energies have been computed using HF-3c/MINIX. It is worth mentioning that, according to the energy benchmark, the method overestimates the BE by 115 K with respect to ωPBE-D3BJ/def2-TZVP result.

The BE distribution computed with the 4 set of clusters are reported in Figure 3.20. The mean BE (μ) and standard deviation (σ) values of the distributions are listed in Table 3.5. There is a shift to high BE for the larger systems: μ increased by around 500 K passing from the 22-water molecules clusters to the 60-molecules ones. In term of shape, the BE distributions also get wider: σ of $NH_3 - ASW_{60}$ distribution is the double of $NH_3 - ASW_{22}$ one.

Another important effect of altering the size of the clusters is found in the proportion of the binding modes. Taking into account only BM1a and BM1b (see Figure 3.14), the modes that constitute the majoritarian peak of

Table 3.5: Mean (μ) and standard deviation (σ) of the Gaussian fit applied to the BE distribution of the systems NH_3 – ASW_S, where S = 12, 22, 37, 60 is the number of water molecules, calculated at HF-3c/MINIX//HF-3c/MINIX level of theory, without including ZPVE correction. The last column reports the ratio between the amount of structures that belong to BM1a and BM1b binding modes, described in the previous section.

System	μ [K]	σ [K]	BM1a / BM1b
NH_3 – ASW_{12}	4684	436	0.6
NH_3 – ASW_{22}	4742	317	0.5
NH_3 – ASW_{37}	5074	809	0.4
NH_3 – ASW_{60}	5287	742	0.2

the BE distribution, the evolution of the ratio (BM1a/BM1b) has been followed along the different models. In the smaller systems ($ASW_{12,22}$) binding sites where NH_3 is interacting solely via HB acceptor interactions with the ice (BM1a) are more frequent (ASW_{12}, ratio: 0.6) or equal (ASW_{22}, ratio: 0.5) to the binding sites where the molecule is more inserted in the ice HB network (BM1b). On the other hand, in the larger systems the proportion changes, reaching a ratio 0.2 for ASW_{60}, in favor of the hyper-coordinated motif. As pointed out in the previous section, BM1b structures are more affected by the site surroundings, therefore display a wider range of BEs that includes more strongly bound structures. The fact that BM1b is the favored pattern in the larger systems suggests that the cluster morphology allows higher insertion of the species in the HB network, besides, it enables the appearance of a high BE tail. Another explanation for the change in the modes ratio is relative to the amount of dangling-H atoms identified for the different clusters. The dangling-H sites on the ice surface appeared to diminish as the size systems

increased (Section 3.1). A binding mode such as BM1a, therefore, is expected to not be favored on larger clusters, since it requires the presence of a dangling-H pointing almost perpendicularly with respect to the surface plane (see Figure 3.14c). On the other hand, the dangling-H in BM1b is placed obliquely, in a configuration that might have been obtained by coerced deformation of the ice network, caused by NH_3 binding process.

In summary, the BE distributions obtained for the system NH_3 bound to ASW clusters of different sizes display a similar shape and average BE values. However, the larger systems ($ASW_{37,60}$) present a high BE tail, which suggest that the cluster morphology allows stronger interaction of the molecule. In fact, analysis the ratio between the major binding modes, revealed that, for larger systems, the motif where NH_3 is creating a larger number of HB interactions, i.e. is more inserted in the ice network, is the most frequent.

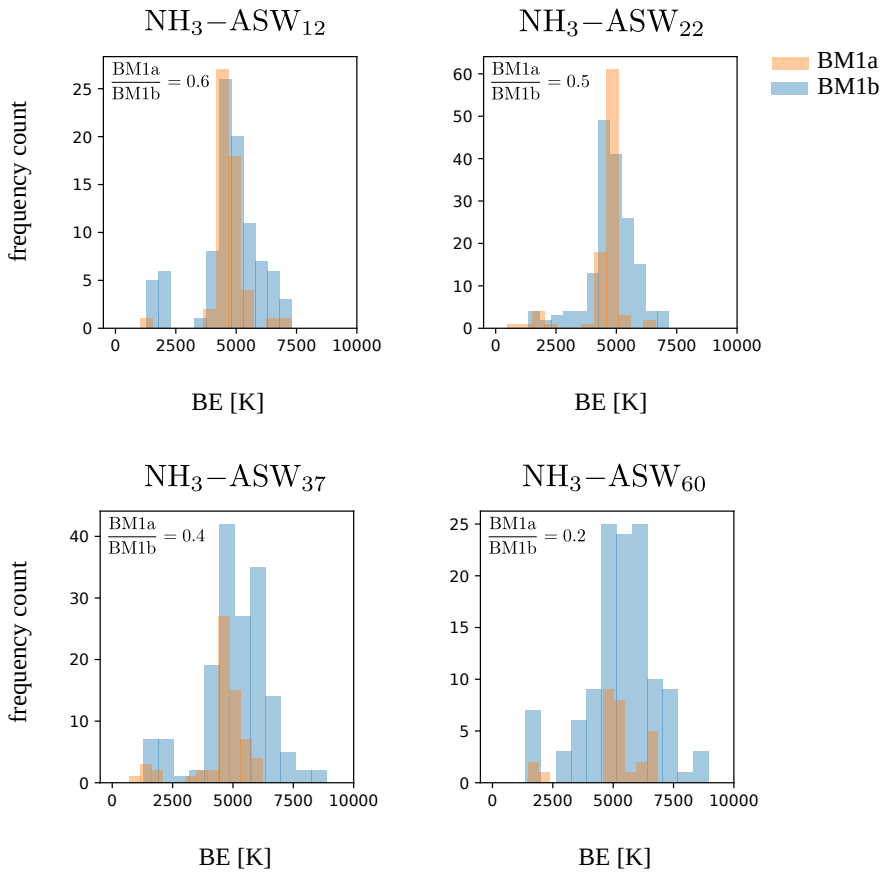


Figure 3.20: Histogram of BE distributions obtained for the system: NH_3 – ASW_S where S = 12,22,37,60 is the number of water molecules, calculated at HF-3c/MINIX//HF-3c/MINIX level of theory, without including ZPVE correction. Each plot also reports the ratio between the amount of structures that belong to BM1a (orange) or BM1b (blue) binding mode: BM1a/BM1b, described in the previous section.

3.3.6 Effect of surface topology on the Binding Energy

Finally, a more realistic ice model has been taken into account, aiming to analyze the effect of non-neglectable changes of the surface topology on the BE. The ice model, generated using a Machine-learned Potential (MLP) trained *ad hoc* (Section 2.2), is a amorphous periodic slab of 500-water molecules (Figure 3.4). The ice surface presents a variety of local environments, such as hollows regions ('valleys') and peaks ('crests'). The set of binding energies collected upon adsorption on such accurate ice model is expected to reflect the subtleties of the ice-molecule interaction, and provides a thorough description of it. The species in study is formyl radical (HCO), since it has been ubiquitously detected in the ISM and is known to be an intermediate in the grain surface formation of several prebiotic molecules. The computational procedure used to compute BE distributions using set-of-cluster model has been used as starting point, although adjustments have been necessary, especially for the sampling. A proper sampling of the periodic system required the definition of a geometrical grid on the surface, to make sure to include all potential binding sites. Specifics about the updated pipeline are reported in Section B. Binding sites optimization and BE calculation have been carried out using the MLP trained for the system in analysis, namely HCO – ASW.

Figure 3.21 reports the comparison between the BE distribution obtained

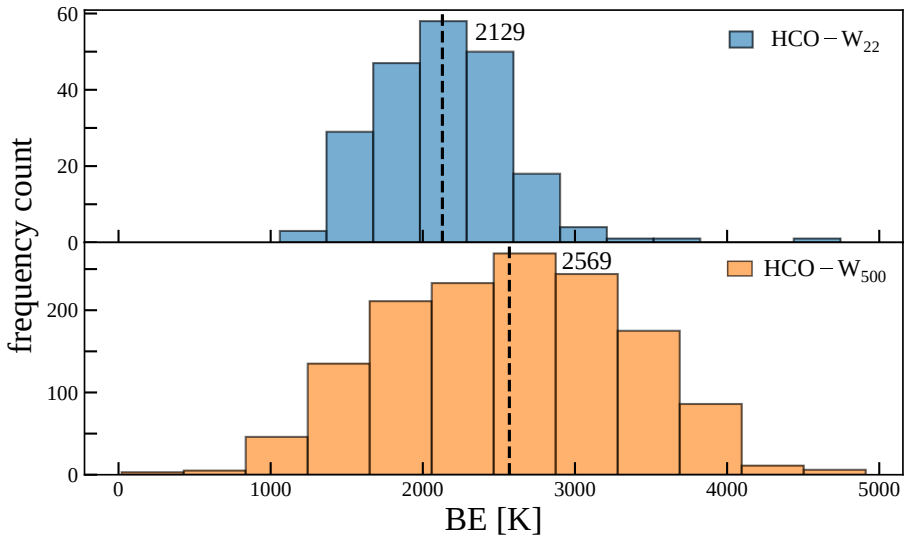
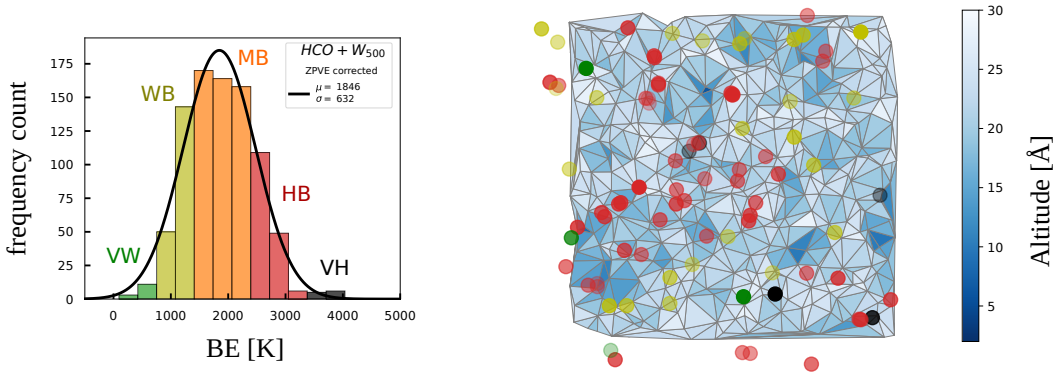


Figure 3.21: Comparison between the BE distribution obtained for the system HCO – ASW_{22}, set of clusters, upper panel, and HCO – ASW_{500}, set of periodic surfaces, lower panel, calculated using a machine learned potential specifically trained for the system, without including ZPVE correction. The figure also reports the average BE value, as a dashed line.

with the set of 22-water molecules clusters (re-calculated using the MLP) and the one obtained with the 5 periodic surfaces. The figure also includes the average BE of both distributions. The average value shifts to higher BEs in the larger systems, by about $\sim$ 450 K, which is imputed to the presence of a notably high BE tail. The result underscores the importance of using a accurate surface morphology, as the realistic model offers a diverse assortment of adsorbing situations on the surface, permitting stronger binding of HCO.

Analysis of the binding modes of the molecule did not disclose any new motives with respect to the findings for the set-of-clusters approach, hinting that the high BE structures are not due to more energetically favoured arrangements. Instead, they appear to be enabled by the surface morphol-

ogy. The BEs have been categorized into four categories depending on their strength: very weak binding (VW, up to 0.4 of the mean BE) weak binding (WB, 0.4–0.7 range), medium binding (MB, 0.7 – 1.2 range), strong binding (HB, 1.2–1.8) and very strong binding (VH, higher than 1.8 of the mean BE). Figure 3.22a displays a histogram of the BEs where the energies that belongs to each group are shaded in distinct colors. In order to understand the nature of the strong interactions, the spatial distribution of the different BE groups is investigated for one of the surfaces that constitute the HCO – ASW model. Figure 3.22b reports the altitude plot of the model, which is a image of the top view of the surface, employing a color scale representing the Z axis. The plot provides a visual measure of the deepness of the surface regions. Representation of the adsorbates center of mass as a dot of matching color with the color scheme defined above, allows to locate the set of binding sites obtained with the specific ice in their topological position on the surface. It is reminded that each binding site is obtained individually, the entire set has been condensed in one figure for the benefit of the visualization. The results indicate that binding sites associated to the high BE group (HB) are mostly located on the valley portions (darker color regions), while the weaker binding sites (WB/VW) are distributed on the higher lying crests (lighter color regions). The result is confirmed by the analysis of average altitude (Z coordinate of the center of mass) of each group of molecules. The average altitude of species belonging to the low BE tail (loosely bound) is about 4 Åhigher



a) b)

Figure 3.22: a) Histogram of the BE distribution obtained for the system HCO–W considering the 5 periodic surfaces, including ZPVE correction. Total of 880 binding sites. The figure also report the color code used to identify the BE groups, illustrated in the paragraph. b) Topological disposition of the different binding group for one of the 5 ice surfaces. The color code reflect the attribution in a). Each dot corresponds to the center of mass of the adsorbate in a binding site. MB sites are omitted for clarity.

than in high BE ones.

The findings reinforce the hypothesis that molecules adsorbed on the crests lack of strong interactions with the ice, hence, they have higher mobility, being the first to diffuse or desorb upon temperature increase of the substrate.

The topological position of the adsorbate can also be considered a index of the capacity of surface insertion of it in the surrounding water network. In fact, although hollow regions like concavities and small porous were already present on the ice surface, part of the high BE values has a large deformation energy (DE) associated to them. The DE is related to the BE according to

Equation 2.51. The fact that strong adsorption carries out higher DE, implies that the existence of strongly bound species is subordinated to the alteration of the water environment upon binding. This final conclusion opens to further investigations of the conditions necessary for strong adsorption, and the role and efficiency of diffusion processes, triggered by temperature increase, in 'filling up' ice nano-porous.

3.3.7 Discussion

Comparison with experimental results and previous theoretical studies

The BE values calculated in this thesis have been compared with available experimental results. Making a meaningful comparison of calculated BEs with experimental data is challenging, due to the variety of conditions under which the experiments are performed. In addition, the experimental data strongly depend on the pre-exponential factor used in the Polanyi-Wigner equation employed to derive the BEs[222] and the fitting procedure for obtaining BEs from TPD temperature curves. Among the experimental works, He et al. [57] has been taken into account. They presented TPD measurement of BEs of relatively simple molecules (N_2 , H_2, CO, CH_4, and CO_2) on a non-porous ASW (np-ASW) surface at monolayer (ML) and submonolayer coverage. In He et al. experiments, it is possible to distinguish between two situations in terms of the coverage (θ) of the target molecule on the surface. The low coverage limit ($\theta \rightarrow 0$), represents a situation in which mostly the binding sites of high BE would be occupied, corresponding to the high energy tail of the BE distribution. On the other hand, BE values obtained at the monolayer regime ($\theta \simeq 1$ ML) can be related with the mean of a BE distribution, where a variety of adsorption sites with different energies are occupied. The comparison shown in Figure 3.23, where the BE distributions are represented as box plots. Overall, the experimental results of these limiting coverage cases

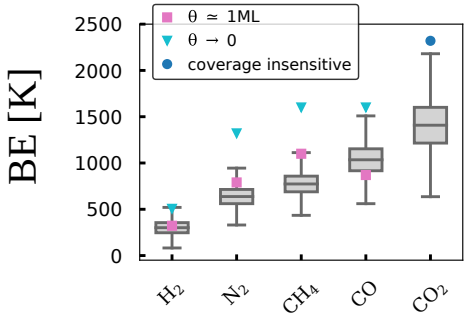


Figure 3.23: Box plot comparison between BE distributions presented in this thesis and experimental results[57], relative to different situations in terms of coverage (θ) of the target molecule adsorbed on the surface. The box plot median corresponds to the mean of the main binding mode.

coincide well with the computational values obtained in this thesis. The comparison is particularly good for H_2, N_2 and CO, (a difference of < 155 K in the low coverage regime and < 170 K in the ML regime) while the error for CH_4 is larger (a difference in low coverage regime of 207 K, and a difference in ML regime of 337 K). The results suggest that the approach of sampling a number of independent ASW clusters of a limited size (22 water molecules) properly reproduces the statistical nature of the interaction of those molecules with an actual ice surface.

Astrophysical implications

When comparing results between chemical experiments and quantum chemistry computations, a difference of 0.2-0.3 kcal mol^{-1}(corresponding approximately to 100-150 K) in the final BE is not substantial. However, for astrochemistry modelling, a few tens of Kelvins could largely affect the final out-

come. The molecular desorption is described by the Polanyi-Wigner equation, where its dependence on the exponential of the BE plays a crucial role in determining the efficiency of the process. To show this effect on a realistic, yet idealized, astrophysical case, the sublimation radius (the so-called snow line) has been calculated for a protoplanetary disk[65,223].

As expected, the position of the snow lines is heavily affected by the assumed BE. The sublimation radius changes by a factor of a few, for example in the case of H_2O, CO and CH_4, among the most important molecule involved in the process of planet formation, up to an order of magnitude in the case of e.g. CH_3OH or NH_3. Other species, like H_2, show larger differences; however they are not expect to form observable snow lines, since they are involved in other chemical processes that are not captured by a simplified disk model like the one taken into account. With a binding site distribution, within the framework of this idealized disk model, the snow line is expected to get smoother, thanks to the interplay between the temperature density profile, and the BE distribution. Increasing the distance from the star, and consequently decreasing the dust temperature, the number of available lower-energy binding sites will grow, depending on the broadening of the distribution. An accurate determination of the BE is then fundamental to quantitatively assess snow line positions in planet-forming regions and evaporation fronts during star-formation.

3.4 Strecker-type reactions on small water clusters: propaedeutic study

This thesis is focused on the Strecker synthesis of glycine. The synthesis is comprised of 4 stages, which have been listed in Section 3.2.2. Each one is studied employing a set of small water clusters ("Small systems", Section 3.5), and on the ASW model composed of 22-water molecules (Section 3.5.

Studying the small systems is propaedeutic for simulating the larger systems, particularly in elucidating the reaction mechanisms involved. Analyzing these mechanisms through the Intrinsic Reaction Coordinate (IRC) provides a detailed outline of the events occurring along the reaction coordinate, enabling meaningful comparisons when investigating reactions on the ASW surface.

3.4.1 Stage 1: Aminomethanol formation

The first stage (S1) of the synthesis is the nucleophilic addition of NH_3 to H_2CO to give aminomethanol (NH_2CH_2OH). Three different systems are studied: (NH_3 + H_2CO + W_X)$_{X=0-2}$ corresponding to ($S1-W_X$)$_{X=0-2}$, where $S1-W_0$ is the reaction in gas-phase, and $S1-W_{1-2}$ represent situations where the reactants are coordinated to a minimal water cluster (monomer and dimer, respectively). The reaction in study is composed of one elementary

step; energy barriers and reaction energies are reported in Table 3.6. Structural parameters of the optimized structures can be found in Appendix, Table E.1.

Table 3.6: Energy barriers ($\Delta E^{\ddagger}$) and reaction energies (ΔE^{o}) computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries. Values in kcal mol^{-1}.

System	$\Delta E^{\ddagger}$	ΔE^{o}
$S1-W_0$	32.8	-9.9
$S1-W_1$	14.6	-11.3
$S1-W_2$	10.1	-12.3

The role of the water molecules in $S1-W_{1,2}$ is assisting in proton transfer reactions, establishing a bridge which mediates the passage of the proton from one reactant to the other. In the reaction in study, one of NH_3 protons is transferred to the O-carbonyl of H_2CO, through the water cluster, which entails a considerable decrease in the reaction barrier, passing from 32.8 (gas-phase) to 10.1 kcal mol^{-1} ($S1-W_2$). Besides, the exoenergetic character of the reaction is maintained in all systems, with slightly increasing values of ΔE^{o}. The zero-point vibrational energy of the reactions (ZPVE) has not been included.

The presence of assisting water molecules exerts a effect on the geometry of the reaction site as well as on the electrostatic properties of the reactants. In terms of geometry, the structural changes required in order to reach the transition state (TS) geometry are smaller in $S1-W_{1-2}$ compared to the gas-

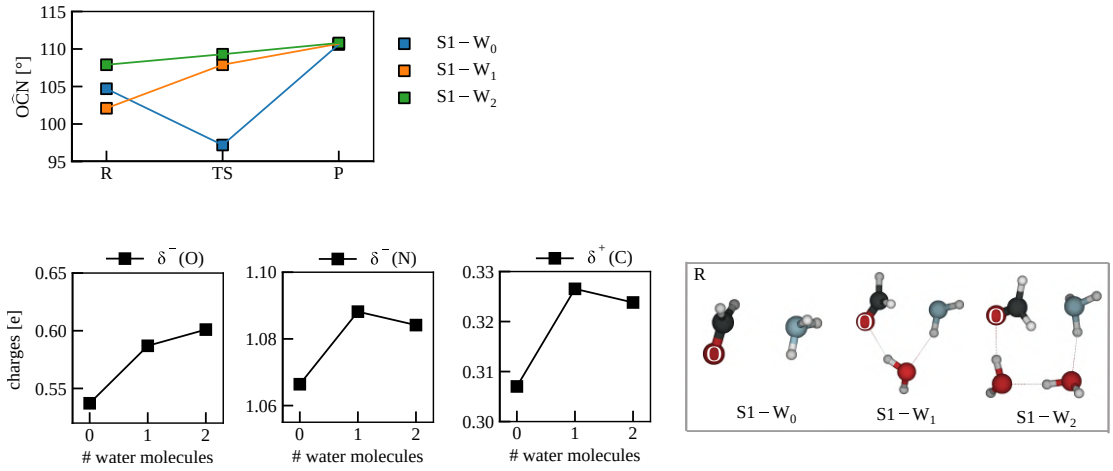


Figure 3.24: Upper panel: variation of the bond angle, A(O-C-N), for the small systems along S1 reaction coordinate. Lower left panel: NAO partial charges on the main atoms for the systems, in the reactant geometries (R), shown in lower right panel.

phase. As an example, Figure 3.24, upper panel, reports the variation of the bond angle A(O-C-N) along the reaction coordinate. Among the three systems, $S1-W_0$ shows the largest deformation in passing from reactants to TS geometry, and also in passing from TS to product geometry, which corresponds to a larger activation energy.

The participating water molecules also alter the electrostatic properties of the reactive centre. For the water-assisted processes to take place, the water molecule(s) has to be placed as hydrogen bond (HB) donor with respect to H_2CO, as well as HB acceptor to NH_3 group, to mediate the proton transfer from the latter to the former. The effect on H_2CO is the polarization of the carbonyl group, increasing the electrophilic character of the C-atom and facilitating the formation of $C-N$ bond. Water also acts as a HB acceptor on

NH_3, slightly increasing the electronegativity of N-atom and exerting a weakening effect on one of the N – H bonds, in turn favoring the proton transfer. Those inductive effects can be estimated by evaluating partial charges on the reactants geometries. For this study, the natural population NAO charges are taken into account (Methods, Section 2.7). The tendency, displayed in Figure 3.24, upper panel, is the increase of the partial charges on the backbone atoms of the reactants (O, C, N) with the number of water molecules present in the reaction environment. The increment is up to 12 % for the O-carbonyl atom, while less pronounced on N and C atoms.

Overall, the water environment provides a catalytic effect, facilitating the structural rearrangements necessary to reach the activated complex and increasing NH_3 nucleophilic and H_2CO electrophilic character respectively, which results in lower reaction barriers.

Stage 1: Reaction mechanism

In order to elucidate the mechanism of the reaction, the intrinsic reaction coordinate procedure (IRC = ξ) is used to locate the reaction paths. Energy and reaction force profiles (see Method Section 2.4) are presented in Figure 3.25 for $S1-W_{1-2}$, and Appendix Figure E.1 for $S1-W_0$. The analysis of the reaction force profiles, lower panel, allows to define different reaction events that take place along the reaction coordinate. The TS region of the IRC profile is defined within the minimum and maximum of the reaction force pro-

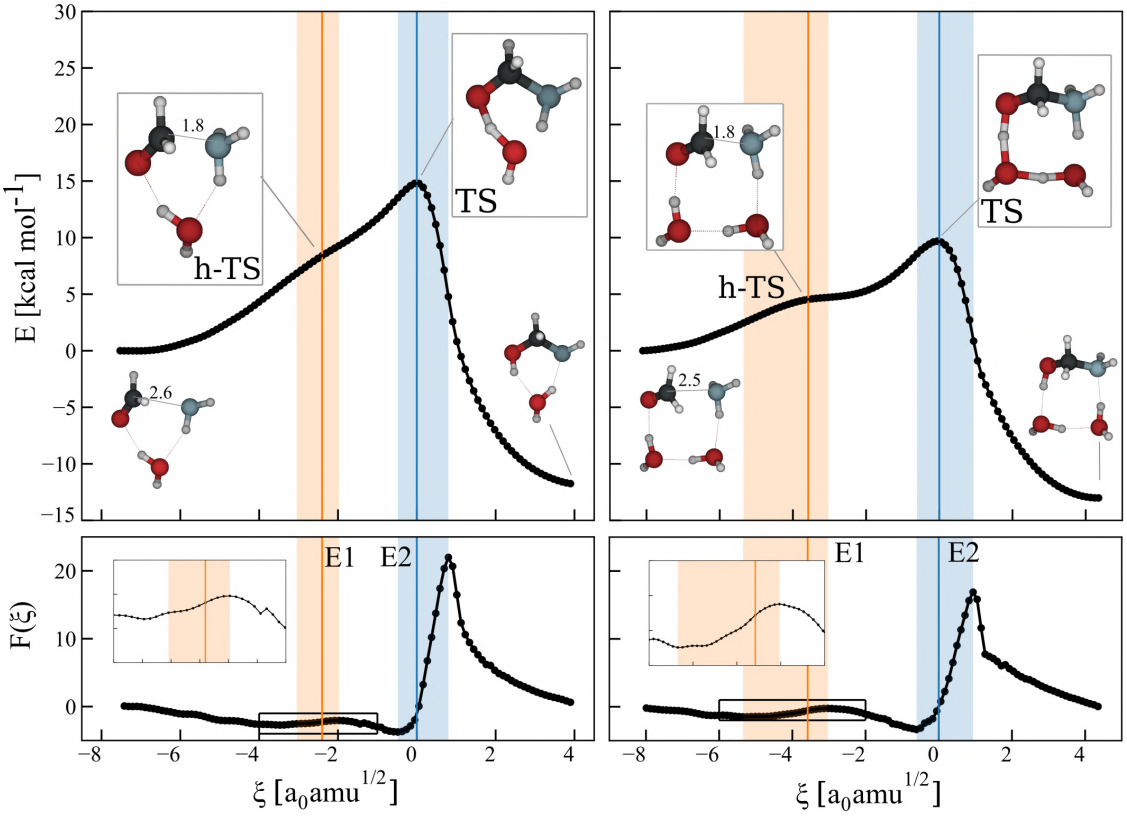


Figure 3.25: Energy (upper panels) and reaction force profiles (lower panels) for S1 – W_{1-2}, using BMK/def2-TZVP//BHANDHLYP-D3BJ/def2-SVP levels of theory. Blue and orange lines represent TS and h-TS. TS and h-TS regions are displayed as orange and blue shadowed areas.

file, and is displayed as a blue shadowed area. It has been pointed out[176] that primarily electronic rearrangements occur in the TS region, whereas outside this region, structural modifications are predominant. However, additional inflection points in the reaction force profile (inset Figure 3.25 lower panel) suggest that a second incipient reaction event might be present before the TS region. In fact, all three systems present additional critical points, located before the TS, that can be used to define a 'pseudo TS'. This event corresponds to a shoulder on the IRC profiles, clearly visible for S1 – W_{1-2} and less pro-

nounced in S1 – W_0. Such events are considered *transient* point along the reaction path and are associated with hidden transition states (marked as 'h-TS') and corresponding hidden intermediates ('h-I'), which can be converted into real TS and I in presence of a change in the environment conditions or substitution pattern of the reactive fragments[177]. Similarly to the TS regions, event regions can be defined for the h-TS as well, by means of the local minimum and maximum of the force profile. Those are shown as orange shaded regions in Figure 3.25.

In order to correlate the bond breaking/forming processes of S1 to the two events identified along the reaction path, the relevant bond distances evolution along the reaction coordinate are analyzed. As displayed in Figure B.1 upper panel, C – N distance (solid line) decreases linearly until overcoming the h-TS region, followed by a change in the slope, meaning that the bond is established. Consequently, C – N formation (event 1, E1) is correlated to the h-TS for the three systems. On the other hand, the bond distances related with the proton transfer (N – H rupture from the ammonia side and O – H formation on the formaldehyde side, dashed and dotted lines) show meaningful slope alterations entering the TS region, thus linking the second event (E2) to the TS. Schematic representation of S1 – W_2 reaction events is reported in Figure 3.27.

The nature of S1 events is confirmed also by the Wiberg bond orders and bond order derivatives along the reaction coordinate, that have been calcu-

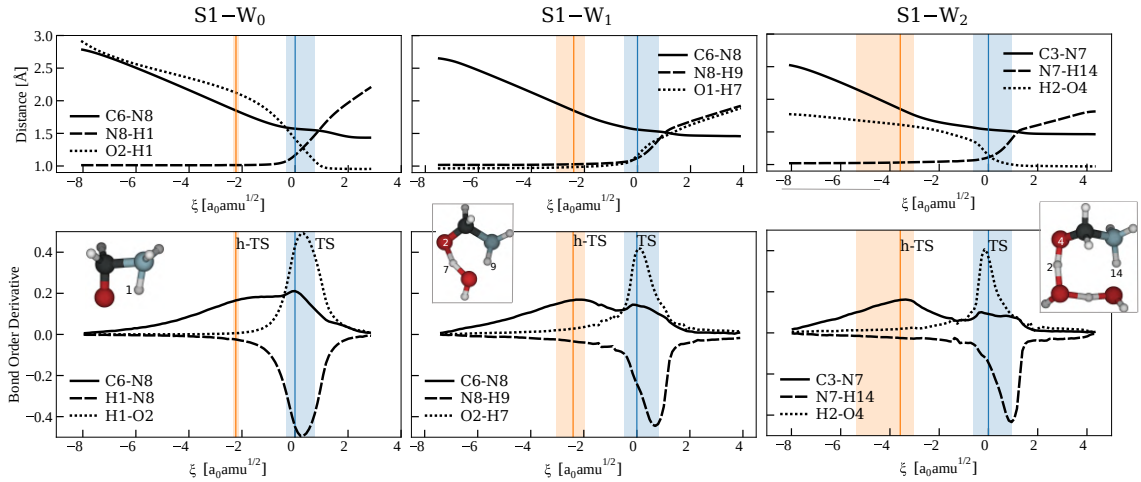


Figure 3.26: Upper panels: relevant bond distances evolution along the reaction coordinate for $S1-W_{0-2}$. Lowe panels: corresponding bond order derivatives. Blue and orange lines represent TS and h-TS. TS and h-TS regions are displayed as orange and blue shadowed areas.

lated for the main bond involved in the reaction (Figure B.1, lower panel). A negative sign in the derivative indicates bond weakening or dissociation, while a positive sign accounts for bond formation or strengthening. The major change in the bond order describing the formation of the C-N bond is located within the h-TS region, while the proton transfer processes are located in the TS region. The bond order derivative plots also provide information about the synchronicity of the processes in E2. For the system $S1-W_0$, the major change in both bonds coincides; while for the water-assisted systems the carbonyl protonation takes place first: $O-H$ forming peak is found to be at the beginning the TS region, while $N-H$ breaking is closer to the end. The asynchronicity is more pronounced in $S1-W_2$. Finally, estimation of the electrostatic potential of $S1-W_2$ allows to characterize the critical points

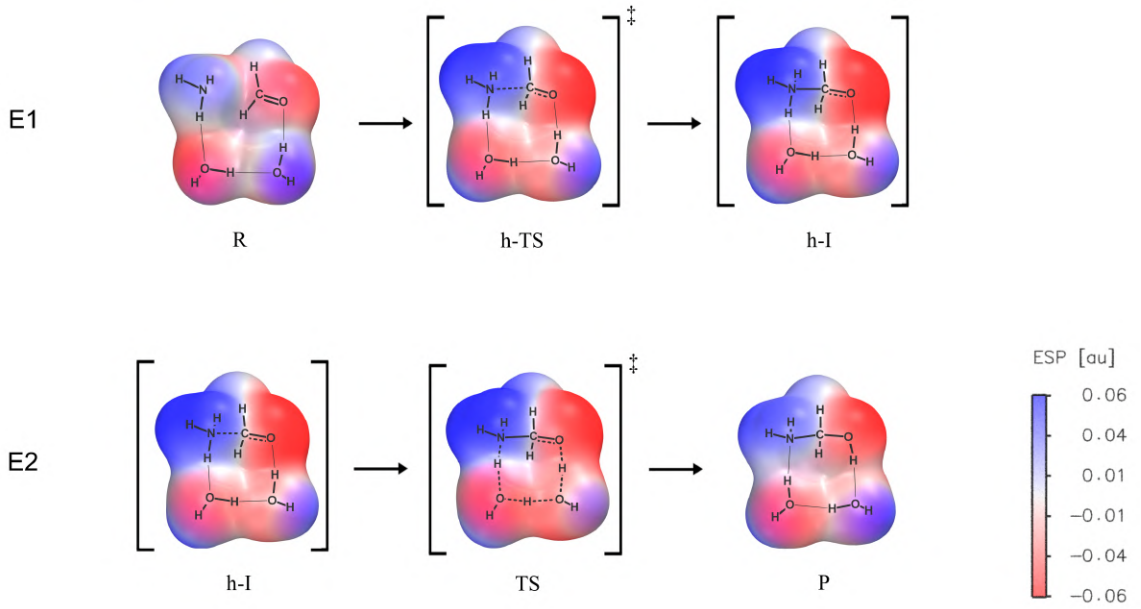


Figure 3.27: Schematic reaction mechanism for S1 – W_2, and MEPs of the minimum energy points. Blue corresponds to electron-poor regions while red to electron-rich ones. The structures have been extracted from the IRC plot. Hidden TS (h-TSs) correspond to inflection points on the reaction force while hidden intermediates (h-Is) have been identified among the points immediately following the event region's area.

of the energy profile. Figure 3.27 displays the corresponding Maps of Electrostatic Potential (MEPs) which use color to indicate electron-rich (red) and electron-poor (blue) regions. All the structures have been extracted from the IRC; h-TS corresponds to a inflection point of the reaction force while the tentative hidden intermediate (h-I) has been selected among the points immediately following E1 region area, as it is not an well defined state.

E1, C – N bond formation, corresponds to a significant shift of the electron density in H_2CO fragment toward the O-carbonyl, as indicated by its much more vivid red color compared to reactants configuration. NH_3 frag-

ment is also strongly electron-depauperated in creating C–N bond, such that the dipolar hidden intermediate h-I ($^{-}OCH_2\cdots NH_3^{+}$) results to be the structure that presents the largest intramolecular polarization. Once the bond is established, the proton transfer processes take place (E2). The MEP of the TS reflects the delocalization of O-carbonyl electron density in forming the O–H bond (attenuation of the red colored area). Also, it is possible to appreciate the change in the position of the lone pairs on the oxygens of the assisting water, as the proton transfer events finalize and the aminomethanol is formed.

In summary, Stage 1 mechanism presents two reactive events (E1 and E2) which are asynchronous: E1, C–N bond formation, is the first to happen and leads to a dipolar hidden intermediate which is then converted to the TS, followed by E2, the proton transfer, which connects the TS to the product.

Stage 1: Partition of the reaction barrier

The integration of the Reaction force profiles allows to quantify the energy associated to the different reaction events (see Methods, Section 2.4). The total energy barrier for this stage, $\Delta E^{\ddagger}$, is therefore partitioned as:

$$\Delta E^{\ddagger} = \Delta E^{\ddagger}_{E1} + \Delta E^{\ddagger}_{E2} \tag{3.7}$$

Where $\Delta E^{\ddagger}_{E1}$ is the energy barrier associated to formation of the hidden TS and $\Delta E^{\ddagger}_{E2}$ is the energy necessary to convert the hidden intermediate to the

TS.

Table 3.7: Partition of the energy barriers ($\Delta E^{\ddagger}$) according to Equation 3.7, in kcal mol^{-1}, as calculated from the force profile in Figure 3.25. Reaction energies (ΔE^{o}) have been included as well.

System	$\Delta E^{\ddagger}_{E1}$	$\Delta E^{\ddagger}_{E2}$	$\Delta E^{\ddagger}$	ΔE^{o}
$S1-W_0$	15.1	17.0	32.9	-8.9
$S1-W_1$	7.8	6.7	14.8	-11.7
$S1-W_2$	4.3	5.2	9.7	-13.0

As can be seen from Table 3.7, the reaction in gas-phase displays high value for both E1 and E2 barriers: 15.1 and 17.0 kcal mol^{-1} respectively. For the water assisted systems, while the magnitude of E2 barrier is comparable (6.7 and 5.2 kcal mol^{-1}), E1 barrier diminishes considerably in $S1-W_2$, being of only 4.3 kcal mol^{-1}, vs 7.8 kcal mol^{-1} of $S1-W_1$, meaning that the difference in the total S1 barrier resides in the entity of E1 barrier. Such decrease in E1 barrier can be imputed to to the comparatively lesser strained configuration of $S1-W_2$'s h-TS. In fact, $S1-W_2$'s h-TS is characterized by a 5-members ring (taking into account all atoms but the hydrogens), compared to the 4-members ring in $S1-W_1$ and the 3-members ring in $S1-W_0$, where there are no water molecules involved.

Stage 1 on ice mixture

A set of small systems $(NH_3+H_2CO+X_2)_{X=W,NH_3}$ is treated as a model system for the reaction involving the larger ice mixture surface (Section 3.5.1.2).

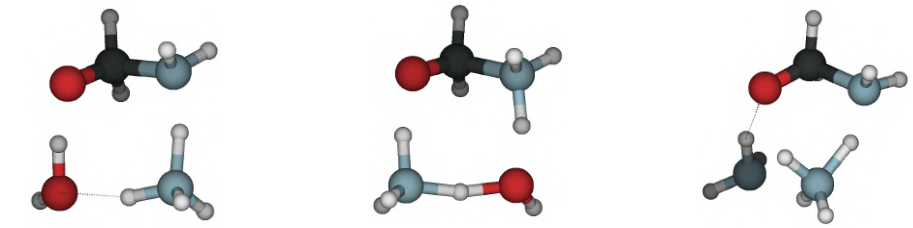


Figure 3.28: TS geometries computed at BHANDHLYP-D4/def2-SVP level of theory for the three systems in analysis: a) $S1-W-NH_3$, where NH_3 is HB-bonded to surface's NH_3; b) $S1-NH_3-W$, where H_2CO is HB-bonded to surface's NH_3; c) $S1-2NH_3$, where both reactants are HB-bonded to surface's NH_3 molecules.

Table 3.8 reports energy barrier results, while TS geometries are shown in 3.28. The exothermicity of S1 process is maintained and all three cases display larger TS barrier than $S1-W_2$. The most favorable situation appears to be Figure 3.28a, where NH_3 (surface) is coordinated to NH_3 (reactant), acting as HB acceptor, with a energy barrier of 13.7 kcal mol^{-1}. On the other hand, the least energetically favorable is $S1-(NH_3)_2$ ($\Delta E^{\ddagger}$ = 17.1 kcal mol^{-1}). This system is also not realistic, representing a reactive site with three NH_3 molecules in close proximity to each other.

Table 3.8: Energy barriers ($\Delta E^{\ddagger}$) and reaction energies (ΔE^{o}) computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries. N_X indicate the number of molecules involved in the proton relay. Values in kcal mol^{-1}.

System	N_X	$\Delta E^{\ddagger}$	ΔE^{o}
a) $S1-W-NH_3$	2	13.7	-12.6
b) $S1-NH_3-W$	2	16.3	-12.1
c) $S1-2NH_3$	1	17.1	-10.6
S_1-W_2	2	10.1	-12.3

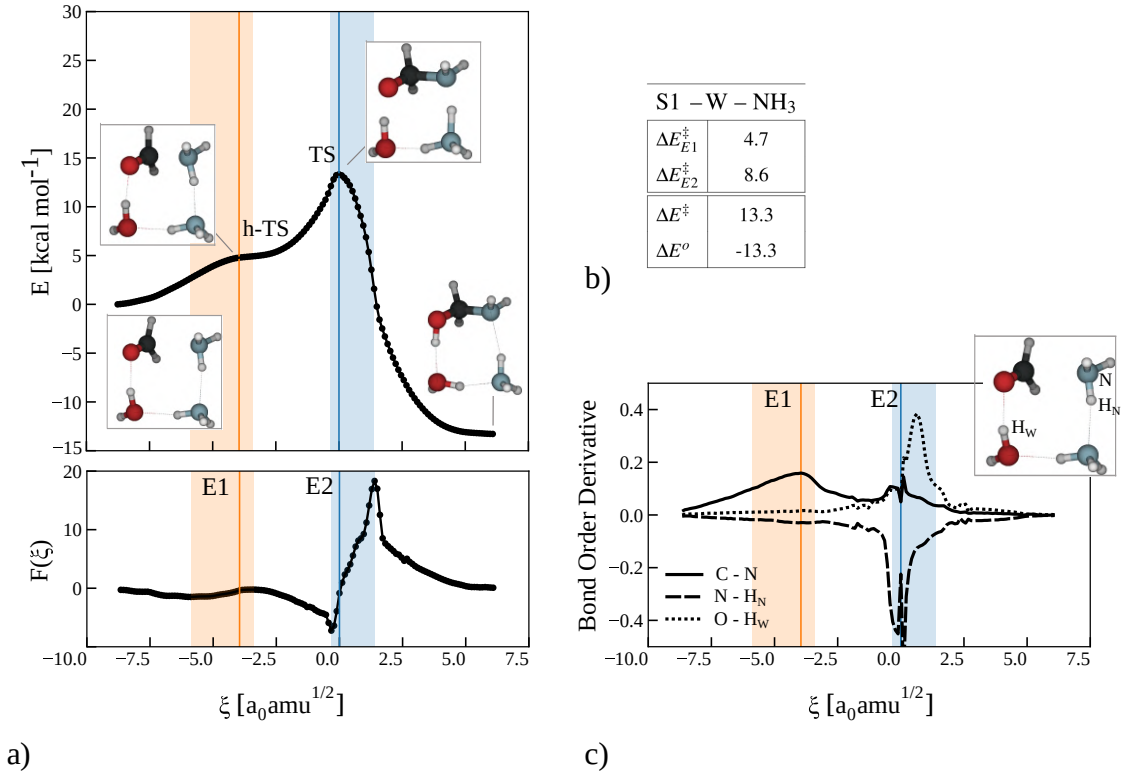


Figure 3.29: a) Energy (upper panel) and reaction force profiles (lower panel) for S1 – W – NH_3, using BMK/def2-TZVP//BHANDHLYP-D3BJ/def2-SVP levels of theory. Blue line represent the TS. TS region is displayed as shadowed area. Minimum energy structures have been reported as extracted from the energy profile. b) Table that reports TS energy ($\Delta E^{\ddagger}$) and reaction energy (ΔE^{o}) extracted from the IRC, as well as the partition of the barrier in $\Delta E^{\ddagger}_{1}$ (relative to h-TS) and $\Delta E^{\ddagger}_{2}$ (TS). c) Bond order derivative for the main bond distances involved in S2.

IRC analysis for system a) S1 – W – NH_3, the one that displays the lowest energy barrier, (Figure 3.29a) confirmed that mechanism is concerted and composed of the aforementioned events E(1,2). The main difference from S1 – W_2 is in the the bond breaking/forming steps of E2: as illustrated by the bond order derivative analysis in Figure 3.29c, the peak corresponding to N – H rupture is anticipated with respect to O – H forming. This is likely

due to the surface's NH_3 exerting a stronger inductive effect on the reactant, compared to the assisting water in $S1-W_2$. The hypothesis is confirmed by the NAO partial charges computed on NH_3 atoms for the two systems, showing larger intramolecular polarization of $N-H$ bond for $S1-W-NH_3$ than in $S1-W_2$.

3.4.2 Stage 2: Methanimine formation

The second stage (S2) is the dehydration of aminomethanol, product of S1, yielding methanimine ($NHCH_2$). The reaction occurs *via* the migration of one of the protons of the amino groups to the alcoholic group of the molecule. IRC calculation for the system $(NH_2CH_2OH+W_2)$, $S2-W_2$, is carried out in order to elucidate S2 mechanism. The plot is reported in Figure 3.30a, upper panel. Analysis of the reaction force profile (Figure 3.30a, lower panel) and the bond order derivative (Figure 3.30c) shows that the dehydration proceeds in a single and asynchronous reactive event (E1).

The bond order derivative peaks indicate that $C-O$ rupture takes place first (solid line), followed by the protonation of the alcoholic group (dotted line) and finally the breaking of $N-H$ bond. That is, the proton relay is initiated by the water molecule acting as HB donor on the alcoholic moiety of the aminomethanol.

In summary, S2 mechanism is concerted and comprised of a single slightly asynchronous event. The first steps to take place are $C-O$ rupture in the

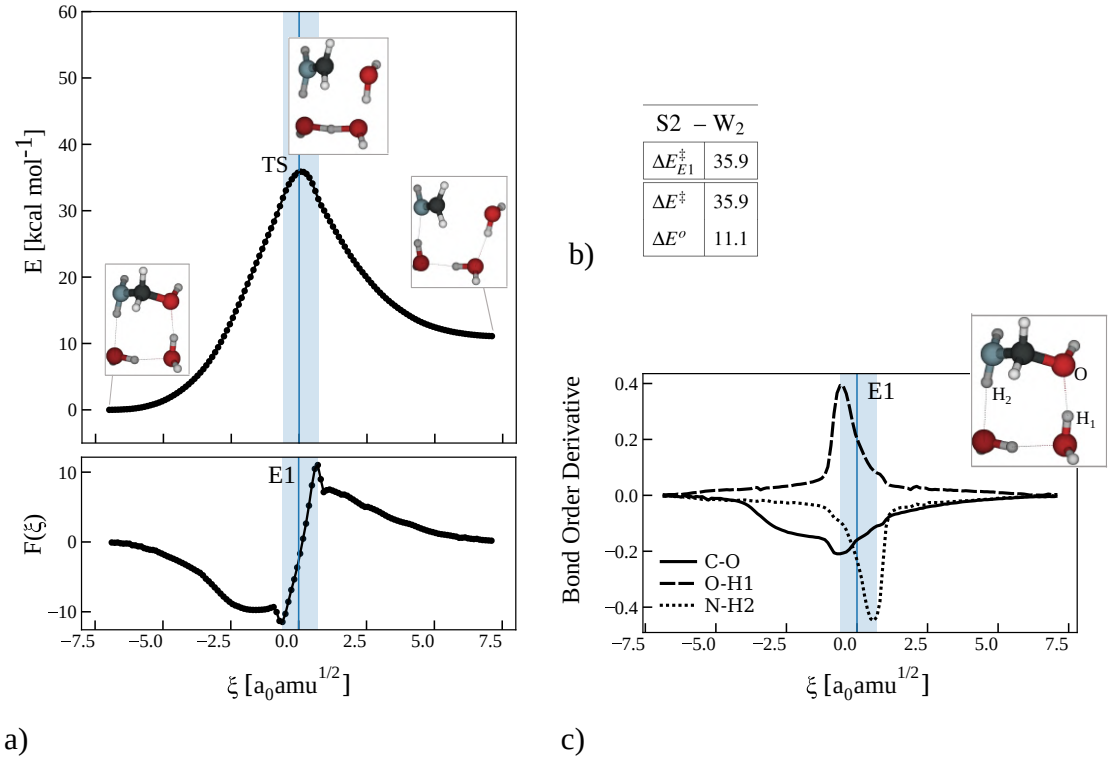


Figure 3.30: a) Energy (upper panel) and reaction force profiles (lower panel) for S2W$_2$, using BMK/def2-TZVP//BHANDHLYP-D3BJ/def2-SVP levels of theory. Blue line represent the TS. TS region is displayed as shadowed area. Minimum energy structures have been reported as extracted from the energy profile. b) Table that reports TS energy ($\Delta E^{\ddagger}$) and reaction energy (ΔE^{o}) extracted from the IRC. c) Bond order derivative for the main bond distances involved in S2.

aminomethanol backbone, along with the protonation of the alcoholic group, leading to the conclusion that the reaction mechanism present a *late* character in terms of the water formation: once the TS geometry is reached, the water molecule has already been released.

3.4.3 Stage 3: Aminoacetonitrile formation

Stage 3 (S3) is the nucleophilic addition of HNC to methanimine (produced in S2), yielding aminoacetonitrile (NH_2CH_2CN).

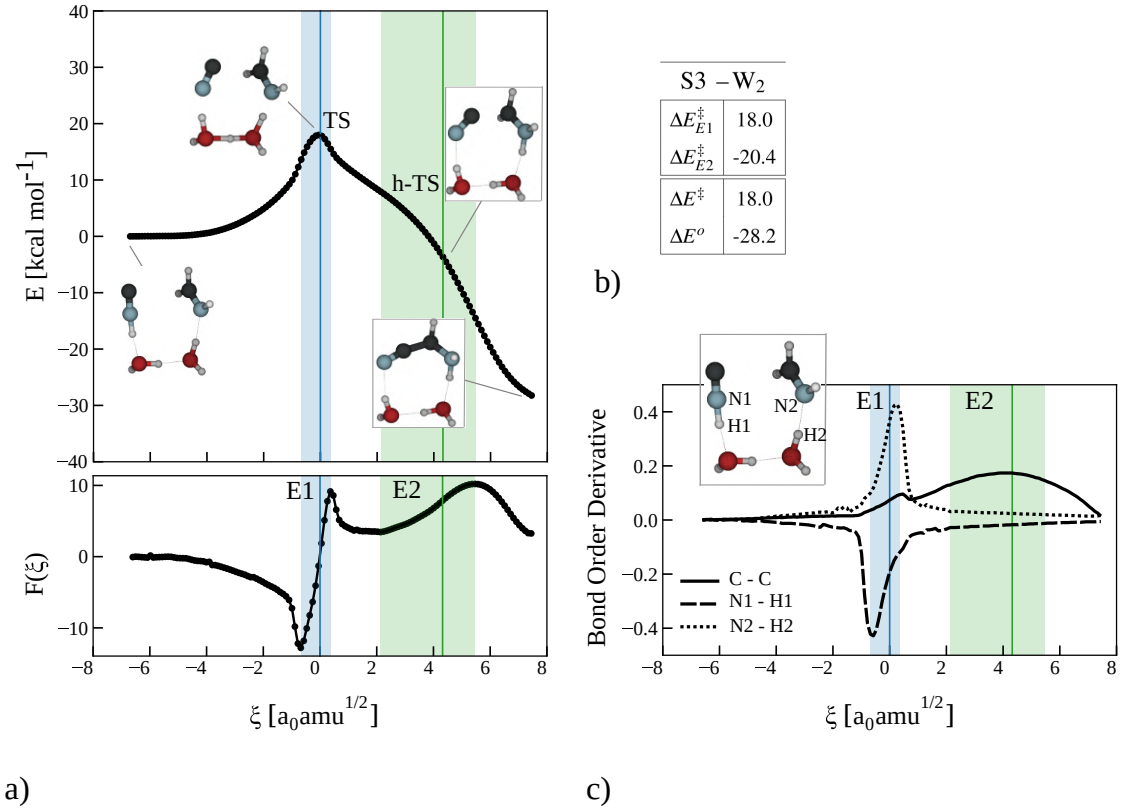


Figure 3.31: a) Energy (upper panel) and reaction force profiles (lower panel) for $S3W_2$, using BMK/def2-TZVP//PW6B95-D3BJ/def2-SVP levels of theory. Blue and green lines represent TS and h-TS respectively. TS and h-TS regions are displayed as blue and green shadowed areas. R, TS, h-TS and P structures have been reported as extracted from the energy profile. b) Table that reports TS energy ($\Delta E^{\ddagger}$) and reaction energy (ΔE^{o}) extracted from the IRC, as well as the partition of the barrier: $\Delta E^{\ddagger}_1$ is relative to the TS, while $\Delta E^{\ddagger}_2$ (h-TS) has been included for completeness. c) Bond order derivative for the main bond distances involved in S3.

The reaction mechanism is analyzed for $(HNC + NHCH_2 + W_2)$, $S3 - W_2$.

The analysis of energy and reaction force profiles (Figure 3.31a), revealed that the reactive events in the nucleophilic addition are asyncrounous.

The first event (E1) is the proton transfer between the reactants and it is associated to the TS region. From the bond order derivatives (Figure 3.31c), it appears that the proton relay is initiated by the loss of the proton of HNC, followed by $NHCH_2$ protonation. The second event (E2) is the aminoacetonitrile backbone formation and it manifests as a shoulder in the energy profile, corresponding to a h-TS located in the relaxation part of the reaction coordinate. The feature associated to this event in Figure 3.31c, is $C-C$ bond formation, a broad peak in correspondence of the h-TS region. It is noteworthy that the order of the nucleophilic addition events is reversed with respect to the founding for previous stages, where the backbone bond forming (S1) or breaking (S2) happened first with respect to the proton transfer. The reason for the change of mechanism in S3, lies in the high reactivity and relatively strong acidity of HNC, which is more susceptible to proton loss upon perturbation[224], compared to NH_3 in S1 or H_2O in S2.

The schematic reaction mechanism of S3 is reported in Figure 3.32 along with MEPs of the relevant structures as extracted from the IRC profile. The reactants configuration presents clear charge separation in the system: HNC side of the pre-reactive complex corresponds to a much more electron-rich area (red), compared to $NHCH_2$ side (blue). The proton transfer between the two molecules exacerbates the system intermolecular polarization, as E1 step

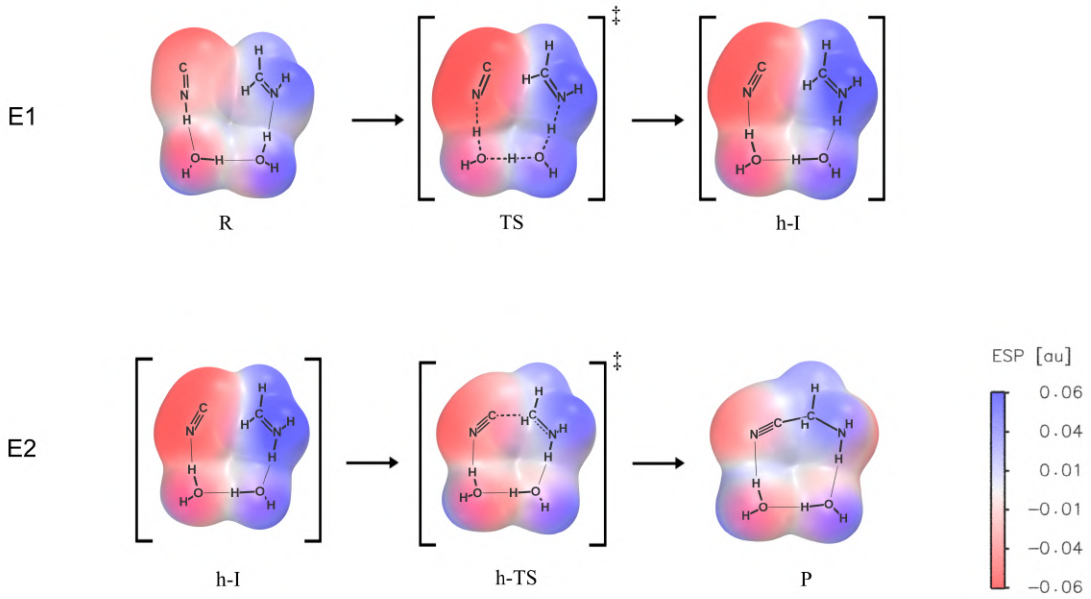


Figure 3.32: Schematic reaction mechanism for S3 – W_2, and MEPs of the minimum energy points. Blue corresponds to electron-poor regions while red to electron-rich ones. The structures have been extracted from the IRC plot. Hidden TS (h-TS) corresponds to a inflection points on the reaction force, while the hidden intermediate (h-I) has been identified among the points immediately preceding the h-TS region).

leads to a dipolar h-TS of zwitterionic nature, where most of the electron-rich regions are located on the deprotonated HNC. E2 is basically the reaction between two ionic fragments of opposite charge. Finally, the product is the structure where the charge is most delocalized, although the cyano group retains negative polarization in the molecule. It should be advised that part on the electron density is located on the restored lone pair on the amino group, although, it can only be partially noted, as the lone pair is pointing backward with respect to the figure orientation.

3.4.4 Stage 4: Glycine formation

The final stage is the hydrolysis of NH_2CH_2CN, which involves the nucleophilic attack (S4a) of a water molecule on the C-atom of NH_2CH_2CN, to yield a hydroxy imine, which tautomerizes (S4b) to the more stable glycine amide form ($NH_2CH_2C(O)NH_2$), followed by a second nucleophilic addition of water (S4c) to give glycine.

S4a has been studied on two different systems: $(NCCH_2NH_2 + H_2O + W_X)_{X=1,2}$, S4a$-W_{1-2}$, meaning that the hydrolysis of the nitrile is water-assisted and X is the number of water molecules engaged in the proton relay. It is to be noted that water molecules are labelled as 'H_2O' when they refer to reactants and as 'W' when they participate in proton transfer processes. The reaction is the nucleophylic addition of H_2O to the cyano group of $NCCH_2NH_2$ and it is concerted and exothermic for both systems. The energy diagrams are reported in Figure 3.33.

S4a$-$W$_1$ (black) is favored in terms of TS energy, since the barrier is lower than S4a$-$W$_2$ one (red) (38.1 vs 44.3 kcal mol^{-1}), as well as in terms of exothermicity. This result is surprising, since the TS geometry where there is a 1-water proton relay (S4a$-$W$_1$) entails a more strained bond angle with respect to the 2-water molecules proton relay (S4a$-$W$_2$). However, TS energy values for S4a are the highest encountered so far in the synthesis. In order to illustrate the several factors contributing to S4a energy barrier, and

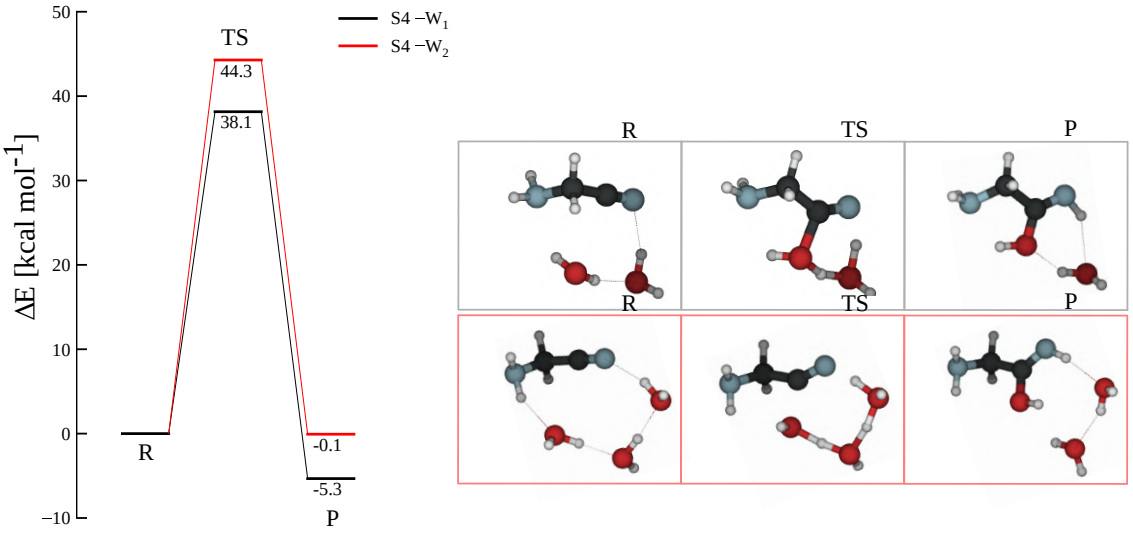
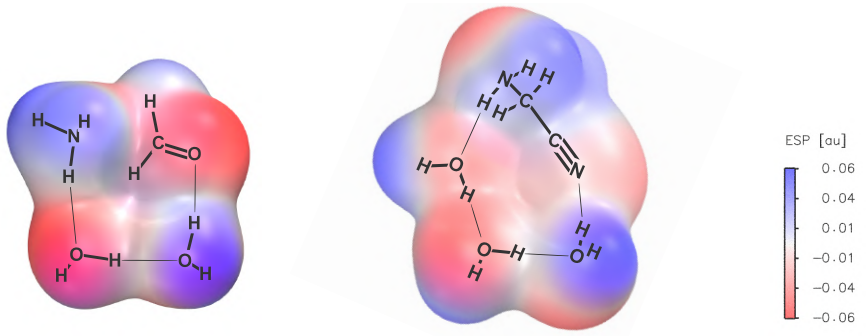


Figure 3.33: Energy diagram for S4a$-$W$_{1,2}$ structures, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. Minimum energy structures - reactants R, transition state (TS) and product (P) - are reported. The color scheme for the atoms is red for O, black for C, blue for N and white for H.

how the assisting water molecules affect it, a detailed comparison with S1 is illustrated, since both stages are nucleophilic additions. From a structural point of view, S4a is characterized by low accessibility of the electrophylic group, due to the fact that the electron-poor C-atom is engaged in a linear bond, therefore the reaction is sterically impeded (Figure 3.33R). Moreover, in order for S4a to take place, the aminoacetonitrile is required to break the linearity of its C-C-N angle, which implies large deformation of the molecule. On the other hand, the participants of S1 have smaller sizes and in the reactant equilibrium structure C-carbonyl is oriented toward NH_3, favoring the addition. In terms of electron affinity, H_2O is known to be a weaker nucleophile than NH_3, therefore its free lone pair is less easily donated to the elec-



S1 - W_2 (R) S4a - W_2 (R)

Figure 3.34: Comparison between MEPs of the reactants of S1 and S4a coordinated to a water dimer. Blue corresponds to electron-poor regions while red to electron-rich ones.

trophilic group. Lastly, from an electrostatic point of view, the cyano group in $NCCH_2NH_2$ is not as polarizable as the carbonyl group in H_2CO, besides, N-atom is a worse proton acceptor compared to O-atom. The comparison between the MEPs of the reactant is S4a-2W and in S1-1A (Figure 3.34) gives a visual measure of how less charge polarization is present in S4a$-W_2$.

Due to all these factors, the water-assisted hydrolysis of $NCCH_2NH_2$ appears to be a very unfavorable process.

Stage 4a: Reaction mechanism

Figure 3.35 reports the IRC profiles for S4a$-$W(1-2) systems. The very sharp shape of the IRC curve revealed a concerted mechanism where the events are almost syncronic.

Information about the order in proton transfer events is provided by the bond order derivative plot: O$-$H breaking in the nucleophilic H_2O takes place

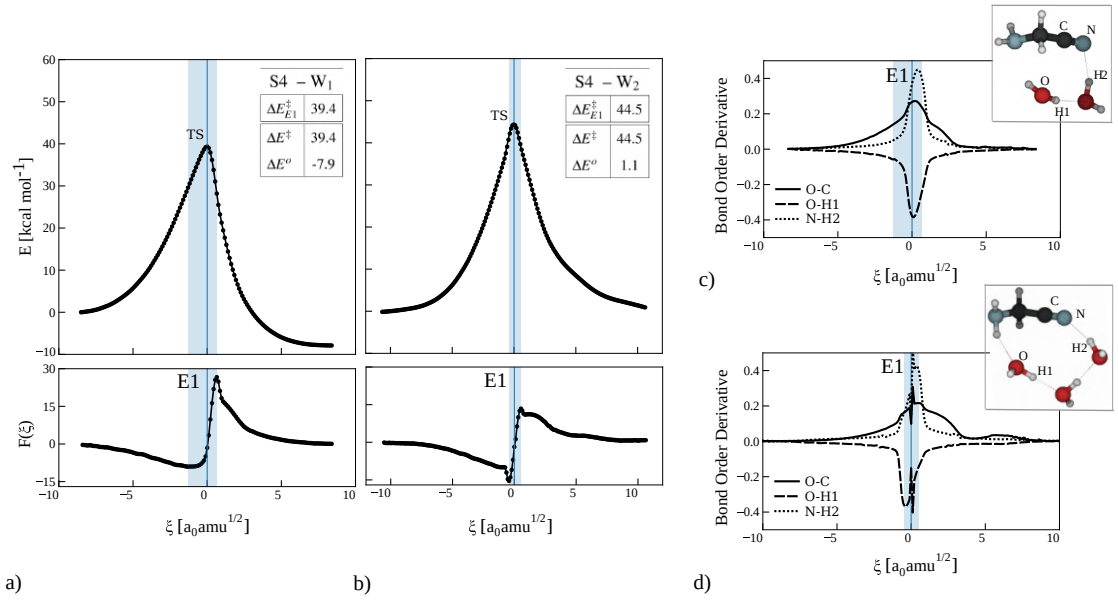


Figure 3.35: a,b) Energy (upper panel) and reaction force profiles (lower panel) for S4a$-$W$_{1,2}$, respectively, using BMK/def2-TZVP//BHANDHLYP-D3BJ/def2-SVP levels of theory. Blue line represent the TS. TS region is displayed as shadowed area. The plots include Tables that reports TS energy ($\Delta E^{\ddagger}$) and reaction energy (ΔE^{o}) as extracted from the IRC. c,d) Bond order derivative for the main bond distances involved in S4a, for the two systems.

slightly before nitrile's N$-$H formation. On the other hand, the addition of the hydroxide group is the event that requires the greates structural rearrangement and displays a broad peak (relative to variations in C$-$O) centered at the TS for both systems. Indeed, C$-$O formation starts before the beginning of the TS region and reaches its final value much later on the reaction coordinate, especially for S4a$-$W$_2$.

Stage 4: complete pathway

For completeness, the entire pathway of S4 has been studied using an intermediate size system constituted by the reactants adsorbed on a water tetramer:

S4 – W_4. The energy diagram is reported in Figure 3.36.

Table 3.9: Energy barriers ($\Delta E_{En}{}^{\ddagger}$) and reaction energies ($\Delta E_{En}{}^{o}$) for the system S4 – W_4, computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries, for step n of S4. Values in kcal mol^{-1}.

System	$\Delta E^{\ddagger}_{E1}$	ΔE^{o}_{E1}	$\Delta E^{\ddagger}_{E2}$	ΔE^{o}_{E2}	$\Delta E^{\ddagger}_{E3}$	ΔE^{o}_{E3}
S4 – W_4	37.9	-8.2	4.9	-10.9	39.5	6.6

The reaction is the conversion of a nitrile (aminoacetonitrile) into a carboxylic acid (glycine). It comprises three steps, the energy barrier for each of them is reported in Table 3.9. The first step (E1) is the already explored nucleophilic addition of water to the polar C – N bond to give the hydroxy imine (I_1). In this system, the proton tranfer is mediated by 2 water molecules. The energy barrier for the process is line with the findings for the smaller clusters. Tautomerization of the intermediate (E2) yields the glycine amide (I_2), in a reaction analogous to the tautomerization of a enol to give a ketone, once again mediated by 2 water molecules. The step requires to overcome a barrier of $\Delta E^{\ddagger}_{E2}$ = 4.9 kcal mol^{-1}. Further hydrolysis of the amide gives the product, glycine, plus ammonia, with a barrier of 39.5 kcal mol^{-1}. The step involves the nucleophilic addiction of water to the C-carbonyl followed by transfer of a proton from the water to N-atom of NH2 moiety, to make the latter a better leaving group, and subsequent elimination. The proton transfer of step E3 is assisted by a single water molecule.

Apart from the tautomerization (E2), the steps E1 and E3, requires them-

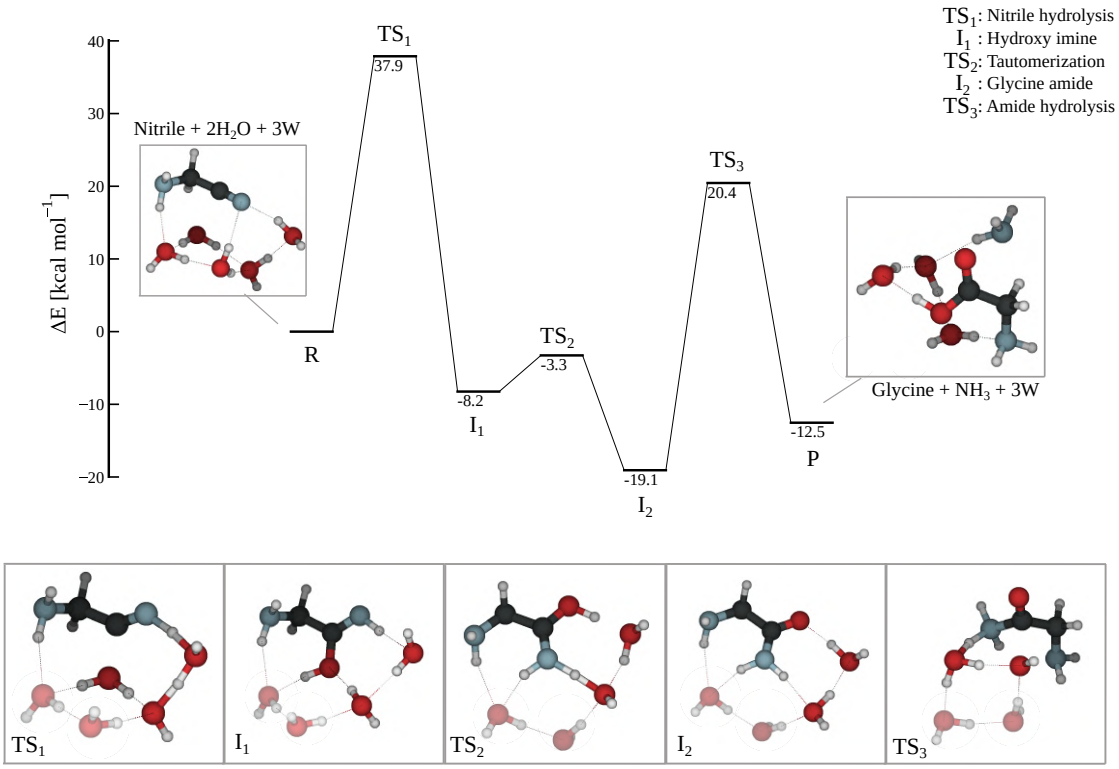


Figure 3.36: Energy diagram for S4 – W_4 structures, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. Minimum energy structures - reactants R, transition states ($TS_{1,3}$) and product (P) and intermedates ($E_{1,2}$) - are reported

selves a amount of energy superior to any other stage of the Strecker synthesis of glycine, suggesting that the feasibility of the S4 in conditions of low temperature is limited.

3.5 Strecker-type reactions on ASW

ASW has no long-range order resulting in greater surface heterogeneity than crystalline-ice, therefore a large variety of adsorption sites exists on the surface. In Section 3.3, the implications of it have been illustrated on the estimation of the binding energies of physisorbed species. The following section addresses how it affects the energetic parameters of reactive processes, such as reaction energies (ΔE^o) and transition state (TS) barriers ($\Delta E^{\ddagger}$).

A mean to account for a most complete census of reactive configurations is selecting a set of suitable sites where the reactants are adsorbed (tightly or loosely), according to their binding sites maps. As it was the case for the binding energy, the TS and reaction energy of processes taking place on ASW are better represented by a distribution of values, where the lower extreme of the range coincides with the most energetically favorable reactive arrangement on the ice.

First and second order surface effects

Different reaction site configurations can be described in terms of the non-covalent intermolecular interactions that the reactants establish with the ice - in form of hydrogen bonds (HB) - and the strength and inductive effects of such coordination. Those effects can be considered as *surface effect of first order*, since they involve the reactants. Besides, it is also important to analyse

a) S1-C3-A b) S1-C3-B

Figure 3.37: Example of coordination anomaly on the bridging water (W−B) on two reactants structures for the first stage of the synthesis. a) structure where W_B presents 2 HB donor groups (D). b) structure where W_B has only one D group. HB acceptor groups are marked as A.

the coordination of the assisting water molecules with respect to the ice HB network. Strecker-type reactions on ASW involve a sequence of proton transfer steps. These processes occur through bridging molecules that belong to an extended HB network, compared to the model system. Thus, it is expected for the surrounding water molecules - not involved in the reactions - to have a specific catalytic effect on each of the proton transfer steps, according to their role in the HB network with respect to the bridging molecules. This kind of surface effect can be considered as *second order*, as it refers to the effects of the ice on the bridging water molecules, instead of the reactants. As an example, Figure 3.37 reports two initial structures for the first stage. Although from a "reactants point of view", structures Figure 3.37a and Figure 3.37b are analogous (H_2CO is establishing two HBs with the surface, while NH_3 only one), in terms of second order effects, they differ considerably. While in

Figure 3.37a the bridging water molecule that will be engaged in the proton transfer (marked as W_B), presents two HB donor groups, in Figure 3.37b, W_B lacks of a HB donor water molecule. This kind of HB network phenomena can have a dramatic effect on the energy barriers of the reactive processes.

The synthesis of glycine is comprised of 4 stages, each one has been studied on a set of small water clusters (Sections 3.4.1, 3.4.2, 3.4.3, 3.4.4). In the following, it will be explored on a series of reactive sites on the ASW model composed of 22-water molecules (Sections 3.5.1, 3.5.2, 3.5.3, 3.5.4). Additionally, the first stage is also studied on a nano-porous model (Section 3.5.1.1) and on the ice mixture model (Section 3.5.1.2).

3.5.1 Stage 1: Aminomethanol formation

General aspects

Studying the reaction on a larger surface model implies dealing with the HB network presents within the ice and the potentially higher number of participating water molecules, giving rise to more complex phenomena. In agreement with the *multi-binding* framework, S1 has been carried out on a variety of sites on the ASW surface. The procedure to obtain the set of transition states has been the following:

1. Selection of three of the ASW_{22} cluster models modelled in Section 3.1.1, and identification of all the sites where a dangling hydrogen atom

('dangling-H') is present. Those are considered to be potential reactive sites, due to the fact that a dangling-H is necessary for the docking of H_2CO: bridging the proton from NH_3 side across the ice to H_2CO. The spreading of the dangling-Hs on a selected surface can be appreciated by the MEP image in Figure 3.38.

2. Extensive sampling of the reactants in the minimum energy configurations corresponding to $S1-W_{1-2}$ fragments, i.e. extracting the geometry of the pre-reactive complex, ($H_3N \cdots C(H_2)=O$) and removing the assisting water molecules. The step provides a set of initial configurations to use as starting point for the TS search.

3. TS characterization, which has been successful for a number of 14 reaction sites.

4. Identification of reactants and products and TS energy calculation.

The set of transition states has been categorized according to the characteristics of their corresponding reactive sites.

Reactive sites categorization

The reactive sites have been characterized in terms of how strongly the reactants are affected by the presence of the surface. As pointed out for the model system, the reactants interact *via* HB with the water (W) molecules of the ice. S1 systems feature primarily three possible types of HB interactions,

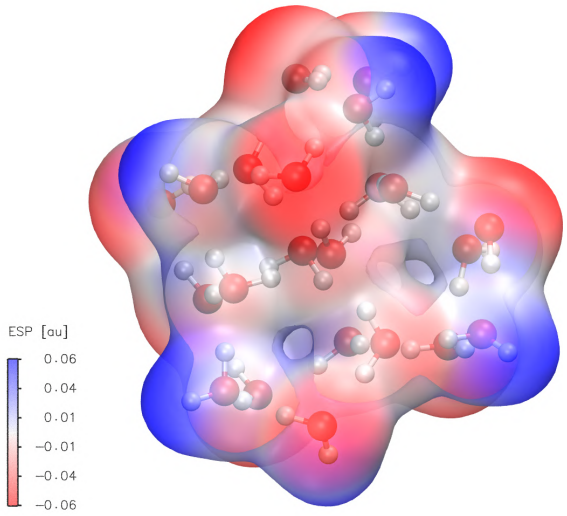


Figure 3.38: Map of electrostatic potential (MEP) calculated for of one of the ASW clusters. Blue corresponds to electron-poor regions while red to electron-rich ones. There is an average of 8 dangling H sites per cluster.

which have been described in Section 3.3.4.1 (namely T1, T2, T3). The total number of HBs, N_{HB} established by pre-reactive complex, can be used as a measure of its insertion into the surface HB network. Larger N_{HB} corresponds to higher insertion. The set of reactants has been divided into 4 cases (C), based on N_{HB} and HB type:

- Case 1 (S1-C1): $N_{HB} = 2$
- Case 2 (S1-C2): $N_{HB} = 3$
- Case 3 (S1-C3): $N_{HB} = 3$
- Case 4 (S1-C4): $N_{HB} = 4\text{-}5$

Moreover, to further quantify the strength of the interaction, the reactant interaction energy - *IE(R)* - between the pre-reactive complex and the ice is

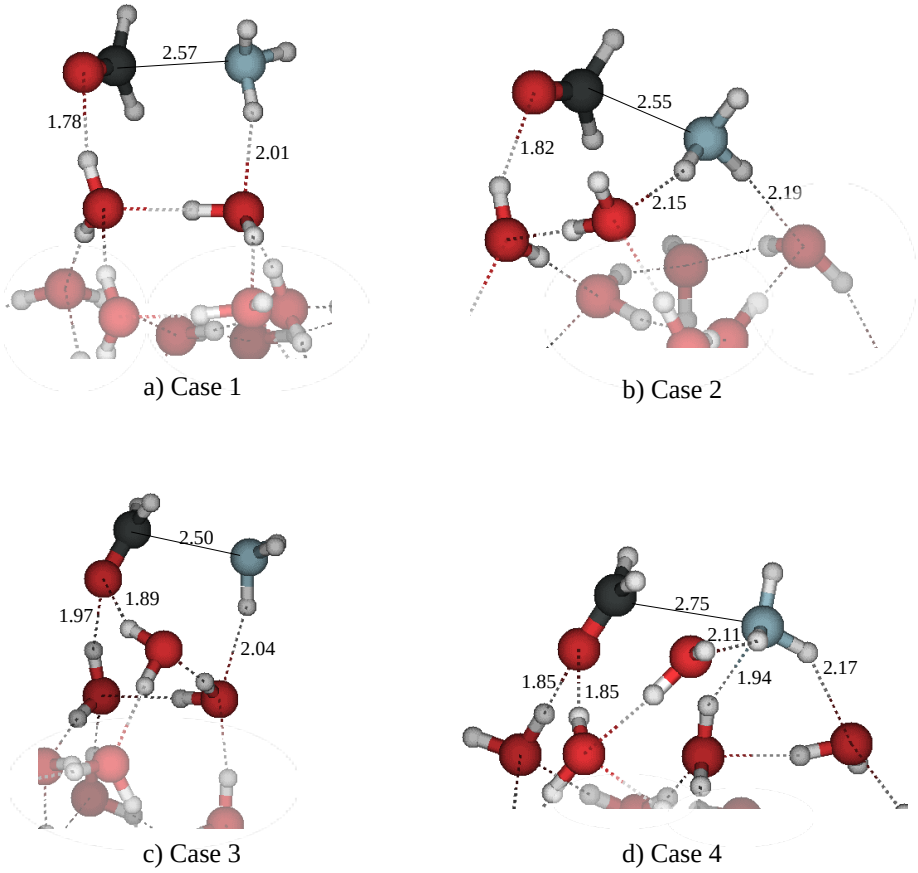


Figure 3.39: Example of structures belonging to S1-C(1-4) groups identified for S1. strucures names are illustrated in Table 3.10. a) S1-C1-B, b) S1-C2-A, c) S1-C3-C d) S1-C4-B.

estimated through a SAPT analysis. The interacting fragments have been defined according to the following expression:

$$IE(R) = -IE(\mathrm{NH_3} + \mathrm{H_2CO} \cdots \mathrm{W_{22}}) \tag{3.8}$$

Table 3.10 reports S1 cases and their characteristics, while Figure 3.39 reports an example structure for each case.

The first case, S1-C1, represents a situation where the effect of the ice on

Table 3.10: Categorization of the reactive sites for Stage 1 (S1) in 4 cases (S1−C(1-4)). First column reports the structures belonging to each case, labelled alphabetically, e.g. S1-C1-(A-E). The categorization is based on: total number of HBs established by the reactants with the ice (N_{HB}, second column), HB type (T_{1-3}, third column), as illustrated in Section 3.3.4.1, and their interaction energy with the surface ($IE(R)$, forth column), as defined in Equation 3.8. The average (Avg) and standard deviation (Std) of $IE(R)$ is reported as well. Values in kcal mol^{-1}.

Case	N_{HB}	$IE(R)$
S1-C1: *mild effect*	2: T2,T3	
S1-C1-A		8.0
S1-C1-B		11.7
S1-C1-C		10.5
S1-C1-D		11.3
S1-C1-E		11.6
	Avg(Std)	10.6 (1.4)
S1-C2: *strong effect on NH_3 side*	3: T2 × 2, T3	
S1-C2-A		13.7
S1-C2-B		11.3
S1-C2-C		12.2
	Avg(Std)	12.4(1.0)
S1-C3: *strong effect on H_2CO side*	3: T2, T3 × 2	
S1-C3-A		16.6
S1-C3-B		18.9
S1-C3-C		17.6
	Avg(Std)	17.7(0.9)
S1-C4: *strong effect on both sides*		
S1-C4-A	4: T2 × 2, T3 ×2	21.7
S1-C4-B	5: T1, T2 × 2, T3 ×2	21.8
S1-C4-C	4: T2 × 2, T3 ×2	30.6
	Avg(Std)	24.7(4.1)

the reaction site resembles the model system: both the reactants establish one HB each with the surface, as they interact solely with the water molecules

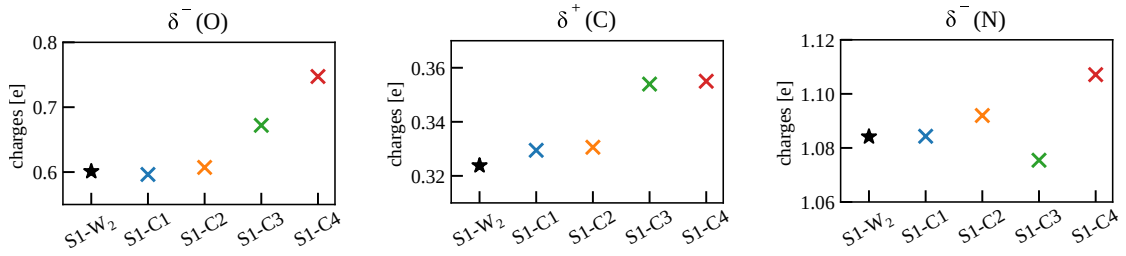


Figure 3.40: Average NAO partial charges estimated on carbonyl-O (left), carbonyl-C (middle), N (right), for the structures of each cases. Color code is blue for S1-C1, yellow for S1-C2, green for S1-C3 and red: S1-C4. Analogous partial charges on the model system atoms are reported for reference (black star).

that will assist the proton transfer: N_{HB} is 2. S1-C1 structures presents the smallest average IE value (10.6 kcal mol^{-1}). Therefore the surface effect on the pre-reactive complex is defined as *mild*.

S1-C2 and S1-C3 cases exhibit structures in which, in addition to the assisting waters, there is another water molecule directly HB-bonded to the reactants. This additional water molecule is specifically coordinated to NH_3, serving as a secondary HB acceptor in the case of S1-C2, or to H_2CO, acting as a secondary HB donor in the case of S1-C3. Therefore, N_{HB} is 3 for both groups and the surface effect is denominated *strong on one side* of the pre-reactive complex. The second water molecule is expected to enhance the inductive effects exert on the reactants by the assisting water molecules, as discussed in Section 3.4.1. The additional HB interactions with the surface are also congruent with the higher $IE(R)$ values in S1-C2 and S1-C3 compared to S1-C1 sites.

Figure 3.40 reports the change in average NAO partial charges for each

case with respect to the model system. Substantially, S1-C1 (blue) behaves as the model system (black star). The presence of a third HB established by NH_3 (S1-C2, yellow), results in the increase of the magnitude of the charge on N-atom. When a third HB is established by H_2CO (S1-C3, green), it results in the increase of the magnitude of C and O charges.

The last case, S1-C4, is constituted by structures which present the combination of the characteristics of the previous two cases, reaching $N_{HB} = 4$ (two HB established by each reactant). There is also a very peculiar case that has been included in S1-C4, where the total N_{HB} is 5: 3 HBs are established on NH_3 side (due to the presence of 2 HB acceptor water molecules and 1 HB donor) and 2 on the H_2CO side. The level of insertion of S1-C4 structures into the ice network is consequently higher, providing *strong surface effect on both sides* of the pre-reacting complex, as indicated also by the $IE(R)$ values for this group, that are far greater in magnitude compared the previous cases, with an average of 24.7 kcal mol^{-1}. S1-C4 structures also present the highest intramolecular polarization on both the reactants, as confirmed by the partial atomic charges in Figure 3.40 (red) that show the largest values.

A careful analysis of the reactive sites shows that structures more inserted in the HB network of the ice (S1-C3 and S1-C4) display the presence of a second dangling-H in the reaction environment. Such dangling-H is responsible of the peculiar coordination of H_2CO, corresponding to BM1a binding mode (an example is displayed in Figure 3.15), which is the highest in BE and also

the rarest, since it requires the presence of two dangling-Hs close to each other in the ice model. It will be illustrated that the reaction mechanisms is highly dependent on such environmental characteristic, therefore, the reaction sites are primarily grouped in situations where there is only one dangling-H (*lone dangling-H sites*, S1-C1, S1-C2), and situations where there are two (*pair dangling-H sites*, S1-C3, S1-C4).

In the following sections, detailed results for the different situation are presented.

Lone dangling-H sites

The reactions in correspondence of lone dangling-H sites includes S1-C1 and S1-C2 structures. Their reactive behaviors exhibit slight differences. Consequently, the results will be presented in terms of energy, followed by a detailed examination of the reaction mechanisms for each of them, separately.

5 structures correspond to case S1-C1: one of them reflects the mechanism of S1-W_1, i.e. the proton transfer is mediated by one water molecule, while the rest reflects S1$-W_2$. Table 3.11 and Figure 3.43, left panel, reports the energy barriers for this stage. The TS barrier for the structure S1-W_1-like is 13.7 kcal mol^{-1}. This value is only marginally smaller than S1-W_1 (14.6 kcal mol^{-1}), meaning that the reaction site is not particularly affected by the ice environment. Structures that present S1$-W_2$-like behavior have an average TS barrier of 10.3 kcal mol^{-1}, in line with S1$-W_2$ model system. The standard deviation of the average TS barrier is very low, suggesting little structural variability in the reactive sites for this case.

There are 3 structures that correspond to S1-C2, TS energy barrier and reaction energy results are reported in Figure 3.42 and Table 3.11. The mechanism is analog to the reaction assisted by 2 water molecules and the energy values are in line with S1$-W_2$, with a average barrier of 10.0 kcal mol^{-1}.

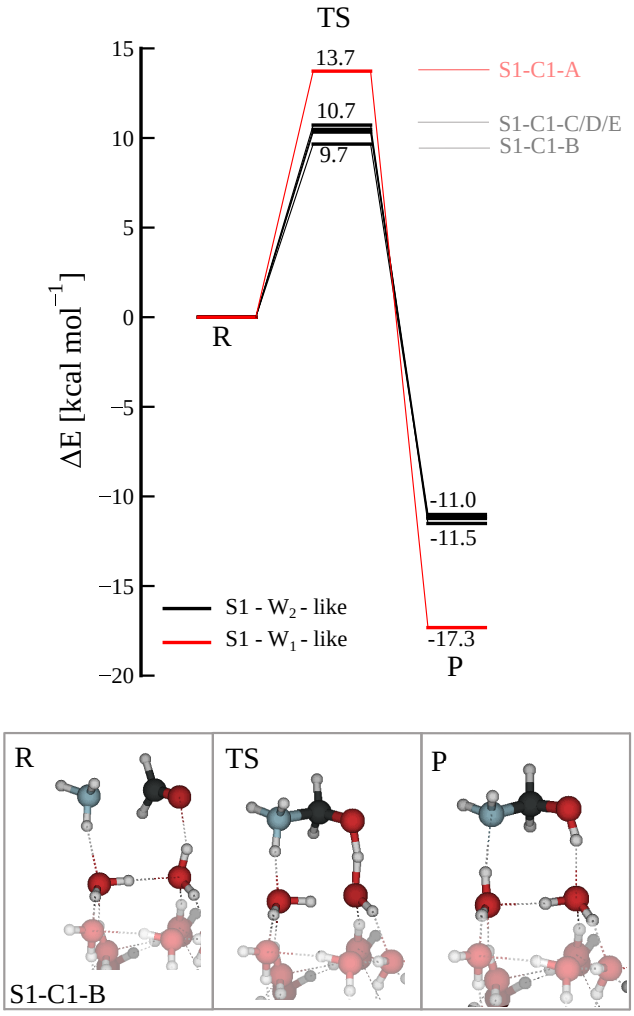


Figure 3.41: Energy diagram for S1-C1 structures, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. Minimum energy structures - reactants R, transition state (TS) and product (P) - are reported for a example case. Energy diagram for S1-C2 structures, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. Solid black lines represents structures where the proton transfer is mediated by 2 water molecules (as in S1$-$W$_2$, hence the name 'S1$-$W$_1$-like'). Solid red line indicate one water-molecule assisted structures (S1$-$W$_1$-like). Minimum energy structures - reactants R, transition state (TS) and product (P) - are reported for a example case. The color scheme for the atoms is red for O, black for C, blue for N and white for H.

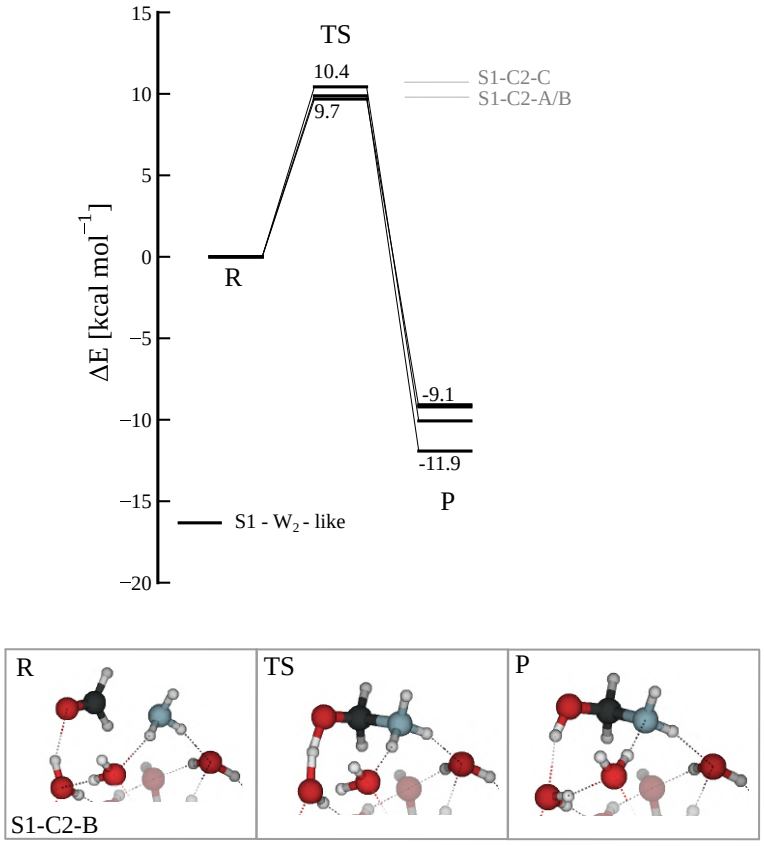


Figure 3.42: Energy diagram for S1-C2 structures, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. See Figure 3.41 caption for more details.

S1-C1: mild surface effect. Reaction mechanism

IRC analysis has been carried out for S1-C1-B, the system that presents the lowest TS barrier, the energy profile is reported in the upper panel of Figure 3.43a. The plot is qualitatively similar to the findings for S1 – W_2. Analysis of the reaction force profile (Figure 3.43a, lower panel) allowed to identify the reactive events, and to partition the energy barrier according to in Equation 3.7. As it was the case in S1 – W_2, the first reactive event associated to

Table 3.11: Energy barriers ($\Delta E^{\ddagger}$) and reaction energies (ΔE^{o}) computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries, for structures that belong to S1-C1, and S1-C2 cases. Column two, W_X, indicates the number of water molecules involved in the proton relay. Average (Avg) and standard deviation (Std) for the set of $S1-W_2$-like cases have been included as well. Values in kcal mol^{-1}.

System	W_X	$\Delta E^{\ddagger}$	ΔE^{o}
S1-C1-A	1	13.7	-17.3
S1-C1-B	2	9.7	-11.2
S1-C1-C	2	10.3	-11.2
S1-C1-D	2	10.5	-11.0
S1-C1-E	2	10.7	-11.5
Avg(Std)		10.3(0.4)	-11.2(0.2)
S1-C2-A	2	9.7	-9.2
S1-C2-B	2	9.9	-11.9
S1-C2-C	2	10.4	-10.1
Avg (Std)		10.0(0.3)	-10.4(1.1)

the h-TS corresponds to the C−N formation while the second reactive event is the proton transfer process. The $\Delta E^{\ddagger}_{E1}$, associated to the h-TS, is of 5.3 kcal mol^{-1}, slightly larger than in the model system (4.3 kcal mol^{-1}). Considering that the only water molecules coordinated to the reactants are the ones involved in the proton relay, as in $S1-W_2$, the effect on the barrier is analogous. However, the rigidity of the HB ice network takes part in increasing $\Delta E^{\ddagger}_{E1}$. Regarding $\Delta E^{\ddagger}_{E2}$, the barrier to convert the dipolar hidden intermediate ($^{-}OCH_2\cdots NH_3^{+}$) to the TS, it is of 3.4 kcal mol^{-1}, slightly smaller than in the model system, meaning that the proton transfer processes are marginally facilitated by the environment.

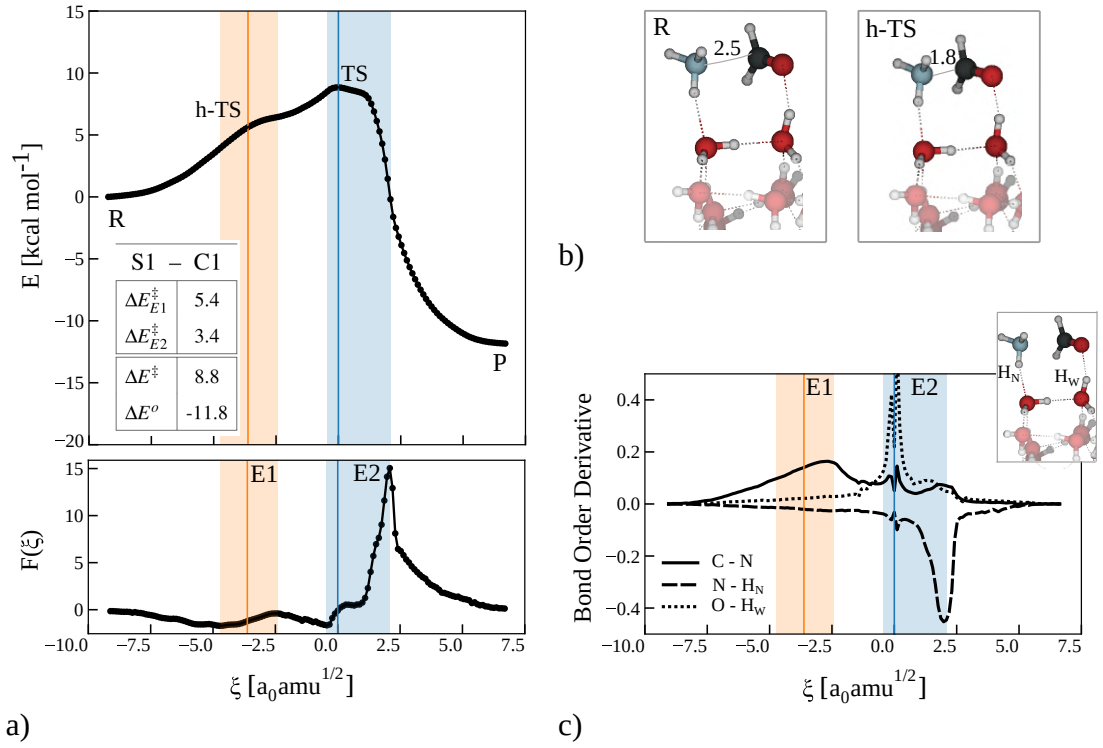


Figure 3.43: a) Energy (upper panels) and reaction force profiles (lower panels) along the IRC for S1-C1-B, using BMK/def2-TZVP // BHANDHLYP-D3BJ/def2-SVP levels of theory. Blue and orange lines represent TS and h-TS, respectively. TS and h-TS regions are displayed as orange and blue shadowed areas. The table reports TS energy ($\Delta E^{\ddagger}$) and reaction energy (ΔE^{o}) extracted from the energy profile, as well as the partition of the barrier in $\Delta E^{\ddagger}_{1}$ (relative to h-TS) and $\Delta E^{\ddagger}_{2}$ (TS). b) Structures relative to R and h-TS, as extracted from the energy profile. C-N bond distance have been highlighted. c) Bond order derivative for the main bond distances involved in S1.

Regarding the reaction mechanism, as shown by the bond order derivative analysis, Figure 3.43b, the order and location of the peaks corresponding to bond breaking/forming phenomena is akin to S1 – W_2. However, the TS region is broader, meaning that the asincronicity of the events in E2 is more pronounced: the legs of the proton transfer are more spaced on the reaction coordinates, with N – H breaking being further apart with respect to O – H

forming, compared to the model system. The reason for such behavior lies on minimal surface second order effects on the bridging water molecules, consistent with the diminished $\Delta E^{\ddagger}_{E2}$. In summary, S1-C1 group of structures presents a behaviour akin to the findings for the small system (average of 10.3 kcal mol^{-1} in TS barrier), showing no major catalytic effect of the surface HB network.

S1-C2: strong surface effect on NH_3 side. Reaction mechanism.

The most relevant consequence of the stronger interaction with the ice on NH_3 side of the reactive complex, is on the reaction mechanism and can be appreciated by taking into account the IRC profile for one of the structures (S1-C2-B), displayed in Figure 3.44a (upper panel).

The left branch of the curve is similar to the case analysed for S1-C1. The shoulder corresponding to E1 is present and slightly more pronounced than in S1-C1 case. The partition of the energy barrier, Figure 3.44b, shows that the energy associated to E1 is 5.3 kcal mol^{-1}, which is larger than in the model system (4.3 kcal mol^{-1}). As previously mentioned, this is probably attributed to the resistance encountered by the approaching NH_3 and H_2CO in the S1-C2 configuration, as they form additional interactions with the ice surface.

On the other hand, the TS region, and the region located passed the TS, differ from S1-C1: the curve presents the appearance of a second shoulder

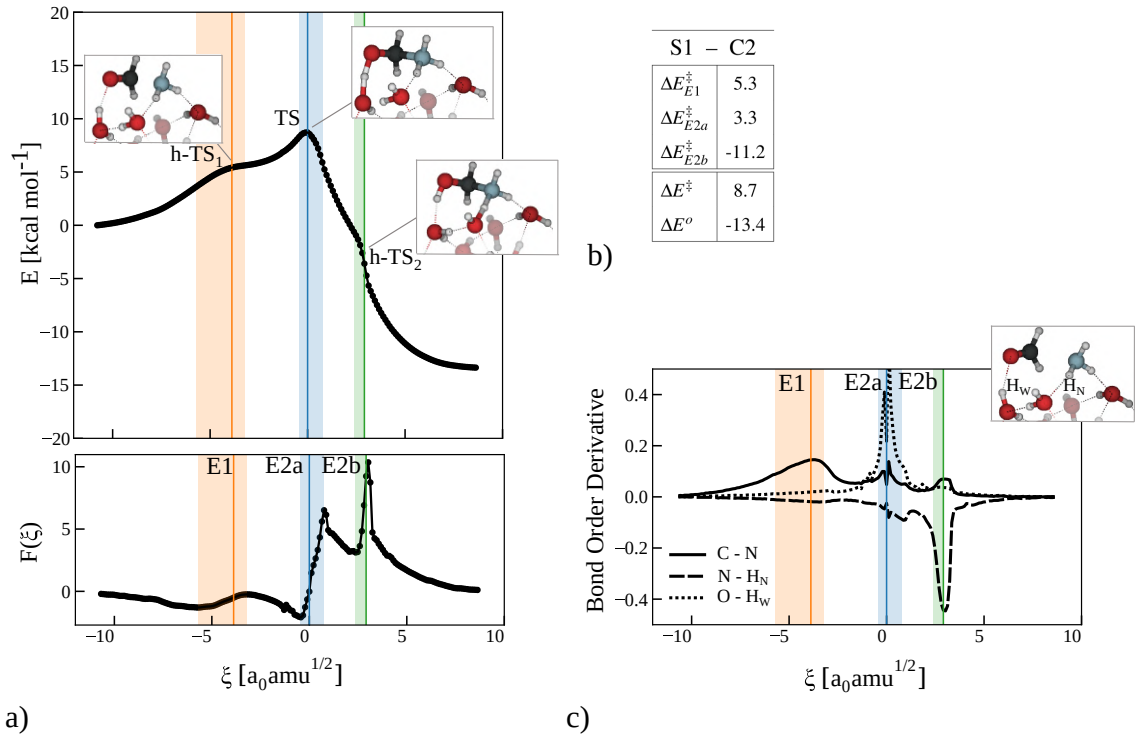


Figure 3.44: a) Energy (upper panel) and reaction force profiles (lower panel) for S1-C2-B, using BMK/def2-TZVP//BHANDHLYP-D3BJ/def2-SVP levels of theory. Blue line represent the TS. Orange and green represent h-TSs. TS and h-TSs regions are displayed as shadowed areas. TS and h-TSs structures have been reported as extracted from the energy profile. b) Table that reports TS energy ($\Delta E^{\ddagger}$) and reaction energy (ΔE^{o}) extracted from the energy profile, as well as the partition of the barrier in $\Delta E^{\ddagger}_{1}$ (relative to h-TS$_1$) and $\Delta E^{\ddagger}_{2a}$ (TS). $\Delta E^{\ddagger}_{2b}$ (h-TS$_2$) has been included for completeness. c) Bond order derivative for the main bond distances involved in S1.

feature, corresponding to a second hidden transition state, h-TS$_2$. The analysis of the inflection points of the reaction force profile, Figure 3.44, (lower panel), confirmed the existence of a third reactive event (green line and green shadowed region). The bond order derivative plot, Figure 3.44b shows that, in S1-C2, the asyncronicity in the proton transfer events increases to such extent that it manifests as two separated reaction events, in which the emerging h-TS can be associated solely to the final step of the proton transfer: the N – H

bond breaking. The result suggest that it is a peculiar feature of S1-C2 and that it follows from the inductive effects exert by the two water molecules acting as HB acceptors on NH_3. They diminish the proton donor character of N-atom, thus delaying the proton transfer from it. Basically the strong surface effect of first order on NH_3 side enhances the asincronicity of the proton transfer processes, therefore $N-H$ breaking is located much later with respect to $O-H$ bond forming (E2a), and constitutes a separate event (E2b).

The updated schematic mechanism for the reactive events of S1-C2 is reported in Figure 3.45, along with the MEPs of the reactive centre along the reaction coordinate, for the example case. As for the model system, Section 3.4.1, Figure 3.27, the structures have been extracted from the energy profile. All but the hidden intermediates (h-Is) correspond to critical points of the energy and the reaction force profiles. The h-Is tentative structures have been selected among the points immediately following the corresponding event region, in the flattest part of the IRC curve. E1 is unaltered with respect to the model system: it is the formation of the dipolar h-I_1, ($^{-}OCH_2\cdots NH_3^{+}$), structure that presents the largest intramolecular charge localization, as the formation of $C-N$ bond has a strong inductive effect on the carbonyl. The two proton transfer steps comprised in E2a lead to the protonation of the O-carbonyl, relocating the electron density from its oxygen's lone pair to the forming σ bond. The delay of the third proton transfer step results in the second hidden intermediate, h-I_2, which presents a relatively electron-rich

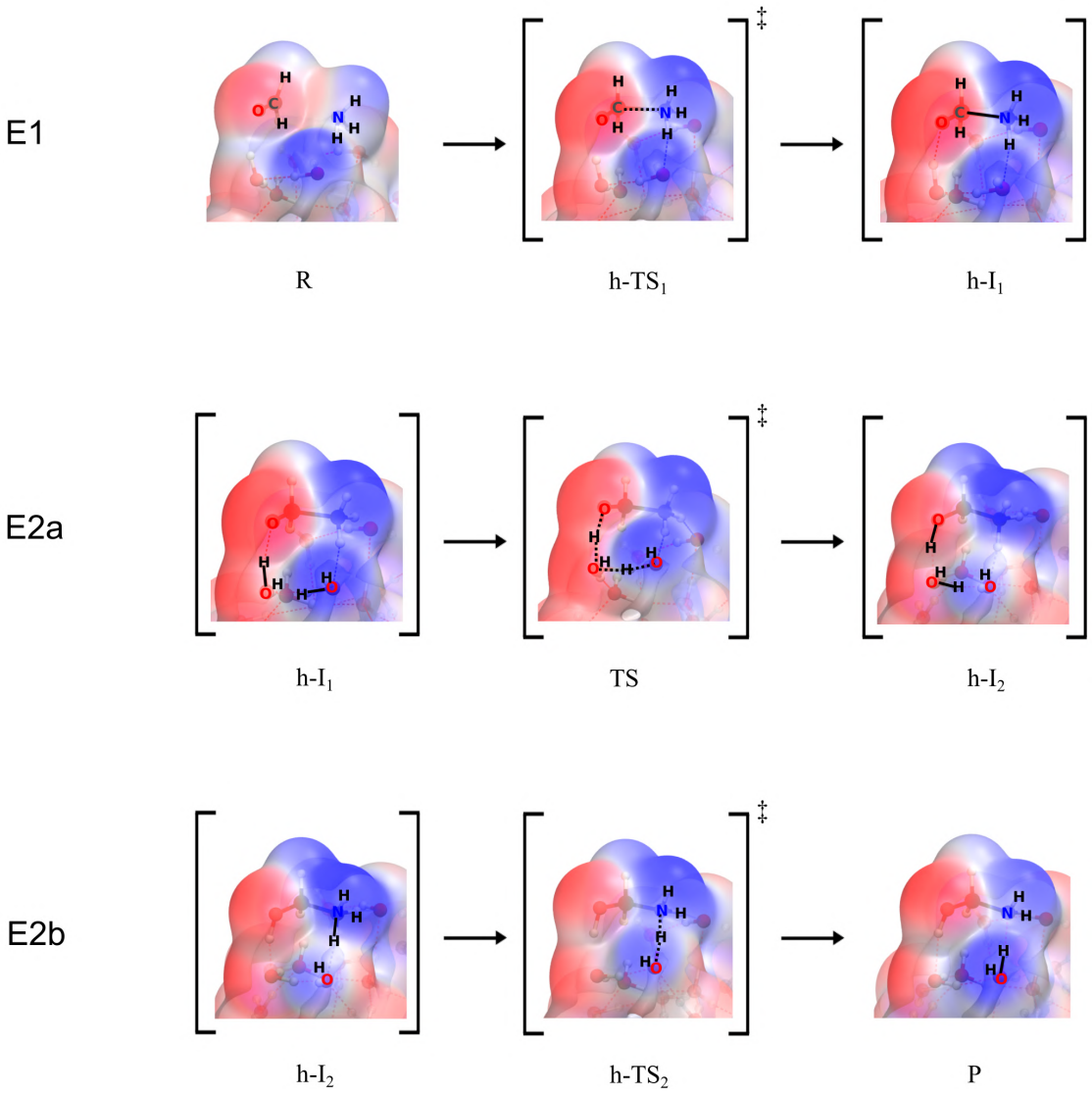


Figure 3.45: Schematic reaction mechanism for S1-C2-B, and MEPs of the minimum energy points. MEP color code: blue corresponds to electron-poor regions while red to electron-rich ones. The bonds involved in each reactive event have been highlighted. The structures have been extracted from the IRC plot. Hidden TS (h-TSs) correspond to inflection points on the reaction force while hidden intermediates (h-Is) have been identified among the point immediately following the event region's area.

area on the deprotonated water acting as HB acceptor on NH_3. It can be appreciated by noting the color change area on the bridging water to the left

of the reactive complex - despite the fact that the orientation of the molecule is not ideal, as the water lone pair is pointing backward - the area passes from vivid blue (h-I_1) to reddish (h-I_2). The third event (E2b) is N – H bond breaking that forms the product, relocating the electron-rich area from the bridging water to the N-atom, where the lone pair is reestablished.

The partition of the energy barrier, Figure 3.44b, shows that the reaction energy associated to E2a is 3.3 kcal mol^{-1}, which is smaller than the model system by almost 2 kcal mol^{-1}. As we pointed out for S1-C1 example case, it might be due to some surface second order effects on the reactive centre.

In summary, S1-C2 structures present an average TS barrier of 10.0 kcal mol^{-1}, hence no consistent surface catalytic effect compared to the model system. The major peculiarity of this group resides in the reaction mechanism that presents the appearing of a third event in the relaxation part of the reaction coordinate. The event is due to the delaying of the last step of the proton relay (N – H bond breaking), caused by the effect exerted by the two water molecules acting as HB acceptors on NH_3.

Pair dangling-H sites

S1-C3 and S1-C4 structures are characterized by the presence of a second water molecule coordinated to the reactants, acting as HB donor. That situation corresponds of having a pair of dangling-H atoms in the reaction site. The additional dangling-H is establishing a second HB donor on H_2CO side. As

a result of the extraordinary polarization of the carbonyl bond (Figure 3.40), once event E1 is reached, the hidden dipolar intermediate, ($^{-}OCH_2\cdots NH_3^{+}$), h-I, is converted into a real intermediate. Therefore, S1 reaction in S1-C3 and S1-C4 sites present a step-wise mechanism - constituted of 2 separated steps.

S1-C3: strong surface effect on H_2CO side

Figure 3.46 and Table 3.12 report the corresponding energy barriers for S1 steps, together with a example case. This group is composed of 4 structures.

Table 3.12: Energy barriers ($\Delta E_{En}{}^{\ddagger}$) and reaction energies ($\Delta E_{En}{}^{o}$) for structures that belong to S1-C3, computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries, for step n of S1. Column two, W_X, indicates the number of water molecules involved in the proton relay. Average (Avg) and standard deviation (Std) have been included as well. Values in kcal mol^{-1}.

System	W_X	$\Delta E_{E1}^{\ddagger}$	ΔE_{E1}^{o}	$\Delta E_{E2}^{\ddagger}$	ΔE_{E2}^{o}
S1-C3-A	2	2.9	2.8	5.0	-12.9
S1-C3-B	2	5.5	5.6	8.0	-11.3
S1-C3-C	1	3.6	4.0	4.1	-11.5
S1-C3-D	1	2.9	2.8	7.6	-11.7
Avg(Std)		3.6(1.1)	3.8(1.2)	6.2(1.6)	-11.9(0.6)

A minimum on the PES of S1-C3 emerges as the dipolar intermediate (I) corresponding to E1. Its appearance forges a PES plateaux, as the intermediate exhibits energy levels akin to those of TS_1, with an average barrier of 3.6 kcal mol^{-1}. This value is slightly lower than the findings for the partial barrier $\Delta E_{E1}^{\ddagger}$ for S1-C1 and S1-C2. E2, the exothermic formation of the product, has a average TS barrier of 6.2 kcal mol^{-1}. For all cases E2 barrier is greater

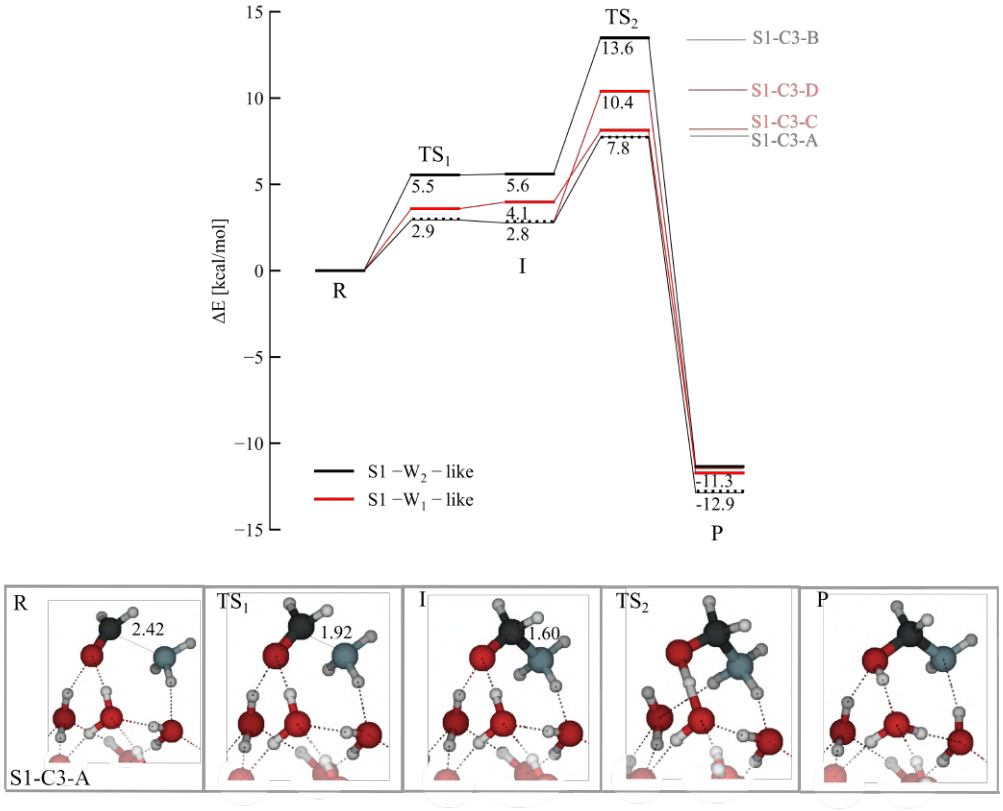


Figure 3.46: Energy diagram for S1-C3 structures, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. See Figure 3.41 caption for more details.

than E1, hinting that the second event is the rate-limiting step of the process. As previously mentioned (Section 3.5), event E2 might be affected by surface second order effects that alter energetic and order of the proton transfer steps in a specific fashion. Therefore, in order to explain the variability in E2 barriers for this group, the comparison between S1-C3-A ($\Delta E^{\ddagger}_{E2}$: 5.0 kcal mol^{-1}) and S1-C3-B ($\Delta E^{\ddagger}_{E2}$: 8.0 kcal mol^{-1}) is illustrated.

Figure 3.47a and 3.47c, show reactants and E2 transition states respectively. This couple of structures has been reported also in Figure 3.37 as

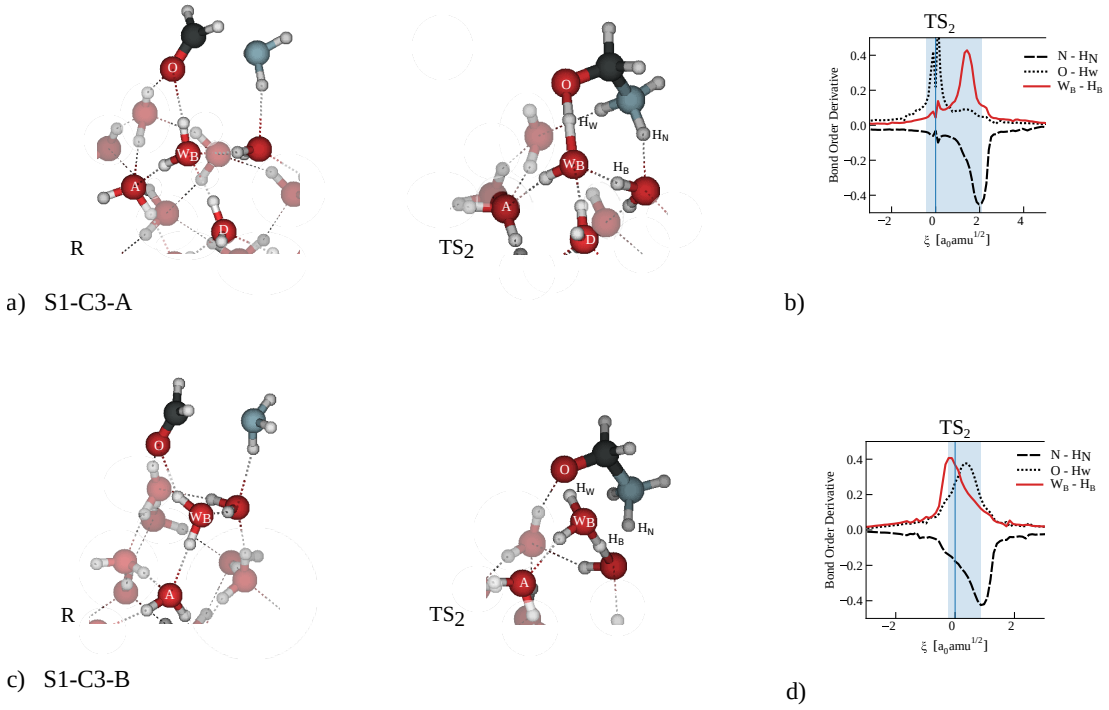


Figure 3.47: Upper panel: a) Geometries relative to S1-C3-A reactants (R, left) and second step TS (TS_2, right). HB donor and acceptor groups are labelled as A, D. W_B stands for bridging water. b) Bond order derivative for TS_2 events for the same system. Atoms involved in bond forming/breaking events have been labelled on TS_2 structure. Lower panel: c) and d) analogous of a) and b), respectively, for S1-C3-B system.

example case of coordination anomaly. In fact, S1-C3-B lacks of a HB donor group (labelled as D) acting on the bridging water (W_B), with respect to S1-C3-A. Thereby, the proton donor character of W_B in S1-C3-B is worsened, with consequent shifting in the order of E2 events. In fact, the protonation of the bridging water W_B, usually the intermediate step in a 2-water proton relay, takes place first with respect to $O-H$ bond formation. The result is confirmed by the bond order derivative analysis in Figure 3.47b and 3.47d. The red solid line representing the bridging water protonation (W_B-H_B bond

formation, in the figure), which in S1-C3-A is the second event to take place, is the first to happen in S1-C3-B, instead.

As a consequence, in terms of structural parameters and electrostatic polarization, TS_2 for the two systems, Figure 3.47a, 3.47c, differ considerably. In fact, the coordination defect in S1-C3-B alters its geometry, breaking the planarity of the 4 heavy atoms establishing the proton relay ($N-O_W-O_W$-O_{H2CO}). Moreover, the oxygen atom of W_B in S1-C3-B presents a lone pair not engaged in any HB, which results in a extraordinary electron-rich region, as it can be appreciated by the MEPs, reported in Figure E.2.

Similar order alteration in the proton transfer events has been observed for the pair of systems where the proton transfer is mediated by one water molecule ('S1-W_1-like' systems), S1-C3-(C,D), as well. Since for both $S1-W_{1-2}$-like cases the proton transfer reversed order mechanism results in higher E2 barriers, it can be concluded that coordination anomalies in the HB ice network, such as lacking HB donor groups acting on the assisting water molecules, is a energetically non favorable characteristic.

This is expected, since the HB network of the ice provides stabilization to the TS, hence coordination defects would lead to higher energy structures and higher TS barriers.

In summary, S1-C3 structures present a step-wise reaction mechanism, comprised of two step, corresponding to events E(1,2) previously identified. The first event, the formation of the dipolar intermediate is associated with

the formation of a PES energy plateaux with a average barrier of 3.6 kcal mol^{-1}, while the formation of the product is exothermic and it is the rate limiting step for all the systems, with an average energy barrier of 6.2 kcal mol^{-1}.

S1-C4: strong surface effects on both sides

Table 3.13: Energy barriers ($\Delta E_{En}{}^{\ddagger}$) and reaction energies ($\Delta E_{En}{}^{o}$) for step n of S1, for structures that belong to S1-C4, computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries. Column two, W_X, indicates the number of water molecules involved in the proton relay. Values in kcal mol^{-1}.

System	W_X	$\Delta E^{\ddagger}_{E1}$	ΔE^{o}_{E1}	$\Delta E^{\ddagger}_{E2a}$	ΔE^{o}_{E2a}	$\Delta E^{\ddagger}_{E2b}$	ΔE^{o}_{E2b}
S1-C4-A	1	0.4	-1.1	7.1	-13.0		
S1-C4-B	2	0.1	-10.7	12.8	-11.5		
S1-C4-C	2	-0.4	-2.1	4.9	1.8	2.6	-10.3

3 structures that correspond to S1-C4 characteristics: high level coordination of both the reactants with the ice surface. Among the reaction sites found for this case group, three involved two ice-surface water molecules in the proton relay (S1$-$W$_2$-like) while in one path the proton relay was established with one water molecule (S1$-$W$_1$-like). As for the previous case, the peculiar binding mode of H_2CO determines a step-wise mechanism comprised of several events. The first one is C$-$N formation, E1. The main difference with S1-C3 is that the dipolar intermediate is formed by a marked exothermic process. Moreover, the barriers $\Delta E^{\ddagger}_{E1}$ (reported in Table 3.13 and Figure 3.48) are close to zero kcal mol^{-1}, meaning that the PES is very flat

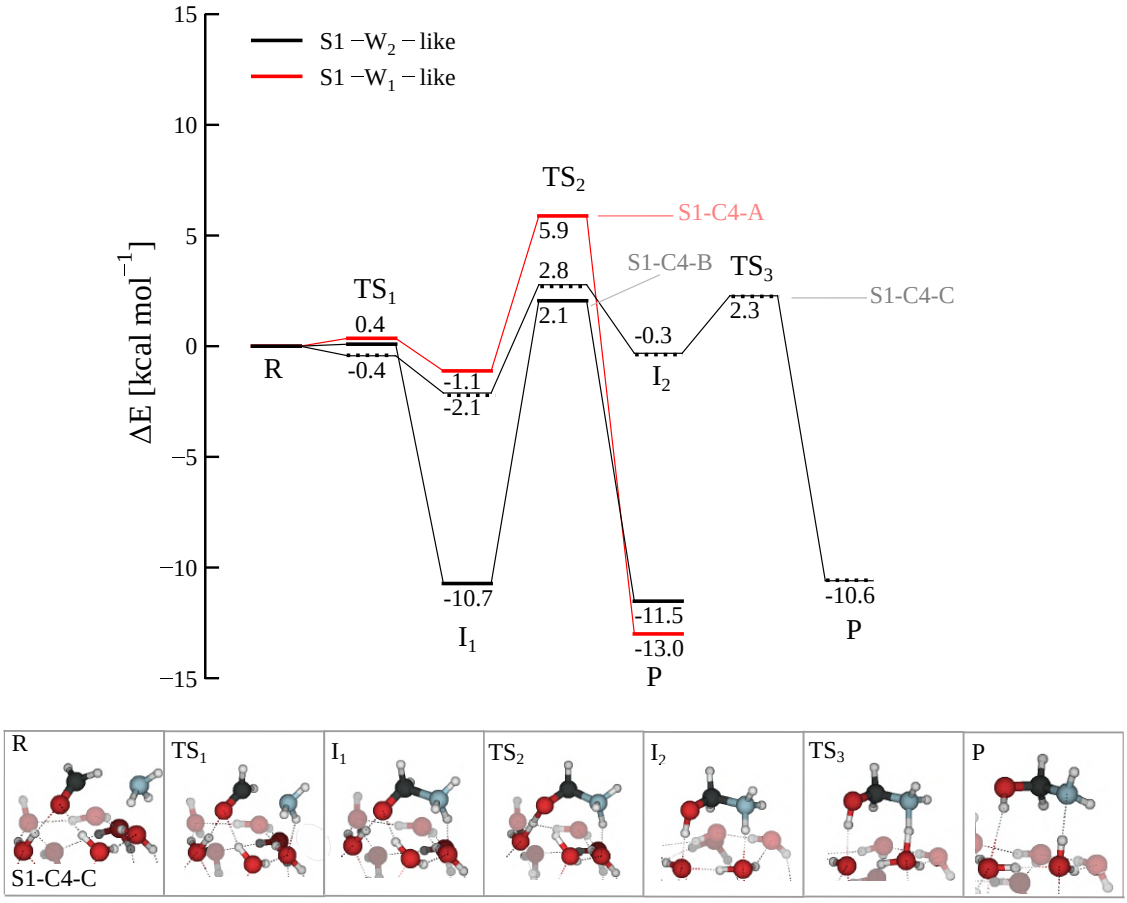


Figure 3.48: Energy diagram for S1-C4 structures, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. See Figure 3.41 caption for more details.

and the formation of the dipolar intermediate is essentially barrierless. The reaction energy of S1-C4-B, one of the reactive sites, is the most exothermic (ΔE^o_{E1}: -10.7 kcal mol^{-1}), due to the extraordinary coordination that the dipolar intermediate establishes with the ice, upon formation. In fact, the H_2CO side of the dipolar intermediate is interacting with three water molecules.

The energy barriers for the proton transfer, E2, range from 5 to 13 kcal mol^{-1},. E2 is the rate limiting step of the process, as it was the case also for

S1-C3. Despite the high E2 barriers, it is possible to argue that the exothermicty of the previous step might supply part of the amount of energy needed to overcome the barrier.

Lastly, structure S1-C4-C displays a different reaction pathway composed of three steps (E1, E2a and E2b). Figure 3.48 also reports the configuration of the different minima for this pathway. E2a, the first leg of the proton relay, is slightly endothermic and leads to the formation of a second dipolar intermediate (I_2). This structure is analogous to h-I_2 that was detected as a shoulder on the energy profile of S1-C2, Figure 3.44. Figure E.3, it is possible to find the MEP calculated for I_2. Analysis the HB network around of the assisting water W_B reveals a extraordinary hyper-coordinated molecule surrounded by three HB donor groups (labelled as D(1-3)). Hence, the successfully isolation of I_2 appears to be due to surface second order effects. S2-C4-C represents a 'paused' S1 mechanism, where the ice environment accommodates the reactive complex at each step. Nevertheless, the E2b, the final step, exhibits significantly greater exothermicity compared to the formation of I_2, making the production of the product more energetically favorable than halting S1 at the second dipolar intermediate. Furthermore, S1-C4-C presents among the smallest energy barriers for each step of S1 taken singularly, hinting that it might reflect the most plausible mechanism for the aminomethanol formation on ASW.

In summary, S1-C4 structures display elevated insertion into the HB net-

work of the ice, reflected by the high number of interactions established by the reactive complex. As a consequence, S1 proceeds through a step-wise mechanism, comprised of two or three exothermic steps, leading to the formation of hyper-coordinated dipolar intermediates that finally convert to the product.

3.5.1.1 Stage 1 inside of a nano-porous

To investigate the impact of the increased insertion of the reactive complex into the ice HB network, also motivated by the finding for S1 in proximity of the double dangling-H sites, S1 is conducted on a surface featuring a nano-porous. Only the inner region of the porous is taken onto account, since the structure has been extracted as a spherical portion of the ASW_{500} periodic surface.

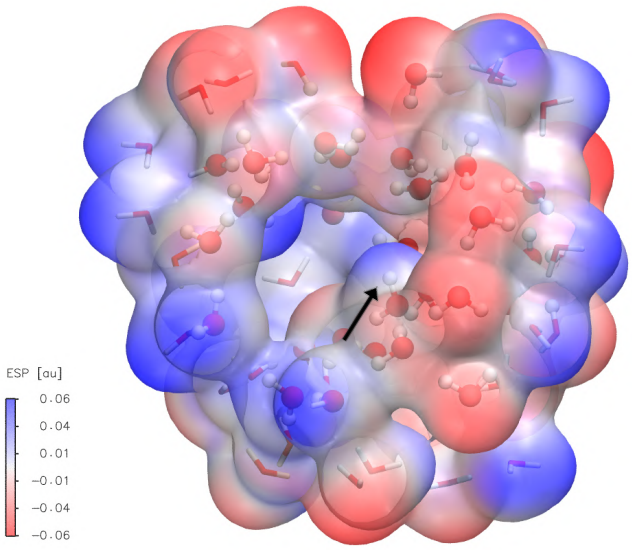


Figure 3.49: Map of electrostatic potential (MEP) calculated for the surface which present a nano-porous. Blue corresponds to electron-poor regions while red to electron-rich ones. The inner surface present two dangling Hs (blue regions); the one that will be used as a docking point for the reaction has been highlighted. A constraint has been applied on the atoms on the outer layer of the surface, represented as sticks.

A constraint has been applied on the atoms belonging to the outer sphere to preserve the porous. Further details about the modelling can be found in Section 3.1.5, and a MEP of the surface is displayed in Figure 3.49.

The inner surface presents only two dangling-H sites; the one on the right part of the nano-porous (highlighted) resulted to be the docking site for the reaction. To attain the initial geometry of the reactive site, the same procedure as in Section 3.5.1 was followed. However, all the sampling attempts using the reactive complex resulted in the barrierless formation of a dipolar intermediate ($^{-}OCH_2 \cdots NH_3^{+}$) I_1 (relative to reactive event E1). It was deduced that the porous environment inclined towards favoring that particular arrangement. Given that it was not possible to optimize both the reactants in-

side of the porous, in a non-interacting configuration, a different strategy has bee used, sampling the porous-inside using solely H_2CO and placing NH_3 on one of the edges. Such initial reactant geometry allowed to estimate the exothermicity of the intermediate formation. As expected, it was very favorable, entailing a energy difference of -16.9 kcal mol^{-1}, upon NH_3 diffusion into the cavity. Figure 3.50 reports the energy diagram for the pathway as well as the structures of the minima. The reaction mechanism is step-wise and substantially corresponds to the three-steps S1-C4 case, S1-C4-C (Section 3.5.1). It is not surprising since the nano-porous surface can be seen as a exacerbated situation of S1-C4, where both sides of the reactive complex are highly coordinated to the ice. The energy barriers for the three steps are reported in Table 3.14, S1-C4-C values are included for comparison.

Table 3.14: Energy barriers ($\Delta {E_{En}}^{\ddagger}$) and reaction energies ($\Delta {E_{En}}^{o}$) for step n of S1 carried out in a nano-porous, computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries. Column two, W_X, indicates the number of water molecules involved in the proton relay. S1-C4-C values are included for comparison. Values in kcal mol^{-1}.

System	W_X	$\Delta E^{\ddagger}_{E1}$	ΔE^{o}_{E1}	$\Delta E^{\ddagger}_{E2a}$	ΔE^{o}_{E2a}	$\Delta E^{\ddagger}_{E2b}$	ΔE^{o}_{E2b}
$S1-ASW_{porous}$	2	-	-16.9	4.9	0.2	4.1	-5.5
S1-C4-C	2	-0.4	-2.1	4.9	1.8	2.6	-10.3

As previously mentioned, I_1 formation (E1) is barrierless and much more exothermic when taking place into the nano-porous, than on the ASW surface, hinting to a important surface catalytic effect. Regarding E2a, the first

part of proton relay, the barrier is almost 5 kcal mol^{-1} and it is is slightly endothermic, as it was the case for S1-C4-C. On the other hand, the last event (E2b, the formation of the product) has a TS barrier of 4 kcal mol^{-1}, larger than S1-C4-C, hinting that the hyper-coordinated intermediate I_2 is more stabilized by the nano-porous than by the ASW cluster. In conclusion, the nano-porous presented a dangling-H acting as docking site located around the centre of the inner surface. This characteristic strongly favored the barrierless formation of the reactive complex between the reactants (I_1). The mechanism to convert the intermediate to the product proceeded *via* a two steps-proton relay with TS barriers of 4-5 kcal mol^{-1} each. Nevertheless, the considerable energy surplus derived from the highly exothermic first step (estimated by around ΔE^o_{E1} = -17 kcal mol^{-1}) might be used for the subsequent ones.

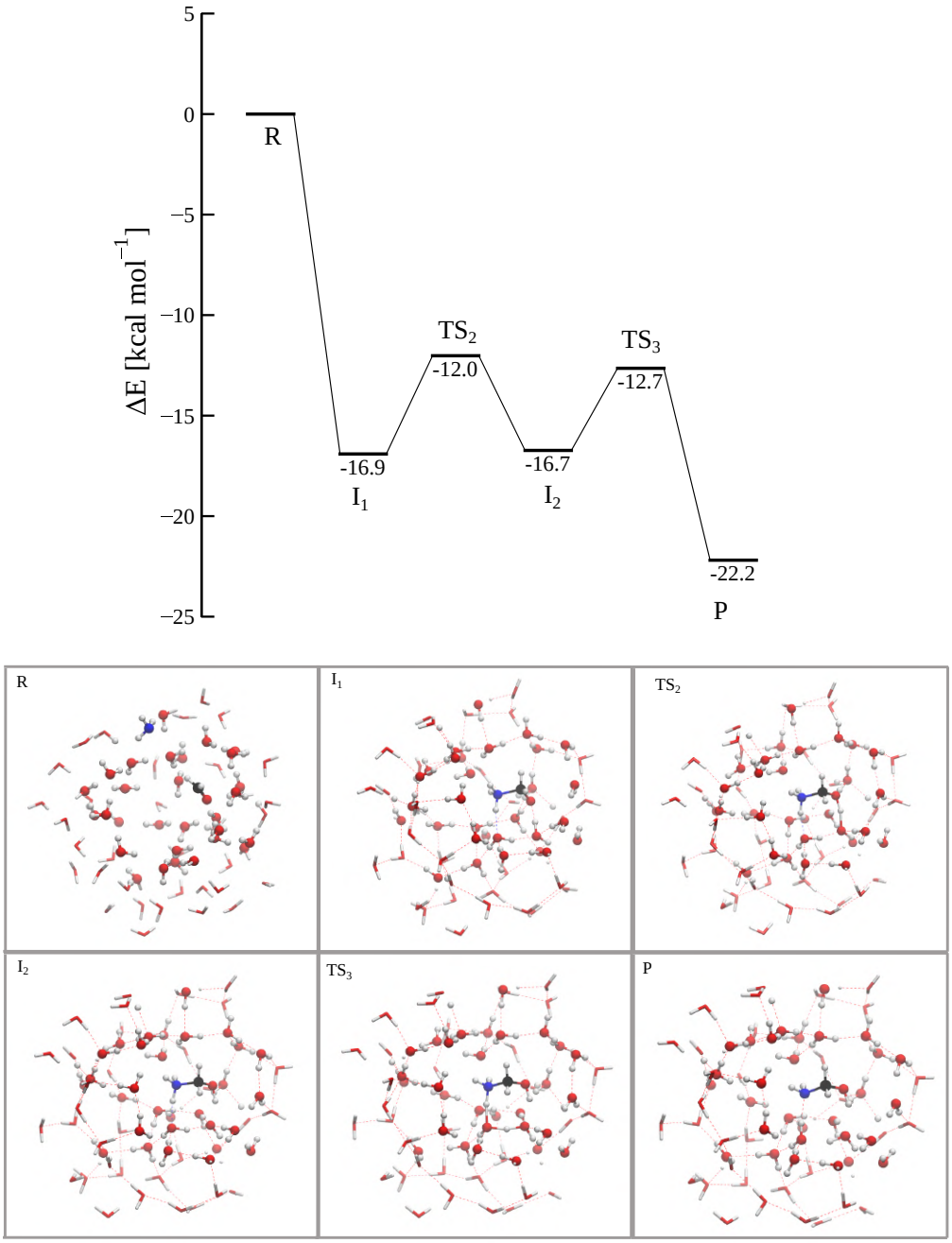


Figure 3.50: Energy diagram for S1 inside a nano-porous. Energies values computed at BMK/def2-TZVP // BHANDHLYP-D3BJ/def2-SVP level of theory. Energy minima are reported in the lower panel. Water ice molecules that are kept constrained during optimization are represented as sticks.

3.5.1.2 Stage 1 on ice mixture

The last section illustrates the results obtained for S1 using a ice mixture surface H_2O : NH_3 = 2.5:1. Only the first stage has been explored since it is the one where NH_3 participates in the reaction, therefore, increasing its availability in the environment might result in a higher probability of the process. However, the analysis has not statistical purpose, being focused mainly on evaluating eventual changes in energetics and mechanism of S1 reactions caused by the ammonia-rich environment.

The ice mixture surface presents the following composition: $(H_2O)_{15}$: $(NH_3)_6$, details about the modelling can be found in Section 3.1.3. Compared to the pure water clusters, the ice mixture presents more dangling-Hs, a total of 10 per surface and less spaced on the surface, as it can be evinced also from the MEP image in Figure 3.51. Due to the large variety of possible reactive combinations, only three limiting cases of such a broad scenario are taken into account:

- S1 – mixture – C1: surface's NH_3 molecules are present in the *bulk* of the ice, but do not interact with the reactants or take part of the reaction
- S1 – mixture – C2: one surface's NH_3 does participate of S1, assisting in the proton transfer.
- S1 – mixture – C3: several surface's NH_3 molecules are coordinated to the reactants, but are not engaged in the proton transfer.

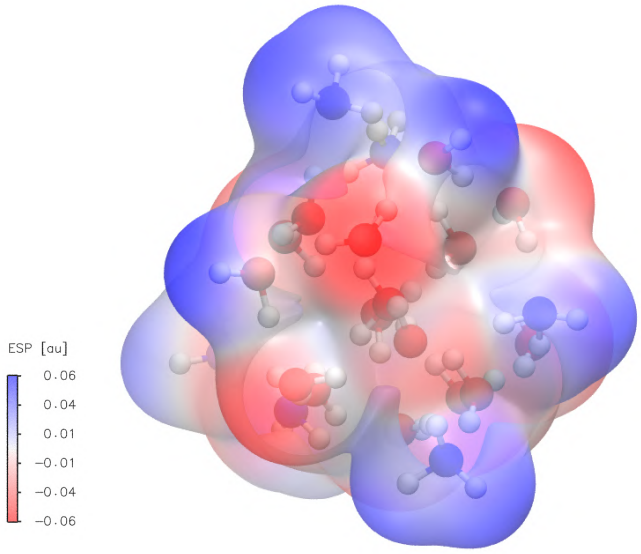


Figure 3.51: Map of electrostatic potential (MEP) calculated for the ice mixture surface. Blue corresponds to electron-poor regions (dangling-H bonds) while red to electron-rich ones.

The reactive complex of each case is reported in Figure 3.52 and the energy barrier results are listed in Table 3.15 and Figure 3.53.

The first case, S1 – mixture – C1, the energy barrier is 12.0 kcal mol^{-1}. In terms of reactant geometries, this situation is analogous of S1-C1 on ASW, illustrated in the previous section. The comparison of the energy barrier with the average found for that group (10.3 kcal mol^{-1}), reveals that the ammonia-rich environment is energetically unfavorable for the aminomethanol formation, increasing the barrier by almost 2 kcal mol^{-1}.

The same conclusion applies to Case 2, being the one with the highest barrier (13.52 kcal mol^{-1}) among the situations considered. In terms of geometry, this case resemble the small system S1 – NH_3 – W (b) (Section 3.4.1), where the NH_3 participating of the proton transfer is coordinated to

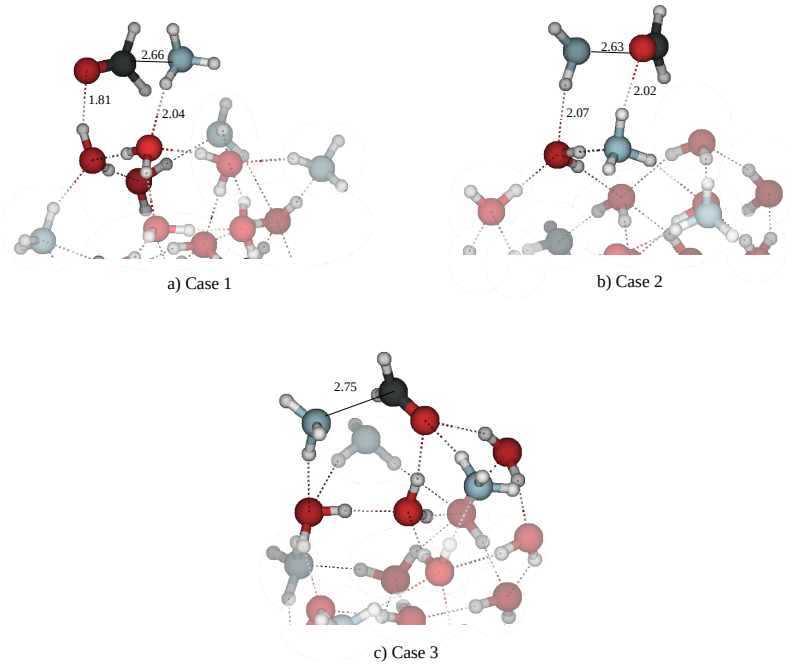


Figure 3.52: Reactive situations taken into account for S1 on ice mixture. a) S1-mixture-C1, b) S1-mixture-C2, c) S1-mixture-C3, according to the description in the text.

the H_2CO. In comparison to the energy barrier of small system, the reaction carried out on the large mixed ice surface presents a smaller barrier by almost 3 kcal mol^{-1}. This result, despite the amount of the overall barrier, provides the evidence of a certain catalytic effect of the larger ice mixture environment applied to $S1-NH_3-W$ (b) reaction mechanism.

Lastly, the effect on the ammonia-rich environment in providing energetic stabilization to the reactants is addressed. In Case 3 structure, H_2CO is creating 3 HBs with the surface: 2 with water molecules and 1 with a NH_3 molecule. In terms of geometry, this kind of coordination is typical of the pair dangling-H sites encountered for the pure ASW surface (S1-C3, S1-C4).

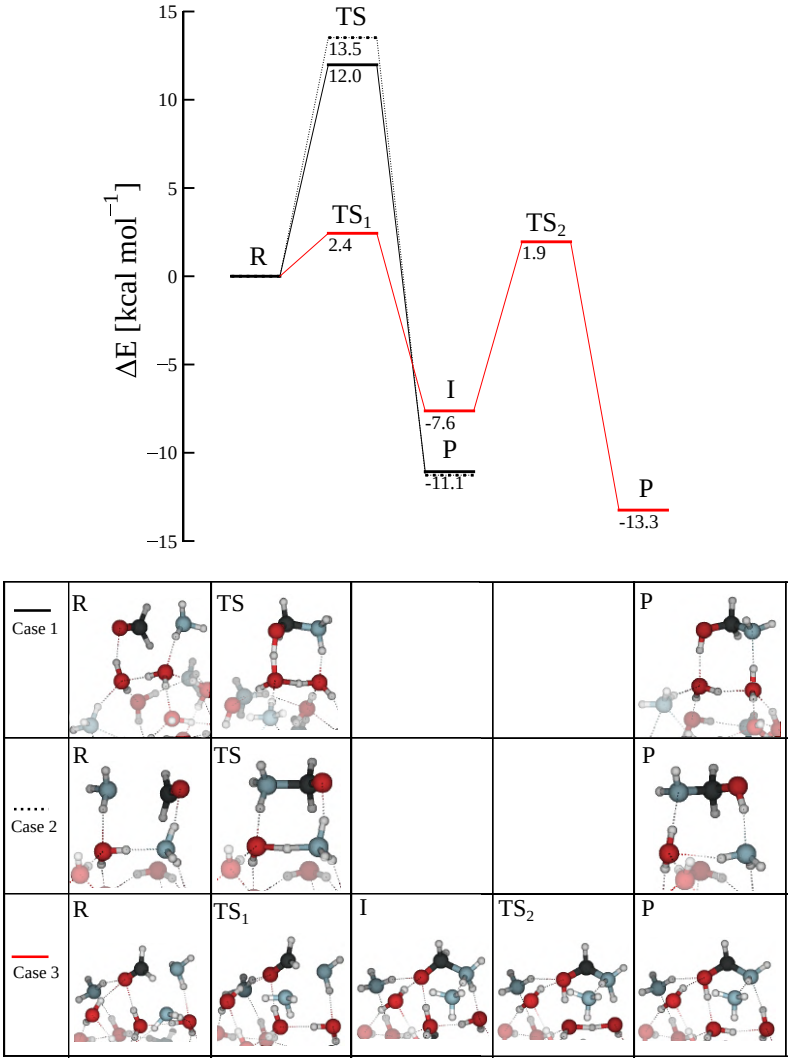


Figure 3.53: Energy diagram for S1 on ice mixture, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. Solid black line: S1-mixture-C1; dashed black line: S1-mixture-C2; solid red line: S1-mixture-C3. Minimum energy structures - reactants R, transition states ($TS_{1,2}$), intermediate (I) and product (P) - are reported for each case.

Table 3.15: Energy barriers ($\Delta E_{En}{}^{\ddagger}$) and reaction energies ($\Delta E_{En}{}^{o}$) for step n of S1 carried out on ice mixture, computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries. Column two, N_X, indicates the number of molecules involved in the proton relay. Values in kcal mol^{-1}.

System	N_X	$\Delta E_{E1}^{\ddagger}$	$\Delta E_{E1}{}^{o}$	$\Delta E_{E2}^{\ddagger}$	$\Delta E_{E2}{}^{o}$
S1-mixture-C1	2	12.0	-11.1		
S1-mixture-C2	2	13.5	-11.1		
S1-mixture-C3	2	2.4	-7.6	9.6	-5.6

In fact, the system in analysis shares the same step-wise mechanism. Both steps are largely exothermic, as in S1-C4, but E1 barrier is not as shallow (2.4 kcal mol^{-1}). Besides, the second step is confirmed to be the limiting step of the process, with a barrier of 9.6 kcal mol^{-1}.

In summary, the energy barriers for the aminomethanol formation on the ice mixture are found to be higher than the pure ASW surface for all the three cases taken into account. A reaction mechanism where the surface NH_3 is participating of the proton transfer is highly unflavoured, as shown also by the preliminary studied carried out for the small systems. Furthermore, the presence of a mixture ice bulk exerts a slightly negative effect on the energy barrier, compared to a S1-C1 type reaction. Once again, the most energetically favorable situation is the hyper-coordinated reactants arrangement, where the surface NH_3 is directly interacting with the reactants. This last case is relative to a step-wise mechanism, analogous of S1-C3,4 type reactions. Despite the greater polarization effects provided by the ammonia-rich

ice on the reactants, the two steps barriers are larger and not as exothermic as for the pure ASW.

3.5.2 Stage 2: Methanimine formation

In order for the assisted dehydration of the aminomethanol (S2) to take place, the molecule needs to be HB-bonded to the surface in a specific arrangement: the amino-group requires a surface HB acceptor group to receive the proton and bridge it to the alcoholic group of the molecule. According to binding mode analysis in Section 3.3.4.1, such binding motif corresponds to BM1(a,b) binding modes, Figure 3.16. 8 suitable structures have been extracted from the aminomethanol BE distribution. 7 of them are of BM1b type, which is the most common in the distribution, and one is BM1a type, much rarer, but more strongly bound to the ice. Such high BE is given by the hypercoordination of the molecule, guaranteed by the presence of two dangling-Hs (Figure 3.16 BM1a), instead of one for BM1b cases. Therefore, in analogy with Stage 1, the results will be illustrated in terms of lone dangling-H sites and pair dangling-H sites.

Lone dangling-H sites

S2 is concerted for all the systems and the set of TS energies is reported in Table 3.16. The average value is 27.3 kcal mol^{-1} with a standard deviation of 2.5 kcal mol^{-1}.

In order to rationalize the energy barrier difference between the systems, the structural parameters of the reactants are analyzed. The plot in Figure

Table 3.16: Energy barriers ($\Delta E^{\ddagger}$) and reaction energies (ΔE^{o}) computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries, for S2. The structures have been labelled alphabetically. Average (Avg) and standard deviation (Std) have been included, as well. Values in kcal mol^{-1}.

System	$\Delta E_{S2}{}^{\ddagger}$	$\Delta E_{S2}{}^{o}$
S2-A	28.8	11.9
S2-B	22.8	6.1
S2-C	26.5	7.5
S2-D	24.6	11.7
S2-E	30.7	10.0
S2-F	28.1	4.8
S2-G	28.8	8.9
Avg (Std)	27.2(2.5)	8.7(2.5)

3.55, left upper panel, shows a tendency between shorter bond distances for the alcoholic group protonation and smaller TS barriers. The shorter is the distance between the dangling-H and the alcoholic group, the easier is the protonation.

A even better tendency can be found analyzing the breaking backbone angle O-C-N , 3.55, right upper panel. Cases where the aminomethanol presents a less strained bond angle (more similar to the model system, 117 degree, marked as star in the plot) show larger TS barriers. On the other hand, higher angle deformation (values between 112 and 113 degree) corresponds to lower TS barriers. As it seems to be the case, reactant structures where the bond breaking angle is bent in a way that minimize the O-H bond forming distance are more reactive with respect to the dehydration.

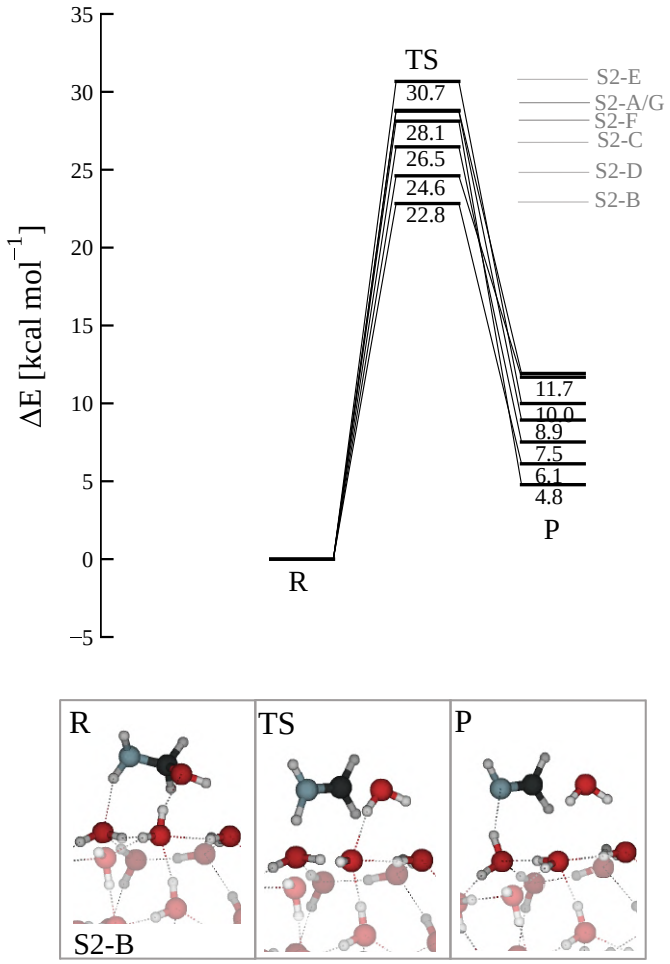


Figure 3.54: Energy diagram for S2−ASW in lone dangling-H sites, computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries. Minima are reported for S2-B, the system that shows lower energy barrier.

Finally, the interaction energy (IE) between the aminomethanol and the ice surface in the set of reactive sites has been calculated as:

$$\mathrm{IE(R)} = \mathrm{IE}(\mathrm{NH_2CH_2OH}\cdots\mathrm{W_{22}}) \tag{3.9}$$

Analysis IE values revealed that most of the structures that presents low TS

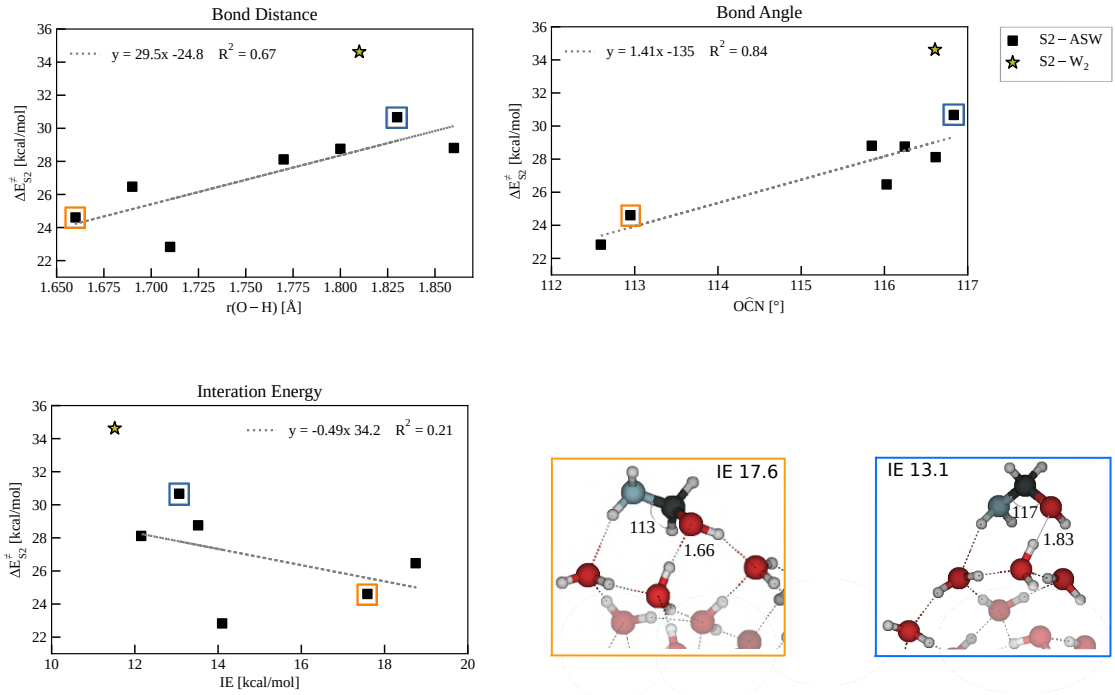


Figure 3.55: Upper panel: O-H bond distance for the alcoholic group protonation and bond angle A(OCN) versus TS energy for the set of S2 – ASW reactants. Lower panel: reactant interaction energy (IE) versus TS energy (left). Example of two reactants at the extreme of the IE distribution (right): high IE, orange, low IE, blue. Structural parameters of the two structures have been highlighted as orange and blue squares in the corresponding plots as well. Parameters of the model system (S2 – W_2) are reported in the plots (yellow star) but have not been used for the linear regressions (dashed line).

energy also shows high IE (> 14 kcal mol^{-1}, Figure 3.55, lower panel, left). In Figure 3.55 lower panel, right, we reported two structures at the extremes of the IE distribution: high IE (orange), low IE (blue). The structural parameters for the two structures have been highlighted as colored squared in all the plots, revealing that they are located at the extremes of the range of each structural parameter, representing two examples of respectively favorable and unfavorable S2 reactants.

In conclusion, the analysis of S2 reactive sites in presence of a lone dangling-H, allowed to define certain structural and energetic characteristics that are the most favorable to undergo aminomethanol dehydration. First, having a highly strained aminomethanol backbone angle that bends the molecule toward the HB donor assisting water. Such feature increases the deformation energy of the system and minimizes the distance to cover in protonating the alcoholic group, which is the step that initiates the proton relay. The short distance between the alcoholic group and the assisting water is reflected by high interaction energies between the aminomethanol and the surface, hence, structures that are strongly adsorbed are energetically favored.

Pair dangling-H sites

Carrying out the reaction in a surface site where a pair of dangling-H is present determines a change in the reaction mechanism, similarly to the results for the first stage.

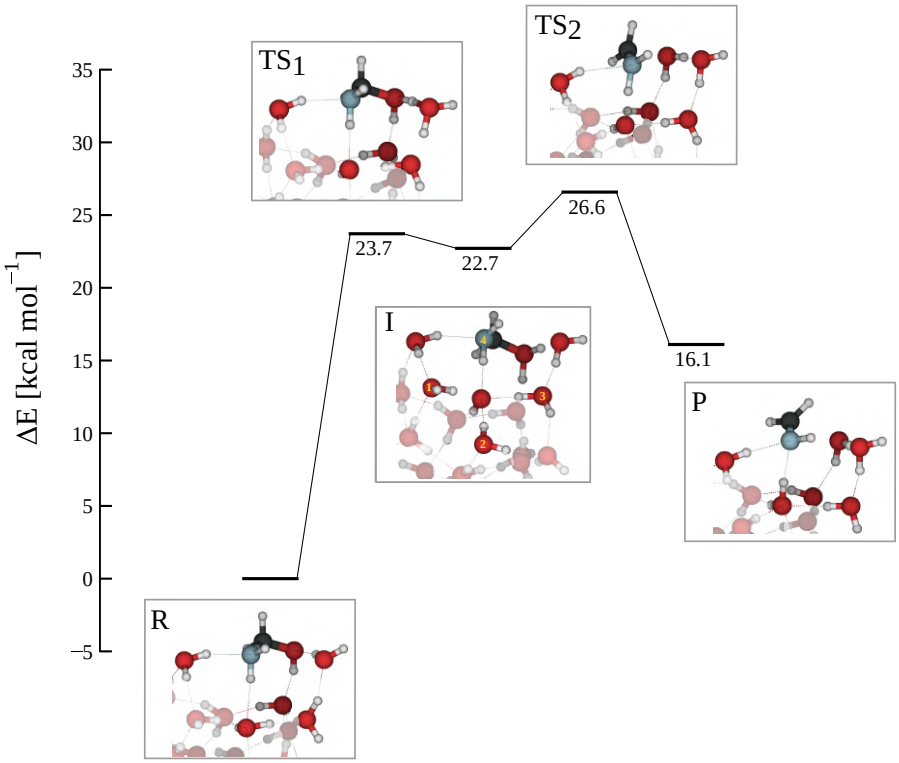


Figure 3.56: Energy diagram for S2 – ASW in a pair dangling-H site, computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries. The groups acting as HB-donor on the OH^- have been highlighted in yellow in the intermediate (I) inset.

The second dangling-H is coordinated to the amino group of the aminomethanol, establishing a HB-donor interaction, corresponding to the binding mode BM1a identified for the species (Figure 3.16). The binding mode corresponds to the strongest arrangement of aminomethanol on ASW.

The dehydration proceeds in two separated steps, TS and reaction and energies are reported in Table 3.17. However, the mechanism - in terms of the order of bond breaking/forming processes - is in agreement with the finding for the small system (Section 3.4.2) and for previous case. The first step (E1)

Table 3.17: Energy barriers ($\Delta E_{En}{}^{\ddagger}$) and reaction energies ($\Delta E_{En}{}^{o}$) for step n of S2 carried out in a site containing a pair of dangling-H bonds, computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries. Column two, W_X, indicates the number of water molecules involved in the proton relay. Values in kcal mol^{-1}.

System	W_X	$\Delta E^{\ddagger}_{E1}$	ΔE^{o}_{E1}	$\Delta E^{\ddagger}_{E2}$	ΔE^{o}_{E2}
S2 pair dangling-H site	3	23.7	22.7	3.9	6.6

is the protonation of the alcoholic group, yielding a dipolar intermediate, I ($HO^- \cdots NH_2CH_2OH_2^+$). The step corresponds to part of the proton transfer steps comprised in the proton relay that bridge the proton from the amino group to the opposite part of the molecule. The negative charged fragment is coordinated to the amino group. The reason for the isolation of the OH^- in that specific ice site is the especially favorable coordination offered by the ice network. In fact, the OH^- in that position is surrounded by 4 HB-donor groups (highlighted in yellow in Figure 3.56, I inset). The first step is the most expensive energetically, requiring the overcoming of a barrier of 23.7 kcal mol^{-1}. The second step, E2, is the elimination of the protonated alcoholic group, along with regeneration of the assisting water ($\Delta E^{\ddagger}_{E2}$ = 3.9 kcal mol^{-1}). The sum of the barrier values of the two steps falls in the range of TS energy found for the lone dangling-H sites, meaning that the hyper-stabilization provided by the ice does not considerably lower the energetic requirements for the process. Besides, this kind of reactive sites are rarest, as they correspond to the high BE tail of the distribution of the aminomethanol.

Nevertheless, passing from a concerted mechanism to a step-wise one where each leg of the proton relay is isolated entails notable advantages, which will be discussed at the end of the chapter (Section 3.5.5).

3.5.2.1 Stage 2 following from Stage 1

Local chemistry mechanism

S2 is also studied using one of the species produced in S1, in order to have a complete picture of the process of aminomethanol synthesis and dehydration on ASW. The idea was to consider a *local* chemistry pathway, where a product of S1 undergoes S2 without the need for structural rearrangements or diffusion to a different ice site. In doing so, the system is expected to take full advantage of the exothermicity of S1. Unfortunately, among the several structures obtained in S1, very few presented the correct configuration to undergo the dehydration. This is expected, as NH_2CH_2OH tends to keep the arrangement in which it has been generated, i.e. with the alcoholic group pointing toward the cluster (resulted of the carbonyl group protonation).

The structure that was used as reactant is the product of S1-C4-B, that belongs to S1-C4 case, providing a extraordinary coordinated reactive site, where 3 dangling-H sites are present in the reactive environment, hence, the aminomethanol produced in S1-C4-B was very much inserted into the HB surface network. The feature turned out to be particularly valuable, since S1-C4-B presents a large number of water molecules surrounding the

aminomethanol, apart from the ones involved in S1, and a pair of them are HB-bonded in the correct arrangement to assist in the dehydration. The energy barriers for the full local pathway is reported in Table 3.18 and Figure 3.57.

Table 3.18: *Local* pathway for the generation and dehydration of aminomethanol (S1+S2). Energy barriers ($\Delta E_{En}{}^{\ddagger}$) and reaction energies ($\Delta E_{En}{}^{o}$) for step n of S1 and S2. Energies at ωB97-M/def2-TZVP // BHANDHLYP-D4/def2-SVP level of theory. Values in kcal mol^{-1}.

S1 (S1-C4-B)				S2	
$\Delta E^{\ddagger}_{E1}$	ΔE^{o}_{E1}	$\Delta E^{\ddagger}_{E2}$	ΔE^{o}_{E2}	$\Delta E^{\ddagger}_{S2}$	ΔE^{o}_{S2}
0.1	-10.7	12.8	-0.8	28.8	11.9

S1 steps have been included, as well, so that the energy diagram displays three energy barriers. $TS_{1,2}$ refer to the exothermic formation of the intermediate between NH_3 and H_2CO and of the product, respectively. TS_3 refers to the aminomethanol dehydration, that is a endothermic process and requires to overcome a barrier of 28.8 kcal mol^{-1}.

The entire pathway (S1+S2) is also slighly endothermic by 0.4 kcal mol^{-1}. The comparison of S2 barrier with the S2-TS we calculated in the previous section, shows that the value belongs to the upper part of the TS distribution. Although, the surface exerts a catalytic role, lowering the TS energy by around 6 kcal mol^{-1} with respect to the model system ($S2-W_2$), the dehydratation presents the highest barrier of the full local pathway and is much larger than the exothermic dissipation generated by S1, as it exceed it by 28.0

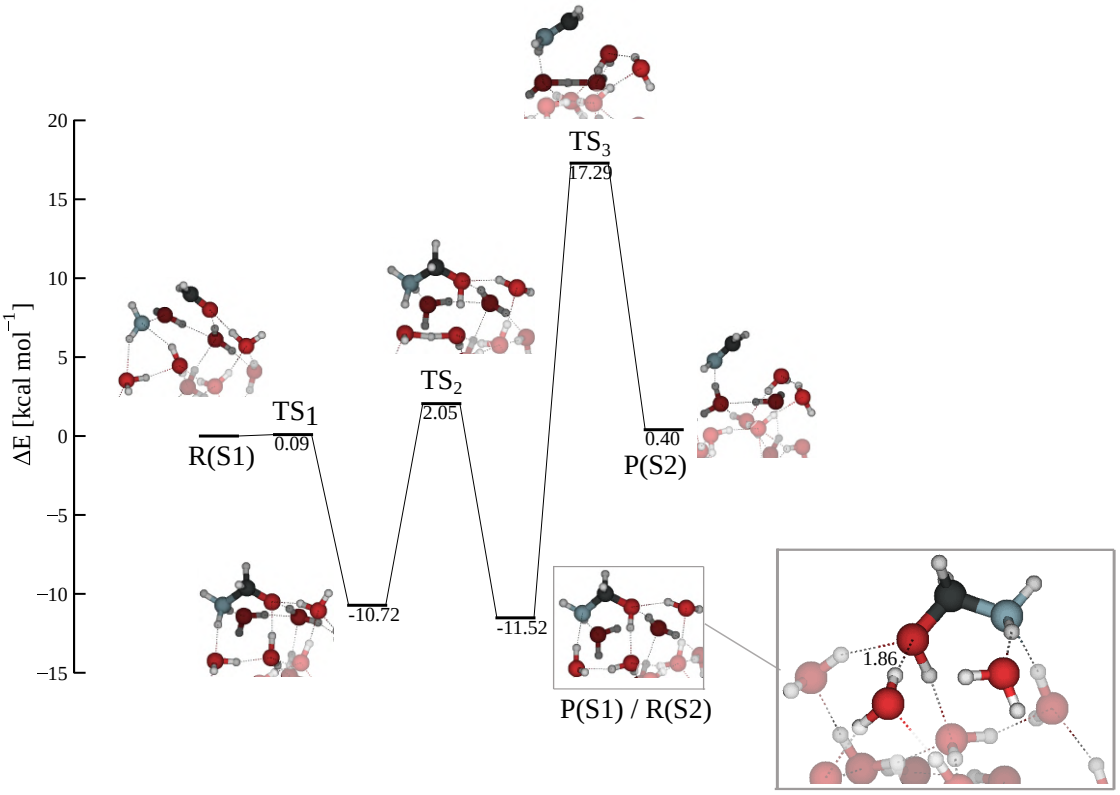


Figure 3.57: Steps for a *local* pathway for the generation and dehydration of aminomethanol (S1+S2). Energy diagram at ωB97-M/def2-TZVP // BHANDHLYP-D4/def2-SVP level of theory. Figure insets: side view, only the water molecules interacting with the reactants have been highlighted. The pair of water molecules involved in S1 are located on the front, while the pair that participates of S2 are on the back. The aminomethanol product of S1 is used as reactant for S2; additional inset on the lower right corner.

kcal mol^{-1}, which makes it a energetically unfavorable process.

In conclusion, we considered a *local* mechanism for the dehydration of the aminomethanol using one of the structures produced in Section 3.5.1. Despite the exothermicity of S1, and the catalytic effect of the ASW surface, the reaction maintains its endothermic character and presents a high TS barrier of 28.8 kcal mol^{-1}, making it an unfeasible reaction pathway.

Non-local chemistry mechanism

A *non-local* mechanism is also considered. It involves a aminomethanol molecules produced in one of S1-ASW cases that undergoes structural recombination to reach the necessary configuration for S2, followed by the dehydration step. There is a great variety of rearrangements that such a 'large' molecule (under astrophysical criteria) might experiment upon diffusion on the surface. In order to facilitate the analysis - and limit the arbitrariness of the recombination - the simplest and most straightforward process identified among the pool of cases at disposal was chosen. The reactant structure is again from S1-C4 group (S1-C4-C, Section 3.5.1), which presented a suitable reaction site for S2 nearby the original S1 site. The new site can be easily reached by rotating the C−N bond, i.e. changing the orientation of the amino group toward the dangling-H of the new assisting water. Basically, only one of the moieties of the molecule is required to recombine to undergo S2, keeping the original site mostly unchanged. Figure 3.58 and Table 3.19 illustrates the complete *non-local* pathway.

As for the *local* case, S1 steps are included for completeness. The S1-C4-C reaction mechanism (TS_{1-3}) has been described in details in Section 3.5.1. TS_4 refers to the structural recombination. This process is slightly exothermic and requires to overcome a barrier of 3.1 kcal mol^{-1}. Finally, the dehydration takes place TS_5, with a barrier of 22.8 kcal mol^{-1}. This value is

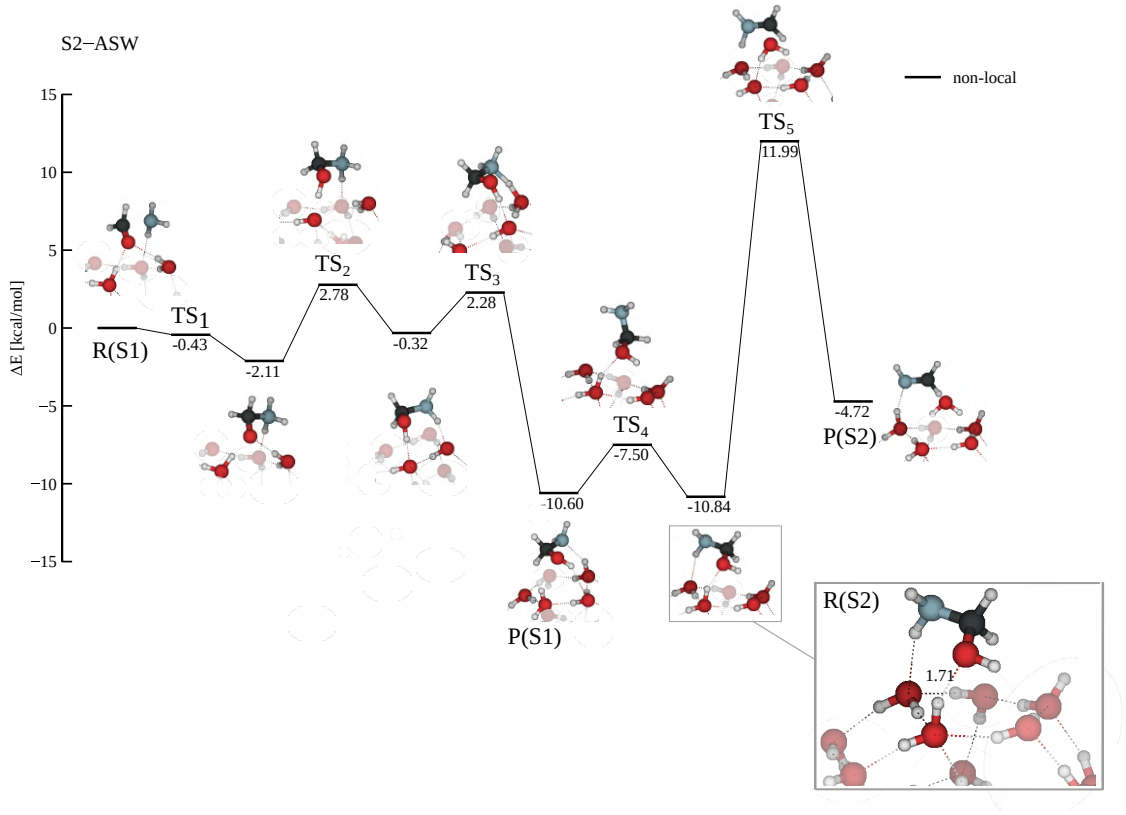


Figure 3.58: Steps for a *non-local* pathway for the generation and dehydration of aminomethanol (S1+S2). In between, the molecule undergoes structure recombination (TS_4). Energy diagram at ωB97-M/def2-TZVP // BHANDHLYP-D4/def2-SVP level of theory. Figure insets: side view, only the water molecules interacting with the reactants have been highlighted.

the lowest we encountered so far for S2, corresponding to a surface catalytic effect of almost 12 kcal mol^{-1} with respect to the model system. Such large decrease in the energy barrier is likely due to the system flexibility that allows the molecule to enhance the interaction with the surface site, compared to the *local* reactive site, whose high level of insertion into the HB network conferred rigidity to system. In fact, analyzing the structural characteristics of the non-local reactant, it can be noted that it meets the criteria previously

Table 3.19: *Non-local* pathway *Non-local* pathway for the generation and dehydration of aminomethanol (S1+S2). In between, the molecule undergoes structure recombination ('Strc. recomb.'). Energy barriers ($\Delta E_{En}{}^{\ddagger}$) and reaction energies ($\Delta E_{En}{}^{o}$) for step n of S1, Strc. recom. and S2. Energies at ωB97-M/def2-TZVP // BHANDHLYP-D4/def2-SVP level of theory. Values in kcal mol^{-1}

S1 (S1-C4-C)						Strc. recomb.		S2	
$\Delta E_{E1}^{\ddagger}$	ΔE_{E1}^{o}	$\Delta E_{E2}^{\ddagger}$	ΔE_{E2}^{o}	$\Delta E_{E3}^{\ddagger}$	ΔE_{E3}^{o}	$\Delta E_{r}^{\ddagger}$	ΔE_{r}^{o}	$\Delta E_{S2}^{\ddagger}$	ΔE_{S2}^{o}
-0.4	-2.1	4.9	1.8	2.6	-10.3	3.1	-0.2	22.8	6.1

listed to be more favorable for S2 (among the lone dangling-H sites).

In summary, S2 *non-local* pathways is comprised of two steps: the diffusion of the reactant to a nearby binding site, suitable for S2, and the NH_2CH_2OH dehydration. The first step is slightly exothermic and presents a barrier of around 3 kcal mol^{-1}. On the other hand, the dehydration is endothermic and presents a barrier of 22.8 kcal mol^{-1}. Despite belonging to the lower part of the TS energy range found for S2 in lone daingling-H sites, the barrier is considerably greater than the energy surplus from S1. It is concluded that both the local and non-local pathways are not feasible formation pathways for methanimine under interstellar conditions.

3.5.3 Stage 3: Aminoacetonitrile formation

The study of S3 mechanism for the model system (Section 3.4.3) revealed that the first event (E1) is HNC deprotonation, which suggests a central role of HNC. Thus, we evaluated a set of possible arrangements of HNC adsorbed on ASW, as starting point for the reaction. According to the binding mode analysis in Section 3.3.4.1, the necessary adsorption motif of HNC to undergo S3 is BM1 (Figure 3.19), and it exists in two variation: BM1a, the highest in BE, and BM1b, the most common in the BE distribution.

The computational procedure to carry out S3 on ASW is the following:

1. selection of a suitable BM1a and BM1b binding sites for HNC
2. random sampling of $NHCH_2$ around HNC in the two binding sites This step provided a set of potential reaction sites.
3. selection of 2 reactive sites where $NHCH_2$ also presents the correct coordination to S3
4. TS search and characterization, which has been successful for both systems; and TS energy calculation

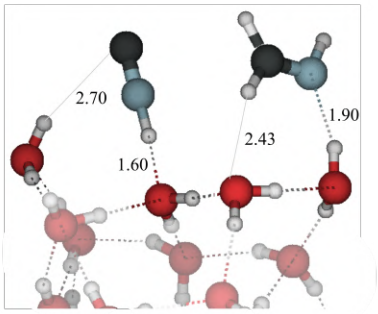
The resulting reaction sites are displayed in Figure 3.59 and corresponds to different situations. Namely, in the first case, Figure 3.59a, both fragments are interacting strongly with the ice. HNC is establishing two HBs, as previusly mentioned, as well as $NHCH_2$. The arrangement of $NHCH_2$ in this site

correspond to the high BE motif in the binding mode analysis: BM1a, Figure 3.9, also suitable to carry out S3. Therefore, $N_{HB} = 4$.

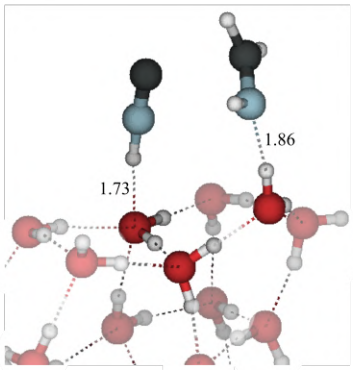
In the second case, Figure 3.59b, both the reactants are loosely bound to the surface: HNC is establishing a single HB (corresponding to its BM1b mode, Figure 3.18), while $NHCH_2$ is interacting solely trough the NH moiety (corresponding its BM1b mode, Figure 3.17). In total $N_{HB} = 2$. The strength of the surface influence on the reactants is reflected by their interaction energy (IE) with the surface, estimated as:

$$IE(R) = -IE(\mathrm{HNC} + \mathrm{NHCH_2} \cdots \mathrm{W_{22}}) \tag{3.10}$$

As expected, IE is larger for S3-C1 (23.2 vs 19.0 in S1-C2 kcal mol^{-1}). According to reactant characteristics, summarized in Table 3.22, the two cases are considered examples of *strong* and *mild* surface effect, respectively. Moreover, the two reactive sites present respectively one and two dangling-Hs, therefore, in agreement with previous stages, the results will be presented in light of that characteristic.



a) S3 - C1



b) S3 - C2

Figure 3.59: Reactive sites selected for S3 on ASW. a) Pair dangling-H site. Both reactants are strongly bound to the surface b) Lone dangling-H site, the reactants are loosely bound.

Table 3.20: Categorization of the reactive sites for Stage 3 (S3) in two cases. First column reports the structures. The categorization is based on: total number of HBs established by the reactants with the ice (N_{HB}, second column), HB type (third column, as illustrated in Section 3.3.4.1, and their interaction energy with the surface ($IE(R)$, forth column), as defined in Equation 3.10, in kcal mol^{-1}.

Case	N_{HB}	$IE(R)$
Lone dangling-H site: mild surface effect	2: T1,T1	19.0
Pair-dangling-H site: strong surface effect	4: T1,T2,T4,T7	23.2

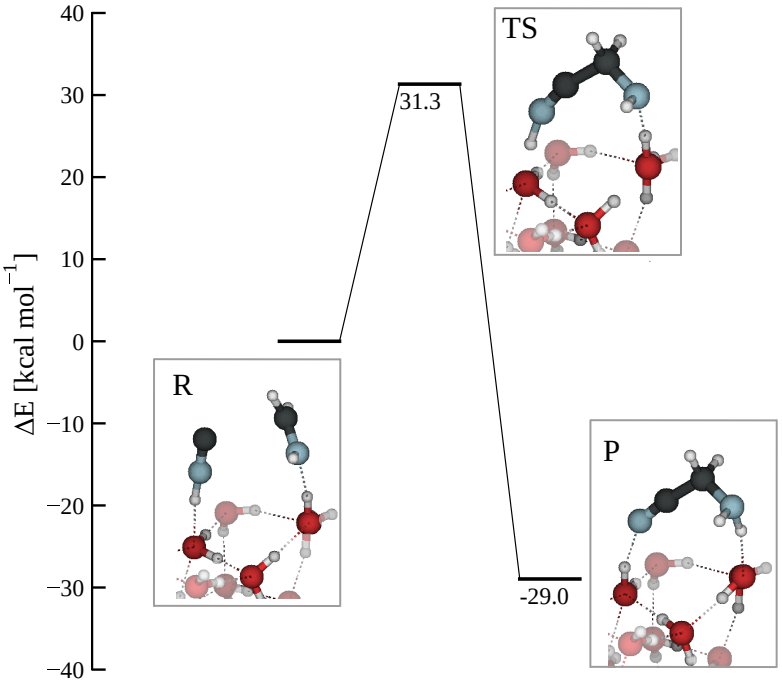


Figure 3.60: Energy diagram for S3 – ASW in a lone dangling-H site, computed at ωB97-M/def2-TZVP level of theory, using PW6B95-D4/def2-SVP geometries.

Lone dangling-H site: mild surface effect

The lone dangling-H reactive site presents both fragments interacting mildly with the surface. The reaction is is concerted and display a high reaction barrier of 31.3 kcal mol^{-1}. This value surpasses the model system by 14 kcal mol^{-1}, suggesting a different reaction mechanism. In fact, analysis of the IRC plot, reported in Figure 3.61a, revealed that the order of reactive events in S3 are reversed with respect to the model system (Section 3.4.3). C – C bond formation takes place first (E1) and is followed by the deprotonation of HNC and the rest of the proton transfer events (E2). An interesting characteristic of the reaction site in study, is that the proton transfer takes place in a 3-

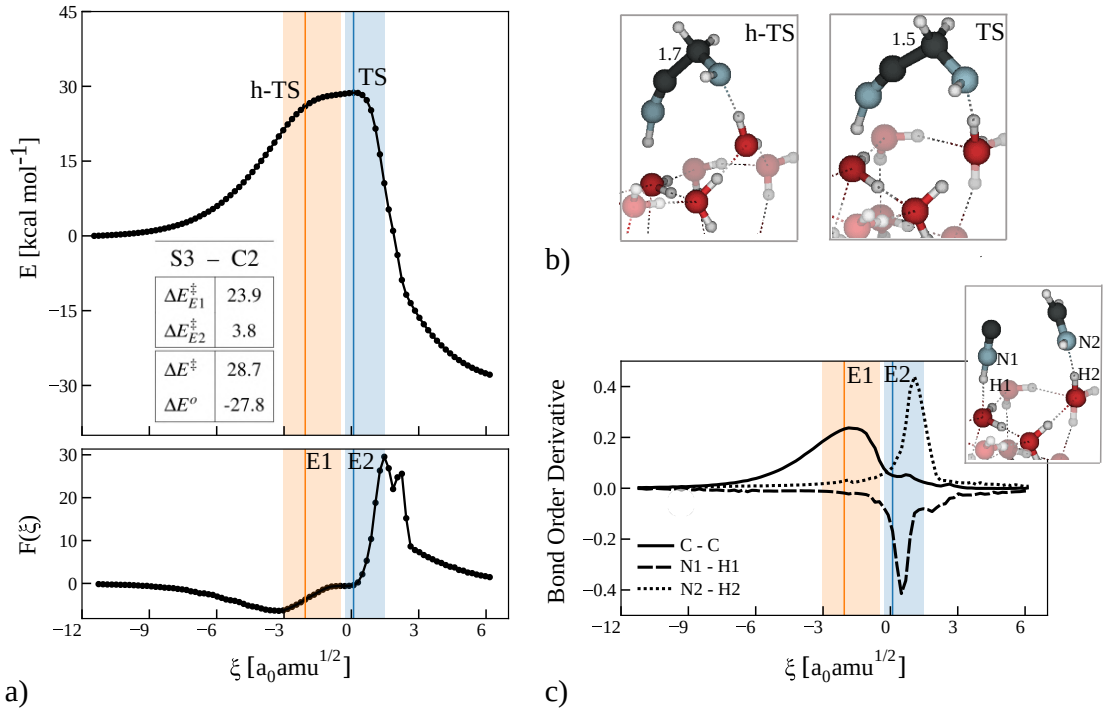
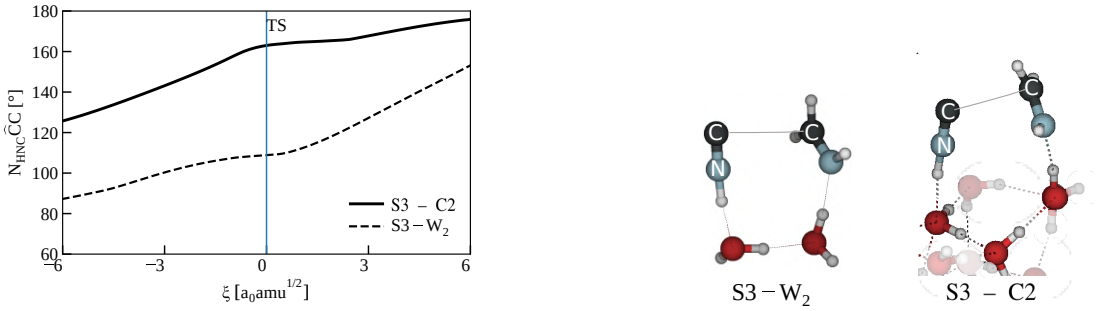


Figure 3.61: a) Energy (upper panel) and reaction force profiles (lower panel) for S3-C1, using BMK/def2-TZVP//PW6B95-D3BJ/def2-SVP levels of theory. Blue and orange lines represent TS and h-TS respectively. TS and h-TS regions are displayed as matching color shadowed areas. The figure include a Table that reports TS energy ($\Delta E^{\ddagger}$) and reaction energy (ΔE^{o}) extracted from the IRC, as well as the partition of the barrier: $\Delta E^{\ddagger}_{E1}$ is relative to the h-TS, while $\Delta E^{\ddagger}_{2}$ to the TS. b) h-TS and TS structures as extracted from the energy profile. c) Bond order derivative for the main bond distances involved.

water-molecules proton relay, unlike all the processes explored so far. Several attempts to perform the reaction in a 2-water proton relay fashion have been unsuccessful. A explanation might be found in the fact that the product has a 4 atoms backbone chain and is almost linear, witch would require a very large deformation of the ice in order to accommodate its formation *via* a 2-water proton relay.

h-TS and the TS regions in the IRC are close by, such that the shape of the IRC is very broad at the top. In order to shed light on the reason behind the change of mechanism, the variation of the bond angle ($N_{HNC}CC$) along the reaction coordinate was examined (Figure 3.62). The comparison with the model system, also reported in the plot, shows that the bond angle at the beginning of the reaction is quite different for the two systems. $S3-W_2$, dashed line, displays a more acute angle (79.1 degrees) which gets wider along the reaction path, reaching the linearity in the product angle. Most of the angle broadening for the model system takes place after the TS, since that is when $C-C$ bond forms. On the other hand, the bond angle in the lone dangling-H ASW site (solid line) presents a large initial value (117.8 degrees) and increases mostly in the region that precedes the TS, in agreement with the anticipated $C-C$ bond formation in the mechanism.

It can be concluded that the reason for the change in mechanism resides in the geometry of the reactive site, as it exhibits a bond angle closer to the value found in the product. Besides, according to the binding modes identification



a) b)

Figure 3.62: a) Comparison between the evolution of the S3 bond angle along the reaction coordinate. b) Systems in comparison: model system (left) and reactants in the lone dangling-H site (right).

illustrated at the beginning of the section, this case corresponds to the reaction between two loosely bound fragments, hinting that their high mobility allow them a certain degree of flexibility on the ice. As a result, the fragments in the initial reactive complex have the possibility to orient toward each other. Such an arrangement favors the backbone C–C bond formation, and leads to a change in the S3 mechanism with respect to the model system. However, the new mechanism shows a very large energy barrier of 32.7 kcal mol^{-1}.

Pair dangling-H site: strong surface effect

Similarly to the findings for S1 and S2, when the process takes place in reactive site that allows the coordination of the reactants to two dangling-H atoms, S3 reaction unfolds in a step-wise manner.

The energy diagram is reported in Figure 3.63. Although the result is different to what was observed for S3 model system, the order in the pathway

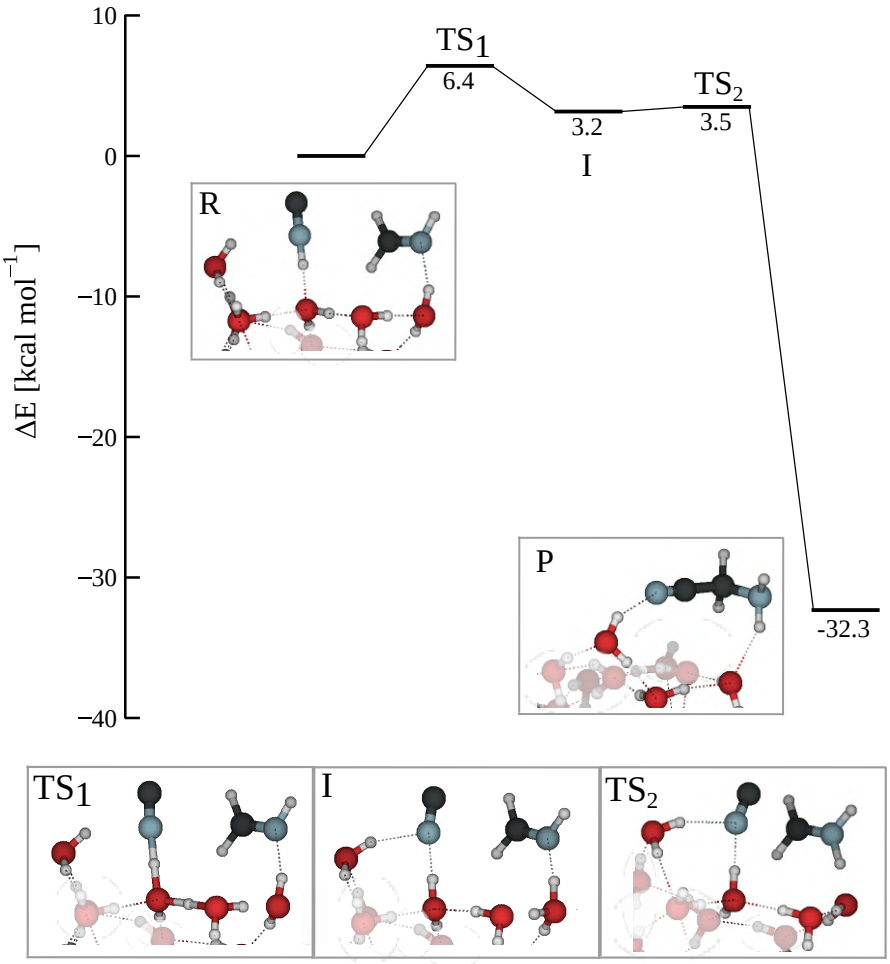


Figure 3.63: Energy diagram for S3 – ASW in a pair dangling-H site, computed at ωB97-M/def2-TZVP level of theory, using PW6B95-D4/def2-SVP geometries.

Table 3.21: Energy barriers ($\Delta E_{En}{}^{\ddagger}$) and reaction energies ($\Delta E_{En}{}^{o}$), computed at ωB97-M/def2-TZVP level of theory, using PW6B95-D4/def2-SVP geometries, for step n of S3 on a lone and a pair dangling-H sites. Column two, W_X, indicates the number of water molecules involved in the proton relay. Values in kcal mol^{-1}.

System	W_X	$\Delta E_{E1}{}^{\ddagger}$	$\Delta E_{E1}{}^{o}$	$\Delta E_{E2}{}^{\ddagger}$	$\Delta E_{E2}{}^{o}$
Lone dangling-H: mild ice effect	3	31.3	-29.0		
Pair dangling-H: strong ice effect	3	6.4	3.2	0.3	-35.5

steps reflect the events previously identified for S3–W_2. In fact, the study revealed a hidden transition state (h-TS, Figure 3.31) that, thank to the stabilization provided by the ASW, becomes a well characterized TS (TS_1 in Figure 3.63). Table 3.21 shows that S3 maintains its overall exothermic character. The first step, E1, HNC deprotonation and proton relay, is endothermic and has a energy barrier of 6.4 kcal mol^{-1}, which corresponds to a catalytic effect of around 12 kcal mol^{-1} with respect to the model system. The proton relay is established among 3 assisting water molecules, as to accommodate the newly forming $NCCH_2NH_2$. The notable barrier decreasing can be attributed to the influence of the additional dangling-H coordinated as HB donor to HNC, as it increases its proton donor character, facilitating the deprotonation. Another consequence of the significant surface effect on the proton-relay event is the emergence of the dipolar intermediate (I), also stabilized by the presence of the aforementioned dangling-H and also of the newly established HB donor coordination of the bridging water.

The intermediate ($NC^{-}\cdots NH_2CH_2^{+}$) was characterized for the model

system, where it appeared as transient point, h-I. Basically, the species presents large intermolecular polarization, as electron-rich regions of the system are located mostly on the deprotonated HNC, while $NHCH_2$ side is electron-depauperated.

Step 2, E2, is the reaction between the ionic fragments of I, it is exothermic and almost barrierless. Basically the system requires only enough energy for the fragments to escape the ice coordination and form the product.

A IRC analysis is carried out solely of the first step of S3, due to the flatness of the potential energy surface of the second step and the practically barrierless nature of that process. The plot is reported in Figure E.6. The main finding of the analysis of the mechanism along the IRC is that the proton transfer events in E1 are also highly asynchronous, to such extend that E1 is separated into two distinct reactive events: the proton transfer is initiated by the deprotonation of HNC (E1a) and the proton migration proceeds until reaching the third assisting water (E1b) where it stops. Exam of the coordination of the last assisting water comprising the proton-relay bridge, unraveled that it has two surface groups around them acting as HB acceptor, worsening its proton donor character. Therefore, the last proton transfer step is delayed due to *second order* surface effects (Section 3.5). The fact that the proton relay steps are spaced on the reaction coordinates confers a very narrow character to the TS region.

In summary, S3 has been studied on ASW in a sites that present a pair

dangling-H bonds, which confers remarkable energy stabilization to the reactants. The reaction proceeds via a step-wise pathway. The mechanism reflects the finding for the model system and is comprised on two events: the proton transfer, initiated by HNC deprotonation, and the establishment of C−C bond. The ASW environment provides such stabilization to the reactants leading to the convertion of the model system dipolar hidden intermediate (h-I) in a real one, similarly to the findings for S1 and S2, in passing from a lone dangling-H site to a pair one. As a consequence, the energy barrier for the proton transfer is sensibly lowered from 18.7 kcal mol^{-1} to only 6.4 kcal mol^{-1}, while the C−C backbone bond formation is almost barrieless.

Surface strong effects leading to an alternative mechanism: HCN deprotonation

Finally, an inquiry was made into the first step of the proton transfer, specifically the deprotonation of HNC. As pointed out, the low character of the TS barrier in the pair danglinh-H site is due to the high BE arrangement of HNC which favors the species deprotonation and, in turn, the stabilization of the dipolar intermediate. Thus, aminoacetonitrile formation was investigated in the same extraordinary favorable reaction site but using HCN as a reactant. HCN and HNC are known to be present in similar amount in the ISM and to isomerize easily on ASW, inter-converting with each other[225].

The alternative mutated reactive site is reported in Figure E.4R. HCN ar-

rangement corresponds to a high BE binding mode for HCN molecule (BE range: 5-7 kcal mol^{-1}), analogous to HNC one in the pair dangling-H reactive site. Solely S3 deprotonaton step was studied (Figure E.4). The calculated energy barrier of 10.3 kcal mol^{-1} is larger than the finding for the deprotonation of HNC (6.4 kcal mol^{-1}), under - it could be safely assumed - the same environmental conditions.

The result is in part expected, as HNC is a more reactive species than HCN and a stronger acid[224]. Besides, the dipolar intermediate, Figure E.4I, obtained by HCN deprotonation, presents a non favorable orientation of the fragments with respect to C – C backbone formation. Therefore, greater structural rearrangements are expected in order to attain the product geometry. Although the test case is not meant for statistical purposes, the finding reinforces the conclusion that S3 reaction mechanism on ASW takes places in a exothermic step-wise mechanism initiated by the deprotonation of HNC, highly assisted by the ice environment.

3.5.4 Stage 4a: Glycoamide formation

S4 is comprised of 3 steps. The first one (S4a) is the hydrolysis of the aminocetonitrile. S4a was studied using small size systems ($NCCH_2NH_2$ + H_2O + $W_{1,2}$, Section 3.4.4). This section illustrate the results on ASW surface. The system is study is: ($NCCH_2NH_2$ + H_2O + W_{22}).

Given the elevated energy barrier observed in the model systems and the low polarizability of the reactive site, emphasis shifted towards the search for alternative proton transfer mechanisms.

Two cases were taken into account, in the first one, S4a-C1, the hydrolysis is assisted by one W, (1-water proton relay). The choice was driven by the fact that S4aW_1 displayed the smallest barrier among the small systems, hence the idea to apply that reaction mechanism to the ASW surface.

The second case, S4a-C2, considered a mechanism where the proton transfer is mediated by a very large number of water molecules, figuratively by the *bulk* of the surface. This hypothesis had its foundation on the attempt to diminish the straining of TS structure in S4a. Therefore, the procedure to identify possible initial reactive site for S4a on ASW has been imprinted to ensure such variability in the TS geometries. The reactants configurations for S4a-C(1,2) have been selected from $NCCH_2NH_2$ BE distribution, according to the binding mode analysis, Section 3.3.4.1, Figure 3.19. The water addition has been carried out using one of the surface molecules as nucleophile

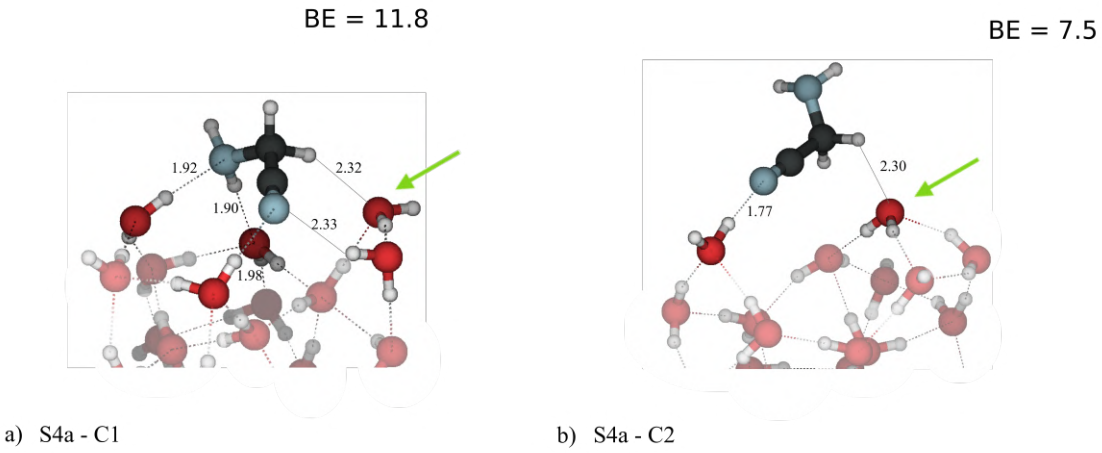
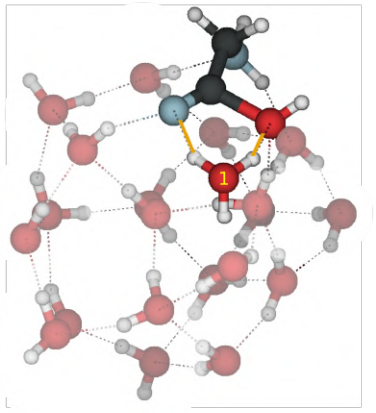


Figure 3.64: Reactants geometries using BHANDHLYP-D4/def2-SVP for a) S4a-C1 ; b) S4a-C2. The H_2O acting as nucleophile has been highlighted with a green arrow. Both structures have been selected from $NCCH_2NH_2$ BE distribution. BE computed at ωPBE-D3BJ/def2-TZVP // HF-3c/MINIX. Values in kcal mol^{-1}.

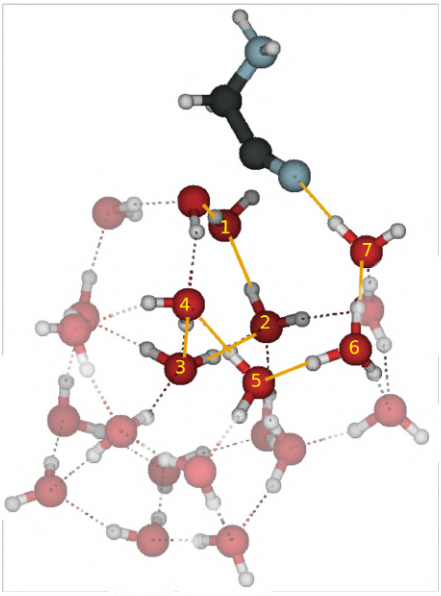
(highlighted in green in Figure 3.64), instead of sampling the nitrile site with H_2O. Structure in Figure 3.64a is engaged in the larger number of interactions, as the reactive site presents three dangling-H bonds. The BE of the molecule is consequently high. On the other hand structure in Figure 3.64a is an example of lone dangling-H reactive site, and is weakly bound to the ice.

Table 3.22: Categorization of the reactive sites for Stage 4a (S4a) in 2 cases (S4a−C(1-2)). First column reports the structures. The second column indicates the total number of HBs established by the reactants with the ice (N_{HB}). The third column reports aminoacetonitrile BE.

Case	N_{HB}	$BE(R)$ [kcal mol^{-1}]
S4a–W_1-like: S4a-C1	5	11.8
bulk H-transfer: S4a-C2	2	7.5



a) S4a - C1 (TS)



b) S4a - C2 (TS)

Figure 3.65: TS geometries at BHANDHLYP-D4/def2-SVP level of theory for a) S4a-C1 ; b) S4a-C2.

Proton transfer mediated by one assisting water

S4a-C1 yielded a reaction barrier of 36.4 kcal mol^{-1} (Table 3.23 and energy diagram in Figure 3.66). The TS geometry is reported also in Figure 3.65a; the molecules that participates of the proton relay have been highlighted and enumerated. The fact that the energy barrier is only almost 2 kcal mol^{-1} smaller than in the model system, suggests a marginal surface catalytic effect. Moreover, the reaction loses its exothermic character, displaying a endothermicity of 3.3 kcal mol^{-1}. Examining the reactive site led to the conclusion that the rigidity of the system (since the nitrile is coordinated to the ice by both extremities) disfavors the reaction, adding up to the characteristic S4a

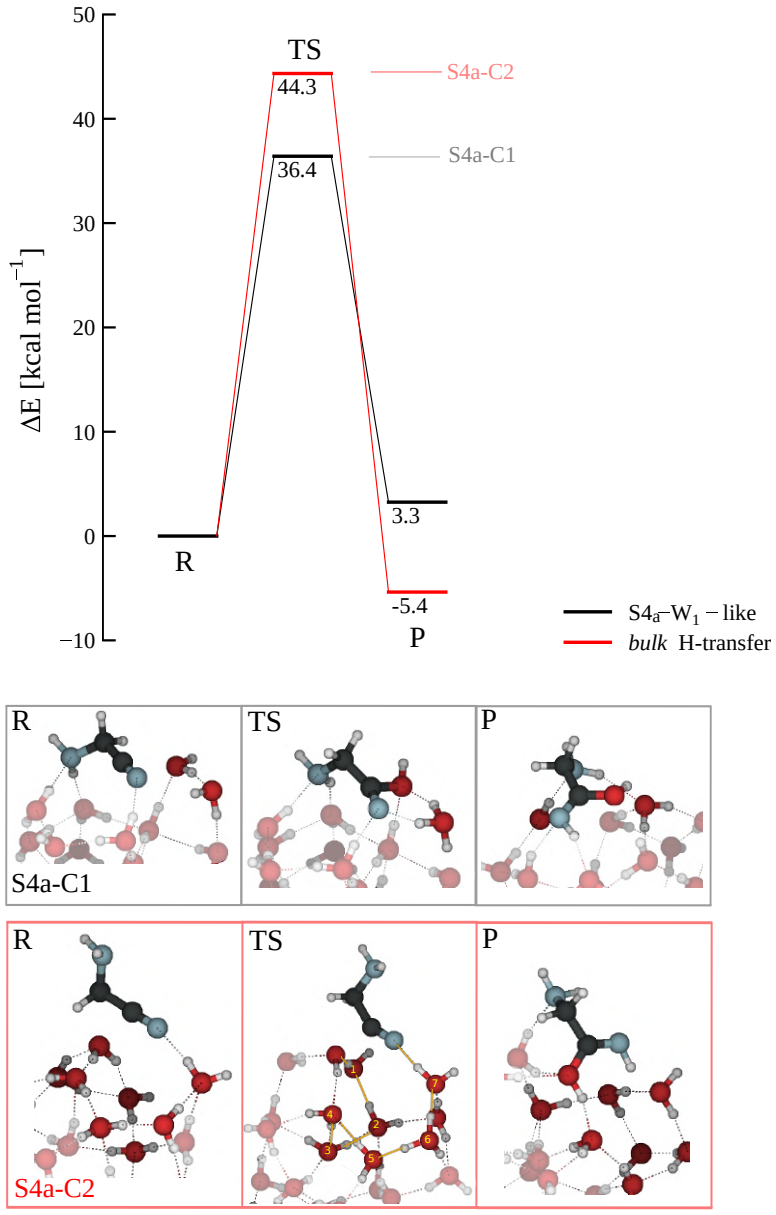


Figure 3.66: Energy diagram for S4a-C(1-2) structures, using BHANDHLYP-D4/def2-SVP geometries. Energies have been computed at ω-B97-M/def2-TZVP level of theory. Minimum energy structures - reactants R, transition state (TS) and product (P) - are reported. The color scheme for the atoms is red for O, black for C, blue for N and white for H.

sterical hindering that has been discussed for the model system.

Table 3.23: Energy barriers ($\Delta E^{\ddagger}$) and reaction energies (ΔE^{o}) computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries, for S4a – ASW. Column two, W_X, indicates the number of water molecules involved in the proton relay. Values in kcal mol^{-1}.

System	W_X	$\Delta E^{\ddagger}$	ΔE^{o}
S4a-C1	1	36.4	3.3
S4a-C2	7	44.3	-5.4

Proton transfer mediated by the ice bulk

Case S4a-C2 took into account a TS that minimizes $NCCH_2NH_2$ deformation, establishing a very large proton relay where the nucleophilic water that initiates the proton transfer and the last proton donor assisting water are not interacting between themselves, as they belong to different regions on the surface. As it can be seen by the TS geometry reported in Figure 3.65b, the pathway that allowed the proton to displace from the H_2O to the cyano group involves 7 water molecules.

Despite the higher flexibility of the nitrile in this reaction site compared to S4a-C1, as it is interacting with the ice only through its cyano group, the TS energy for this case is higher than in S4a-C1 and S4aW_1 model system by 8 and 6 kcal mol^{-1} respectively, although the pathway maintains its exothermic character.

In conclusion, the two cases explored for S4a reaction on ASW present

very meager catalytic effect of the surface with respect to the model system. Furthermore, the 1-W proton transfer pathway performed on the ASW surface, reaction mechanism that was the most favorable among the small systems, resulted in the loss of exothermicity. On the other hand, the attempt to reduce the deformation energy of the system, carrying out a 7-W proton transfer between water molecules far apart on the surface, and displacing the proton into the *bulk* of the ice, did not show promising results either.

In light of those results, it is concluded that the hydrolysis of aminoacetonitrile on ASW surfaces is highly improbable. Therefore, the successive steps of Stage 4 have been excluded from the investigation.

3.5.5 Discussion

Nature of the ice catalytic sites

Within this thesis, a novel approach is introduced to study reactivity processes on interstellar ice models. The application of the methodology to reactions associated with the Strecker synthesis of glycine allowed for a more accurate computation of reaction energies and transition state barriers for processes taking place on a amorphous ice surface. The findings indicate a pivotal role of the morphology of the surface in determining the reaction mechanisms. In light of the comprehensive study, especially for the first stage of the synthesis, a proposed categorization of the ice sites into three types is suggested, based on their catalytic importance:

1. *Common catalytic site*: Essentially, any site where the reaction can occur, characterized by the presence of a dangling-H (so-called lone dangling-H site).

2. *Hyper-catalytic site*: Characterized by the presence of two dangling-H in close proximity, facilitating strong binding of the fragments (so-called pair dangling-H site).

3. *Nano-porous catalytic site* or concavities: dangling-H site within hollow regions.

The first stage was studied on 15 reactive sites. Three variations of the main reaction mechanism were found, depending on the catalytic site involved. The nucleophilic addition reaction of ammonia to formaldehyde proceeds in a concerted manner in the presence of common type catalytic sites (lone dangling-H) and has a barrier in the range of 9-10 kcal mol^{-1}. When the reactive site is hyper-catalytic (pair dangling-H) the strong non-covalent interaction between formaldehyde and the ice site causes hyper-polarization of the molecule. As a consequence, the mechanism converts from concerted to step-wise, since the approach of ammonia leads to the formation (with or without a barrier) of a dipolar intermediate between the two molecules: ($^{-}OCH_2 \cdots NH_3^{+}$). The presence of such an intermediate, stabilized by the ice molecules, had been documented in the works of Chen and Woon [103] and Riffet et al. [100], in the presence of a surface model of 9 and 4 water molecules, respectively. In those studies however, the conditions under which the intermediate was isolated, were not investigated. According to this thesis study, the dipolar intermediate is formed exothermically and without a barrier when both fragments are strongly bound to the surface. It is formed endothermically, instead, when only the formaldehyde presents a high binding energy arrangement, while the adsorption motif of ammonia is of secondary relevance. In that case the formation has a barrier of 3-5 kcal mol^{-1}. The product is obtained from the dipolar intermediate in a subsequent step through a concerted proton transfer, which requires overcoming a barrier of

3-12 kcal mol^{-1}, depending on the reaction site, and is the limiting step of the reaction. An inspection of ASW_{500} periodic surfaces generated in Section 3.1.4, enabled an estimation of the prevalence of common and hyper-catalytic sites. On average, each surface might accommodate $\sim$100 unique sites, 25% of them are estimated to harbour lone dangling-Hs, while 5% being hyper-catalytic sites. Consequently, it is deduced that hyper-catalytic sites do exist on realistic ices, and are present in fair amounts.

When the reaction takes place within nano-pores, a significant variation in the mechanism occurs: not only the dipolar intermediate form without overcoming a energy barrier, but it is also impossible to isolate a non-interacting configuration of the two reactants within the porous. This suggests that nano-pores catalytic sites are primarily important in determining the encounter of molecules, as well as subsequent reactivity. Furthermore, once the dipolar intermediate is formed, the proton transfer step also proceeds in a step-wise manner, as the environment favors the isolation of a second hemiaminal-type intermediate, where the carbonyl protonation has already taken place. This second intermediate had also been found in Chen and Woon [103] work, in the presence of a surface of 12 water molecules. The other steps of the proton transfer take place separately. Another important property of the nano-pores catalytic sites is, therefore, evident: the ability to isolate individual steps of the proton relay in the reaction. Normally, ice-mediated proton transfers occur in a concerted manner, establishing a cyclic ring (proton-relay). The fact

that the steps can be isolated in the presence of pores opens up important consequences for reactivity in the interstellar medium, which will be discussed in the upcoming section. Carrying the first step within the nano-porous requires overcoming two energy barriers of 5 and 3 kcal mol^{-1}, corresponding to the above-mentioned steps, which correspond to the lowest values for the surface formation of aminomethanol. As challenging as it may be to estimate the quantity of pores-like sites in the realistic ice that has been modeled, it can certainly be considered that, when present, they strongly influence reactivity and guide molecules to encounter in the more hollow regions of the surface.

As for the second stage, the dehydration of aminomethanol, it proceeds concertedly at all 7 common catalytic sites considered and presents barriers in the range of 23-31 kcal mol^{-1}. When, instead, the reactive site is hyper-catalytic, the reaction takes place in two steps, similarly to what was observed for the previous stage. The dipolar intermediate in this case corresponds to the protonated aminomethanol: ($NH_2CH_2OH_2^+ \cdots OH^-$), while the negative fragment is located in correspondence of one of the water molecules participating to the proton relay. However, the barrier remains within the energy range found for the lone dangling-H sites. It is concluded that the second stage is the least affected by variations in the surface morphology.

The third stage involves the formation of aminoacetonitrile from HNC and methanimine. The reaction is a nucleophilic addition, like the first step, and is subjected to a similar type of environmental influence. In the com-

mon catalytic sites, the lone dangling-H is located on the side of the $NHCH_2$ species and is used to bridge the proton in the proton relay. The reaction takes place concertedly with a rather high barrier (31 kcal mol^{-1}). On the other hand, if the site is hyper-catalytic, there is also a change in the mechanism. The second dangling-H establishes a binding mode for the HNC molecule where it is strongly bound to the surface. Therefore, the reaction occurs in two stages, in which the deprotonation of HNC takes place first to form the dipolar intermediate ($NC^- \cdots NH_2CH_2^+$), followed by the proton transfer to yield the product. Koch et al. [105] had observed that stage 2 proceeds through a dipolar intermediate, but the necessary conditions for its formation and stabilization had not been thoroughly investigated. In the case of the step-wise mechanism, overcoming a barrier of 6 kcal mol^{-1} is required for the deprotonation, while the formation of the molecule's backbone is barrierless, involving the association of two fragments with opposite charges.

The fourth stage involves, in first place, the concerted hydrolysis of aminoacetonitrile and has been studied in two reactive sites of distinct catalytic efficiency. Due to the nature of the reaction and the involved species (see discussion in Section 3.4.4), the reactive processes do not exhibit catalytic effects compared to the model system, similarly to the dehydration of the second stage. In fact, in one of the cases, there is a opposite effect of increasing the barrier by 2-3 kcal mol^{-1}. As the barrier remains indiscriminately high in both cases - about 40 kcal mol^{-1} - it was not deemed relevant to study the

subsequent steps of stage four.

Comparison with previous works

Regarding the first stage, previous studies primarily focused on reactions on minimal water clusters[97,102], showing results similar to those presented in this thesis for the preliminary study that was propaedeutic to reactions on amorphous surfaces. Among works considering ice models, Rimola et al. [98] simulated the reaction on an 18-molecule crystalline water ice surface model at B3LYP/6-31+G(d,p) level, finding an activation barrier of 9.6 kcal mol^{-1} for the first stage. This thesis work likely constitutes the most comprehensive census of reactive configurations concerning the aminomethanol formation, as well as the first research on a actual amorphous ice model. The reactive configuration encountered by the author[98], upon visual inspection, seems to correspond to a common catalytic site, lone dangling-H, and indeed, their barrier aligns with the results for those specific systems, (S1-C1 and S1-C2). Works mentioned in the previous section[100,226], where the reaction proceeds on 4-12 molecules cluster models through the formation of a dipolar intermediate, exhibit barriers in the range of 5-9 kcal mol^{-1}. However, it is essential to emphasize that in their works, the total barrier of the process, from neutral species has not been calculated, i.e. they assumed the barrierless formation of the dipolar intermediate as reactant. In other words, the comparison with the results presented here, involves only barriers subsequent to the first one.

In that case, the barriers are consistent with hyper-catalytic sites (S1-C3, S1-C4), which present analog mechanism, even considering the discrepancy in the method used (B3LYP/6-31G**).

The second stage, the dehydration of aminomethanol to form methanimine, has been addressed in the aforementioned work of Rimola et al. [98] and also by Riffet et al. [100], using a PCM model at DFT/G3B3 theory level on small water clusters. In both works, despite the catalytic behavior of water, the reaction presented quite high energy barriers (21 and 23 kcal mol^{-1} respectively), which is in line with the results of this thesis for the concerted mechanism (common catalytic site). However, to the best of our knowledge, a hyper-catalyzed step-wise dehydration pathway, facilitated by the presence of the ice has not been presented thus far.

The direct path for aminoacetonitrile formation has been investigated by Koch et al. [105]. Their work considered different situations: minimal water clusters, and the presence of two water molecules acting as proton transfer assistants, plus 12 water molecules acting as ice spectators, at B3LYP/6-31+G(d,p) level of theory. The energy barrier has been found to decrease from 10 kcal mol^{-1} (reaction on water dimer embedded in the PCM solvation) to only 1 kcal mol^{-1} (considering explicit 12-molecules solvation). Rimola et al. [98] simulated the reaction on 18-molecules ice crystalline model, obtaining a higher barrier of 20 kcal mol^{-1}. This thesis results fit into the framework of the previous studies. Specifically, the concerted mechanism

pathway corresponds to the common catalytic reactive site identified in the latter work[98], and the barrier found in this thesis is slightly higher, which is attributed to the different type of surface model and the level of theory. Regarding the intriguing result by Koch et al. [105], it is worth noticing that the 2 + 12-molecule solvation model has been performed *ad hoc*, in order to selectively stabilize the transition state. Therefore, the reactive situation is more akin to carrying out the reaction in the liquid phase.

Finally, the last stage has been simulated by Rimola et al. [98], assuming that the two water molecules participating in the hydrolysis derive from evaporation processes occurring on the ice surface. The first water nucleophilic attack towards the nitrile gives the glycine amide, and displayed a high energy barrier of 38 kcal mol^{-1}. This is the only step of stage 4 that has been explored in this thesis on ASW, and the results align with their, emphasizing a limited catalytic role of the ice in the process.

Astrophysical implications

The reactivity results presented in this thesis lead to interesting consequences in terms of astrophysical implications.

As mentioned above, hyper- and nano-pores catalytic sites on amorphous ice mantles appear to convert concerted water-assisted reaction mechanisms to step-wise. Under these specific circumstances, each step of a proton relay corresponds to a isolated event. As a consequence, it might benefit from

quantum tunnelling effects, allowing reactions involving hydrogen atoms to occur faster than expected from transition state theory. Tunneling effects associated with transfer and abstraction reactions involving hydrogen, have been extensively studied[3,227,228] and are known to play a significant role in the ISM. However, tunneling effects are heavily dependent on the width of the energy barrier. Analysis of the IRC profiles revealed that, in case of isolated or almost isolated proton transfer step (e.g. S1-C2), the top region of the IRC curve gets particularly narrow, compared to the profile of a reaction that involves the entire set of proton relay steps (e.g. S1-C1), opening to the possibility that the reaction rates might actually be enhanced by tunnelling effects, which could be crucial at extremely low temperature (< 10K).

This result is a further evidence in support of one of the main conclusions of this thesis: the aminomethanol, product of the first stage of the synthesis, can be formed through a Strecker-type reaction under interstellar dense clouds condition. The hypothesis that the species is present on the ice, and has long residence time is especially valid, in light of the estimated range of binding energies provided in this thesis (2-15 kcal mol^{-1}). The binding energy distribution of aminomethanol exhibits a rather high average value (9 kcal mol^{-1}) compared to most molecules studied, which can be attributed to the numerous non-covalent interactions that the molecule can establish with the ice mantle; thus, the fact that the aminomethanol has not been observed in the interstellar medium yet, might be due to identification issues

(see discussion in Section 1.1.4.1) or to the possibility that the molecule undergoes destruction on the surface, before desorbing. However, according to the results of this study, dehydration is not among them: the energy barrier is prohibitively high for most interstellar environments. Furthermore, although aminomethanol dehydration, as well as the majority of the other stages of the synthesis, undergoes alterations in the mechanism due to the effect of surface morphology, it does not experience the usual reduction in the extent of the energy barrier. A consequence of the latter is that alternative formation mechanisms for methanimine must be considered, ruling out the Strecker-type reaction. Various pathways have been proposed, including surface reactions involving the hydrogenation of HCN[75]. However, it has been suggested[229] that such a scenario would lead to the formation of NH_2CH_3 as the main product, i.e. the saturated species that concludes the chain of hydrogenation steps, instead of $NHCH_2$. In support of it, corresponding experiments on ice-analogs did not detect methanimine among the products, suggesting that it is a short-lived intermediate. More credited pathways are gas-phase reactions, including ion-molecule or neutral-neutral reactions. The latter seem more probable[230] and also align with the fact that the species has not been observed in dark cloud regions but only in warm star-forming regions[112], where neutral-neutral gas-phase chemistry can occur.

With respect to the various interstellar regions relevant to the synthesis of prebiotic species, it is worth mentioning that some of the species have been

detected only in warm regions near hot cores. Higher temperatures have various effects on reactivity, among these, facilitating surface reactions not energetically favored in extreme conditions, such as the concerted mechanism variation of stages 1 and 3 (energy barrier range 10-30 kcal mol^{-1}). Additionally, in such environments, an increase in the mobility of the species on the ice is expected. This would allow the encounter of fragments that were originally situated in distant sites with respect to each other. Or else, originally located within lower layers of the ice mantle, and subsequent to the evaporation of part of it, are made available to react.

Diffusion is often considered the limiting step in interstellar models of surface chemistry. As mentioned above, reactions in correspondence of pair dangling-H sites, or hyper-catalytic sites, require the presence of one or both species in high binding energy situations. The temperature increase is also thought to favor a change in the type of binding modes[65], meaning that, with an increase in temperature - not sufficient for desorption - the molecules migrate from low BE sites to high BE sites. This could promote, in such warmer environments, the reaction in a hyper-catalytic arrangement of the reactants (e.g. S1-C3, S1-C4 for stage 1). In terms of morphology, the situation corresponds of the species to be in (or to had moved to) the so-called 'valleys' regions of the ice mantle.

On the other hand, reactions taking place around lone dangling-H sites, or common catalytic sites, among molecules weakly bound to the surface

(e.g. S1-C1, S1-C2 for stage 1), would still be more likely to take place in active star-forming regions - than in quiescent dark clouds - since the barriers, although higher, can be more easily overcome. Such situations represent the reaction between loosely bound species that encounter and react in the time-span before reaching sites of higher binding energy, or before desorbing. Topologically, these molecules are originally located on peaks, or 'crests' of the ice mantles. Such topological considerations are suggested by the binding enery results obtained with the realistic ice model, which provides large surface available for absorption, with a greater variety of scenarios. Indeed, molecules on the crests are shown to exhibit lower binding energy than those in the valleys.

As for colder regions, such as dark clouds, Strecker-like neutral-neutral surface reactions are disadvantaged, in favor of non-energetic pathways[96], involving radical species. Nevertheless, there is the possibility that the processes increase their feasibility due to catastrophic events, such as violent impacts of cosmic rays. These events might cause abrupt kinetic energy increase in the impact vicinity, potentially allowing the systems to overcoming barriers comparable to those calculated for stages 2 and 4 (energy range: 30-45 kcal mol^{-1}).

Moreover, changes in the morphology of the ice following cosmic ray bombardment could release species trapped in lower levels of the ice mantle. Catastrophic events could also open up possibilities for new formation

mechanisms, for example regarding $NHCH_2$ species. Recently, it has been proposed that its formation occurs following the impact of an excited-state carbon atoms on top of ammonia molecules adsorbed on the surface, causing atom-insertion. The recombination of the reactive complex, powered by means of external sources of energy, such as cosmic rays, yields methanimine.

However, the possibility that the entire set of reactions associated with a Strecker-type synthesis of glycine occurs following one or more catastrophic events might be unrealistic, as the stages are numerous and contain somewhat intricate steps. Besides, competitive pathways for glycine formation[88,92,96], which present simpler mechanisms, would also receive a energy contribution from catastrophic events, likely emerging as the favored ones. This, in any case, does not rule out the possibility that a certain stage of the Strecker synthesis takes place, individually. One last consideration regarding the chemistry under extreme conditions of quiescent clouds is needed. This thesis study has brought attention to the fact that, for the vast majority of the reactions, there exists a step-wise variant that takes place in hyper-catalytic sites or nano-pores, involving the formation of dipolar intermediates. Once the intermediate is isolated, the reaction is basically the coupling between charged fragments, thus exhibiting either none or minimal barriers. It cannot be excluded that, in dark clouds, suitable charged species may form through alternative pathways - i.e. not originating from neutral species - making them

available for the reaction. For example, due to the secondary ionization field caused by cosmic rays. This induced-UV field is the only type of radiation present in dark clouds, as the high-density conditions of the environment shield it from the radiation common in other regions of the interstellar medium. Therefore, a surface chemistry of ion-ion type could be an alternative for the production of such prebiotic species. Obviously, this hypothesis should consider the adsorption and mobility of charged species - the ease of encountering each other - their lifetime on the surface, and whether that is within the overall dust neutralization time.

Lastly, the study reinforce the idea that, in peculiar situations such as in comets or meteorites precursors, this kind of Strecker-type reactions is viable. These environments include the surface of certain comets that exhibit a scenario composed of patches of ices and exposed refractory material, resulting in generally higher temperatures ($\sim$200 K). Studies of surface models consisting of silica surfaces with traces of watery ices have been recently proposed[231]. Or else, in aqueous alterations of meteorites interstellar-parent bodies (asteroids, comets, or moons) that may have undergone transformations, causing partial melting of water ice of the future meteorite fragments. This allowed for prebiotic synthesis to occur in the lower ice layers. Such processes seem to be at the foundation of the origin of life in the Solar System, as amino acids resulting from these reactions could have been preserved within the meteorites, shielded from radiation, until reaching Earth.

4. Conclusions

The objective of this thesis was to create a robust framework to accurately determine parameters — specifically, binding energies and transition state barriers — that are key to explain the formation and observation of prebiotic species in the interstellar medium. Binding energies are critical as they determine the desorption rate of small molecules from the ice mantle surface, while the transition state barriers are pivotal for calculating the reaction rates between adjacent molecules on the surface. The research aimed to investigate the impact of an amorphous solid water (ASW) surface on these parameters by constructing accurate models of ASW ice and computing these properties with the highest possible precision using methods from electronic structure theory.

Different models of ASW ice were proposed. A set of amorphized water clusters was derived from *ab initio* molecular dynamics simulations and subsequently cooled to 20 K, to mimic surfaces akin to those in dense interstellar clouds. A periodic ASW surface containing 500-water molecules was built using a purposed trained Machine-Learned Potential. The models were characterized in terms of dangling-hydrogen atoms, as they constitute important catalytic sites, where adsorption and reaction preferentially occur.

Considering the computation of binding energies, it was paramount to build a framework to automatize the computation of binding energies on the set-of-clusters ASW models in order to apply it to a large number of molecules of interstellar relevance. Therefore a Binding Energy Evaluation Platform (BEEP) was constructed to implement a protocol to compute binding energies on ASW cluster models using DFT. This protocol included systematic sampling of binding sites on the water clusters and BSSE-corrected binding energy calculations. Furthermore, in order to obtain highly accurate binding energy values, the DFT model chemistries were carefully benchmarked on small water clusters with respect to a CCSD(T)/CBS reference value. It was found that, due to the amorphous nature of the surface, the binding energy is best described by a distribution of values that can be computed using BEEP. The molecules were categorized into two groups, based on their type of interaction with the surface: molecules that are bound primarily through hydrogen-bonds (Group H) and molecules for which dispersion interactions enable binding to the surface (Group D). 21 binding energy distributions of astrophysically relevant molecules were computed. Each distribution contains between 220-230 binding sites. Most molecules in Group H present two distributions, corresponding to different binding modes, while Group D molecules mainly have one. The astrophysical impact of the new binding energy values was analyzed and it was determined that the position of the

snowline of an idealized protoplanetary disk change significantly when using a binding energy distribution instead of single value. Such result suggests that the new multi-binding approach might help in determining the correct position of the sublimation front, and also hints at the importance that providing accurate binding energies might have on key astrophysical observables. Finally, various binding modes of the reactants involved in the Strecker synthesis of glycine were examined, and it was found that not all binding modes facilitate a reactive encounter necessary for the formation of prebiotic species associated with that synthesis.

Lastly, the new multi-binding approach allowed to disclose the effect of the ice morphology on specific reactions mechanism, clarifying its catalytic role. The study took into account reactions associated to the Strecker synthesis of glycine. Each of them has been simulated in a varied set of reactive sites with different characteristics, as offered by realistic ASW ice models. The results indicate that, depending on the nature of the catalytic site, significant variations of the reaction mechanisms occur. Especially, when the processes takes place in so-called pair dangling hydrogen sites or in the interior of nano-porous, many water-assisted reactions convert from concerted to step-wise. That is due to the extraordinary polarization of the reactants caused by the ice environment, which permits the isolation of dipolar intermediates. Under these specific circumstances, each step of a water-assisted

proton transfer corresponds to an isolated event. The latter entails notable advantages: in many cases the step-wise variation presents remarkably lower energy barrier with respect to the concerted one. Moreover, as the reactions involve hydrogen, isolated proton transfer events might benefit from quantum tunnelling effects, providing a viable route to formation at extremely low temperature ($<$10 K).

Regarding reactivity on ASW, a primary result is that a viable route exists, under interstellar conditions, for the first stage of Strecker's synthesis, the nucleophilic addition of ammonia and formaldehyde yielding the product aminomethanol. Moreover, the process of aminomethanol dehydration, constituting the second stage of synthesis, exhibits a prohibitively high energy barrier across all reactive sites. This result, coupled with the analysis of the binding energy values calculated for the species, further strengthens the hypothesis in favor of aminomethanol to be present on the surface of interstellar ice mantles. Also, as a consequence, alternative formation mechanisms to the Strecker synthesis for the observed dehydration product (methanimine) must be taken into account.

The third stage of the synthesis involves the formation of aminoacetonitrile from the nucleophilic addition of hydrogen isocyanide to methanimine. The reaction mechanism is sensitive to changes in the surface environment, and benefits from the presence of ice HB donor groups, which provide stabiliza-

tion to the reactive site, similarly to the findings for first stage. Therefore, it represents a viable synthetic route for aminoacetonitrile, assuming a different source for methanimine. Finally, such considerations are not applicable to the last stage of the synthesis, the hydrolysis of the aminoacetonitrile to attain glycine, which comprises a sequence of steps with high transition state barriers.

The results of the study for the second and the last stages lead to the conclusion that the Strecker synthesis as a pathway for amino acids formation in the coldest regions of ISM, has to be ruled out. Nevertheless, there is the possibility that individual Strecker-type reactions increase their feasibility due to catastrophic events, such as impacts of cosmic rays, which might provide a local energy surplus needed to overcome large energy barriers.

In conclusion, within this thesis, a novel approach is introduced to obtain parameters, which are paramount to understand processes occurring on interstellar ices, such as reactivity and desorption of species, thus contributing valuable insights of possible formation pathways of prebiotic species in the interstellar medium.

Bibliography

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

(2) *Encyclopedia of Astrobiology*; Gargaud, M., Amils, R., Quintanilla, J. C., Cleaves, H. J., Irvine, W. M., Pinti, D. L., Viso, M., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011.

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

(4) Pontoppidan, K. OBSERVATIONS OF HYDRIDES IN DISKS. Space Telescope Science Institute, `https://hydride-toolbox.sciencesconf.org/data/program/pontoppidan_hydrides.pdf`.

(5) Tielens, A. G. G. M. Interstellar Depletions and the Life Cycle of Interstellar Dust. *The Astrophysical Journal* **1998**, *499*, 267–272.

(6) Bartels-Rausch, T. et al. Ice structures, patterns, and processes: A view across the ice-fields. *Reviews of Modern Physics* **2012**, *84*, 885–944.

(7) Maggiolo, R.; Gibbons, A.; Cessateur, G.; Keyser, J. D.; Dhooghe, F.; Gunell, H.; Loreau, J.; Mousis, O.; Vaeck, N. Effect of the Surface

Roughness of Icy Grains on Molecular Oxygen Chemistry in Molecular Clouds. *The Astrophysical Journal* **2019**, *882*, 131.

(8) Palumbo, M. E. Formation of compact solid water after ion irradiation at 15 K. *Astronomy & Astrophysics* **2006**, *453*, 903–909.

(9) Boogert, A. A.; Gerakines, P. A.; Whittet, D. C. Observations of the Icy Universe. *Annual Review of Astronomy and Astrophysics* **2015**, *53*, 541–581.

(10) Whittet, D. C. B.; Poteet, C. A.; Chiar, J. E.; Pagani, L.; Bajaj, V. M.; Horne, D.; Shenoy, S. S.; Adamson, A. J. ICE AND DUST IN THE PRESLAR DARK CLOUD LYNDS 183: PREPLANETARY MATTER AT THE LOWEST TEMPERATURES. *The Astrophysical Journal* **2013**, *774*, 102.

(11) Öberg, K. I.; Boogert, A. C. A.; Pontoppidan, K. M.; Broek, S. v. d.; Dishoeck, E. F. v.; Bottinelli, S.; Blake, G. A.; Evans, N. J. THES-PITZERICE LEGACY: ICE EVOLUTION FROM CORES TO PROTOSTARS. *The Astrophysical Journal* **2011**, *740*, 109.

(12) Gibb, E. L.; Whittet, D. C. B.; Schutte, W. A.; Boogert, A. C. A.; Chiar, J. E.; Ehrenfreund, P.; Gerakines, P. A.; Keane, J. V.; Tielens, A. G. G. M.; Dishoeck, E. F. v.; Kerkhof, O. An Inventory of Interstellar Ices toward the Embedded Protostar W33A*. *The Astrophysical Journal* **2000**, *536*, 347.

(13) McClure, M. K. et al. An Ice Age JWST inventory of dense molecular cloud ices. *Nature Astronomy* **2023**, *7*, 431–443.

(14) Steinhardt, C. L.; Jermyn, A. S.; Lodman, J. Thermal Regulation and the Star-forming Main Sequence. *The Astrophysical Journal* **2020**, *890*, 19.

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

(16) Boogert, A. C. A. et al. The c2d *Spitzer* Spectroscopic Survey of Ices around Low-Mass Young Stellar Objects. I. H $_2$ O and the 5–8 m Bands1,2. *The Astrophysical Journal* **2008**, *678*, 985–1004.

(17) Aikawa, Y.; Kamuro, D.; Sakon, I.; Itoh, Y.; Terada, H.; Noble, J. A.; Pontoppidan, K. M.; Fraser, H. J.; Tamura, M.; Kandori, R.; Kawamura, A.; Ueno, M. AKARI observations of ice absorption bands towards edge-on young stellar objects. *Astronomy & Astrophysics* **2012**, *538*, A57.

(18) Dishoeck, E. F. v. et al. Water in star-forming regions: physics and chemistry from clouds to disks as probed by Herschel spectroscopy. *Astronomy & Astrophysics* **2021**, *648*, A24.

(19) Charnley, S. B.; Tielens, A. G. G. M.; Millar, T. J. On the molecular complexity of the hot cores in Orion A - Grain surface chemistry as

'The last refuge of the scoundrel'. *The Astrophysical Journal* **1992**, *399*, L71.

(20) Caselli, P.; Hasegawa, T. I.; Herbst, E. Chemical differentiation between star-forming regions - The Orion Hot Core and Compact Ridge. *The Astrophysical Journal* **1993**, *408*, 548.

(21) Boss, A. P. TEMPERATURES IN PROTOPLANETARY DISKS. *Annual Review of Earth and Planetary Sciences* **1998**, *26*, 53–80.

(22) Dullemond, C. P.; Isella, A.; Andrews, S. M.; Skobleva, I.; Dzyurkevich, N. Midplane temperature and outer edge of the protoplanetary disk around HD 163296. *Astronomy & Astrophysics* **2020**, *633*, A137.

(23) Pontoppidan, K. M.; van Dishoeck, E. F.; Dartois, E. Mapping ices in protostellar environments on 1000 AU scales. Methanol-rich ice in the envelope of Serpens SMM 4. *Astronomy and Astrophysics* **2004**, *426*, 925–940.

(24) Whipple, F. L. A comet model. I. The acceleration of Comet Encke. *The Astrophysical Journal* **1950**, *111*, 375.

(25) Bar-nun, A.; Kleinfeld, I. On the temperature and gas composition in the region of comet formation. *Icarus* **1989**, *80*, 243–253.

(26) Agency, T. E. S. Rosetta takes comet's temperature, `https://www.esa.int/Science_Exploration/Space_Science/Rosetta/Rosetta_takes_comet_s_temperature`.

(27) Cobb, A. K.; Pudritz, R. E. NATURE'S STARSHIPS. I. OBSERVED ABUNDANCES AND RELATIVE FREQUENCIES OF AMINO ACIDS IN METEORITES. *The Astrophysical Journal* **2014**, *783*, 140.

(28) Altwegg, K. et al. Prebiotic chemicals—amino acid and phosphorus—in the coma of comet 67P/Churyumov-Gerasimenko. *Science Advances* **2016**, *2*, e1600285.

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

(30) Jaccard, C. P. V. Hobbs Ice physics. Oxford, Clarendon Press, 1974. xvii, 837 p. £29. *Journal of Glaciology* **1976**, *17*, 155–156.

(31) Al-Halabi, A.; Fraser, H. J.; Kroes, G. J.; Van Dishoeck, E. F. Adsorption of CO on amorphous water-ice surfaces. *Astronomy & Astrophysics* **2004**, *422*, 777–791.

(32) Watanabe, N.; Kouchi, A. Ice surface reactions: A key to chemical evolution in space. *Progress in Surface Science* **2008**, *83*, 439–489.

(33) Hama, T.; Watanabe, N. Surface Processes on Interstellar Amorphous Solid Water: Adsorption, Diffusion, Tunneling Reactions, and Nuclear-Spin Conversion. *Chemical Reviews* **2013**, *113*, 8783–8839.

(34) He, J.; Frank, P.; Vidali, G. Interaction of hydrogen with surfaces of silicates: single crystal vs. amorphous. *Physical Chemistry Chemical Physics* **2011**, *13*, 15803–15809.

(35) Noble, J. A.; Congiu, E.; Dulieu, F.; Fraser, H. J. Thermal desorption characteristics of CO, O2 and CO2 on non-porous water, crystalline water and silicate surfaces at submonolayer and multilayer coverages. *Monthly Notices of the Royal Astronomical Society* **2012**, *421*, 768–779.

(36) Bovolenta, G.; Bovino, S.; Vöhringer-Martinez, E.; Saez, D. A.; Grassi, T.; Vogt-Geisse, S. High level ab initio binding energy distribution of molecules on interstellar ices: Hydrogen fluoride. *Molecular Astrophysics* **2020**, 100095.

(37) Das, A.; Sil, M.; Gorai, P.; Chakrabarti, S. K.; Loison, J.-C. An Approach to Estimate the Binding energy of Interstellar Species. *The Astrophysical Journal Supplement Series* **2018**, *237*, 9.

(38) Wakelam, V.; Loison, J. C.; Mereau, R.; Ruaud, M. Binding energies: New values and impact on the efficiency of chemical desorption. *Molecular Astrophysics* **2017**, *6*, 22–35.

(39) Molpeceres, G.; Kästner, J. Adsorption of H 2 on amorphous solid water studied with molecular dynamics simulations. *Physical Chemistry Chemical Physics* **2020**, *22*, 7552–7563.

(40) Enrique-Romero, J.; Rimola, A.; Ceccarelli, C.; Ugliengo, P.; Balucani, N.; Skouteris, D. Reactivity of HCO with CH3 and NH2 on Water Ice Surfaces. A Comprehensive Accurate Quantum Chemistry Study. *ACS Earth and Space Chemistry* **2019**, *3*, 2158–2170.

(41) Ferrero, S.; Zamirri, L.; Ceccarelli, C.; Witzel, A.; Rimola, A.; Ugliengo, P. Binding Energies of Interstellar Molecules on Crystalline and Amorphous Models of Water Ice by Ab Initio Calculations. *The Astrophysical Journal* **2020**, *904*, 11.

(42) Perrero, J.; Enrique-Romero, J.; Ferrero, S.; Ceccarelli, C.; Podio, L.; Codella, C.; Rimola, A.; Ugliengo, P. Binding Energies of Interstellar Relevant S-bearing Species on Water Ice Mantles: A Quantum Mechanical Investigation. *The Astrophysical Journal* **2022**, *938*, 158.

(43) Shimonishi, T.; Nakatani, N.; Furuya, K.; Hama, T. Adsorption Energies of Carbon, Nitrogen, and Oxygen Atoms on the Low-temperature Amorphous Water Ice: A Systematic Estimation from Quantum Chemistry Calculations. *The Astrophysical Journal* **2018**, *855*, 27.

(44) Sil, M.; Gorai, P.; Das, A.; Sahu, D.; Chakrabarti, S. K. Adsorption energies of H and H2: a quantum-chemical study. *The European Physical Journal D* **2017**, *71*, 45.

(45) Gould, R. J.; Salpeter, E. E. The Interstellar Abundance of the Hydrogen Molecule. I. Basic Processes. *The Astrophysical Journal* **1963**, *138*, 393.

(46) Watson, W. D.; Salpeter, E. E. Molecule Formation on Interstellar Grains. *The Astrophysical Journal* **1972**, *174*, 321.

(47) Molpeceres, G.; Zaverkin, V.; Furuya, K.; Aikawa, Y.; Kästner, J. Reaction dynamics on amorphous solid water surfaces using interatomic machine-learned potentials: Microscopic energy partition revealed from the P + H $\rightarrow$ PH reaction. *Astronomy & Astrophysics* **2023**, *673*, A51.

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

(49) Tsuge, M.; Molpeceres, G.; Aikawa, Y.; Watanabe, N. Surface Diffusion of Carbon Atoms as a Driver of Interstellar Organic Chemistry. *Nature Astronomy* **2023**.

(50) Tielens, A. G. G. M.; Hagen, W.; Natuurwetenschappen, F. d. W. e. Model calculations of the molecular composition of interstellar grain mantles, Letter to editor, 1982.

(51) Hasegawa, T. I.; Herbst, E. Three-Phase Chemical Models of Dense Interstellar Clouds - Gas Dust Particle Mantles and Dust Particle Surfaces. *Monthly Notices of the Royal Astronomical Society* **1993**, *263*, 589.

(52) Garrod, R. T. A THREE-PHASE CHEMICAL MODEL OF HOT CORES: THE FORMATION OF GLYCINE. *The Astrophysical Journal* **2013**, *765*, 60.

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

(54) McElroy, D.; Walsh, C.; Markwick, A. J.; Cordiner, M. A.; Smith, K.; Millar, T. J. The UMIST database for astrochemistry 2012. *Astronomy & Astrophysics* **2013**, *550*, A36.

(55) Grassi, T.; Bovino, S.; Schleicher, D. R. G.; Prieto, J.; Seifried, D.; Simoncini, E.; Gianturco, F. A. KROME - a package to embed chemistry in astrophysical simulations. *Monthly Notices of the Royal Astronomical Society* **2014**, *439*, 2386–2419.

(56) Minissale, M. Physics and chemistry on the surface of interstellar dust grains: the effect of O-atom diffusion and chemical desorption on the H-C-N-O reaction network. *Université de Cergy Pontoise; Università degli studi (Catane, Italie)* **2014**, 165.

(57) He, J.; Acharyya, K.; Vidali, G. BINDING ENERGY OF MOLECULES ON WATER ICE: LABORATORY MEASUREMENTS AND MODELING. *The Astrophysical Journal* **2016**, *825*, 89.

(58) Schlemmer, S.; Illemann, J.; Wellert, S.; Gerlich, D. Nondestructive high-resolution and absolute mass determination of single charged particles in a three-dimensional quadrupole trap. *Journal of Applied Physics* **2001**, *90*, 5410–5418.

(59) Penteado, E. M.; Walsh, C.; Cuppen, H. M. Sensitivity Analysis of Grain Surface Chemistry to Binding Energies of Ice Species. *The Astrophysical Journal* **2017**, *844*, 71.

(60) Song, L.; Kästner, J. Formation of the prebiotic molecule NH2CHO on astronomical amorphous solid water surfaces: accurate tunneling rate calculations. *Physical Chemistry Chemical Physics* **2016**, *18*, 29278–29285.

(61) Lamberts, T.; Markmeyer, M. N.; Kolb, F. J.; Kästner, J. Formation of Acetaldehyde on CO-Rich Ices. *ACS Earth and Space Chemistry* **2019**, *3*, 958–963.

(62) Duflot, D.; Toubin, C.; Monnerville, M. Theoretical Determination of Binding Energies of Small Molecules on Interstellar Ice Surfaces. *Frontiers in Astronomy and Space Sciences* **2021**, *8*.

(63) Sameera, W. M. C.; Senevirathne, B.; Andersson, S.; Al-lbadi, M.; Hidaka, H.; Kouchi, A.; Nyman, G.; Watanabe, N. CH3O Radical Binding on Hexagonal Water Ice and Amorphous Solid Water. *The Journal of Physical Chemistry A* **2021**, *125*, 387–393.

(64) Watanabe, N.; Kimura, Y.; Kouchi, A.; Chigai, T.; Hama, T.; Pirronello, V. DIRECT MEASUREMENTS OF HYDROGEN ATOM DIFFUSION AND THE SPIN TEMPERATURE OF NASCENT H 2 MOLECULE ON AMORPHOUS SOLID WATER. *The Astrophysical Journal* **2010**, *714*, L233–L237.

(65) Grassi, T.; Bovino, S.; Caselli, P.; Bovolenta, G.; Vogt-Geisse, S.; Ercolano, B. A novel framework for studying the impact of binding energy distributions on the chemistry of dust grains. *Astronomy & Astrophysics* **2020**, *643*, A155.

(66) Choughuley, A. S. U.; Subbaraman, A. S.; Kazi, Z. A.; Chadha, M. S. Transformation of some hydroxy amino acids to other amino acids. *Origins of life* **1975**, *6*, 527–535.

(67) Oba, Y.; Takano, Y.; Naraoka, H.; Furukawa, Y.; Glavin, D. P.; Dworkin, J. P.; Tachibana, S. Extraterrestrial hexamethylenetetramine in meteorites—a precursor of prebiotic chemistry in the inner solar system. *Nature Communications* **2020**, *11*, 6243.

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

(69) Kuan, Y.-J.; Charnley, S. B.; Huang, H.-C.; Tseng, W.-L.; Kisiel, Z. Interstellar Glycine. *The Astrophysical Journal* **2003**, *593*, 848.

(70) Snyder, L. E.; Lovas, F. J.; Hollis, J. M.; Friedel, D. N.; Jewell, P. R.; Remijan, A.; Ilyushin, V. V.; Alekseev, E. A.; Dyubko, S. F. A Rigorous Attempt to Verify Interstellar Glycine. *The Astrophysical Journal* **2005**, *619*, 914.

(71) Jones, P. A.; Cunningham, M. R.; Godfrey, P. D.; Cragg, D. M. A Search for biomolecules in Sagittarius B2 (LMH) with the Australia Telescope Compact Array. *Monthly Notices of the Royal Astronomical Society* **2007**, *374*, 579–589.

(72) Cunningham, M. R. et al. A search for propylene oxide and glycine in Sagittarius B2 (LMH) and Orion. *Monthly Notices of the Royal Astronomical Society* **2007**, *376*, 1201–1210.

(73) Lattelais, M.; Pauzat, F.; Pilmé, J.; Ellinger, Y.; Ceccarelli, C. About the detectability of glycine in the interstellar medium. *Astronomy & Astrophysics* **2011**, *532*, A39.

(74) *Laboratory Astrochemistry: From Molecules through Nanoparticles to Grains*; Schlemmer, S., Mutschke, H., Giesen, T., Jäger, C., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2014.

(75) Woon, D. E. Ab initio quantum chemical studies of reactions in astrophysical ices. 4. Reactions in ices involving HCOOH, CH2NH, HCN, HNC, NH3, and H2O. *International Journal of Quantum Chemistry* **2002**, *88*, 226–235.

(76) Pilling, S.; Baptista, L.; Boechat-Roberty, H. M.; Andrade, D. P. P. Formation Routes of Interstellar Glycine Involving Carboxylic Acids: Possible Favoritism Between Gas and Solid Phase. *Astrobiology* **2011**, *11*, 883–893.

(77) Sato, A.; Kitazawa, Y.; Ochi, T.; Shoji, M.; Komatsu, Y.; Kayanuma, M.; Aikawa, Y.; Umemura, M.; Shigeta, Y. First-principles study of the formation of glycine-producing radicals from common interstellar species. *Molecular Astrophysics* **2018**, *10*, 11–19.

(78) Largo, L.; Redondo, P.; Rayón, V. M.; Largo, A.; Barrientos, C. The reaction between NH_3^+ and CH_3COOH: a possible process for the formation of glycine precursors in the interstellar medium. *Astronomy and Astrophysics* **2010**, *516*, A79.

(79) Shivani; Singh, A.; Gupta, V.; Misra, A.; Tandon, P. Quantum-chemical approach to serine formation in the interstellar medium: A possible reaction pathway. *Astronomy & Astrophysics* **2014**, *563*, A55.

(80) Jeanvoine, Y.; Largo, A.; Hase, W. L.; Spezia, R. Gas Phase Synthesis of Protonated Glycine by Chemical Dynamics Simulations. *The Journal of Physical Chemistry A* **2018**, *122*, 869–877.

(81) Sanz-Novo, M.; Largo, A.; Redondo, P.; Barrientos, C. Formation of Protonated Glycine Isomers in the Interstellar Medium. *ACS Earth and Space Chemistry* **2019**, *3*, 1170–1181.

(82) Strecker, A. Ueber die künstliche Bildung der Milchsäure und einen neuen, dem Glycocoll homologen Körper; *Justus Liebigs Annalen der Chemie* **1850**, *75*, 27–45.

(83) Biver, N.; Bockelée-Morvan, D.; Boissier, J.; Moreno, R.; Crovisier, J.; Lis, D. C.; Colom, P.; Cordiner, M. A.; Milam, S. N.; Roth, N. X.; Bonev, B. P.; Dello Russo, N.; Vervack, R. J.; DiSanti, M. A. Molecular composition of comet 46P/Wirtanen from millimetre-wave spectroscopy. *Astronomy & Astrophysics* **2021**, *648*, A49.

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

(85) Bernstein, M. P.; Dworkin, J. P.; Sandford, S. A.; Cooper, G. W.; Allamandola, L. J. Racemic amino acids from the ultraviolet photolysis of interstellar ice analogues. *Nature* **2002**, *416*, 401–403.

(86) Danger, G.; Borget, F.; Chomat, M.; Duvernay, F.; Theulé, P.; Guillemin, J.-C.; d'Hendecourt, L. L. S.; Chiavassa, T. Experimental investigation of aminoacetonitrile formation through the Strecker synthesis in astrophysical-like conditions: reactivity of methanimine (CH2NH), ammonia (NH3), and hydrogen cyanide (HCN). *Astronomy & Astrophysics* **2011**, *535*, A47.

(87) Borget, F.; Duvernay, F.; Danger, G.; Theulé, P.; Vinogradoff, V.; Mispelaer, F.; Müller, S.; Grote, D.; Chiavassa, T.; Bossa, J.-B. What are the intermediates that could react in the interstellar ices? *Journal of Physical Organic Chemistry* **2015**, *28*, 163–169.

(88) Nhlabatsi, Z. P.; Bhasi, P.; Sitha, S. Possible interstellar formation of glycine from the reaction of CH2NH, CO and H2O: catalysis by extra water molecules through the hydrogen relay transport. *Physical Chemistry Chemical Physics* **2015**, *18*, 375–381.

(89) Nhlabatsi, Z. P.; Bhasi, P.; Sitha, S. Possible interstellar formation of glycine through a concerted mechanism: a computational study on the reaction of CH2NH, CO2 and H2. *Physical Chemistry Chemical Physics* **2016**, *18*, 20109–20117.

(90) Lee, H. M.; Choe, J. C. Formation of glycine from HCN and H2O: A computational mechanistic study. *Chemical Physics Letters* **2017**, *675*, 6–10.

(91) Kayanuma, M.; Kidachi, K.; Shoji, M.; Komatsu, Y.; Sato, A.; Shigeta, Y.; Aikawa, Y.; Umemura, M. A theoretical study of the formation of glycine via hydantoin intermediate in outer space environment. *Chemical Physics Letters* **2017**, *687*, 178–183.

(92) Krasnokutski, S. A.; Jäger, C.; Henning, T. Condensation of Atomic Carbon: Possible Routes toward Glycine. *The Astrophysical Journal* **2020**, *889*, 67.

(93) Krasnokutski, S. A. Did life originate from low-temperature areas of the Universe? *Low Temperature Physics* **2021**, *47*, 199–205.

(94) Krasnokutski, S. A.; Chuang, K.-J.; Jäger, C.; Ueberschaar, N.; Henning, T. A pathway to peptides in space through the condensation of atomic carbon. *Nature Astronomy* **2022**, 1–6.

(95) Molpeceres, G.; Kästner, J.; Fedoseev, G.; Qasim, D.; Schömig, R.; Linnartz, H.; Lamberts, T. Carbon Atom Reactivity with Amorphous Solid Water: H2O-Catalyzed Formation of H2CO. *The Journal of Physical Chemistry Letters* **2021**, *12*, 10854–10860.

(96) Ioppolo, S.; Fedoseev, G.; Chuang, K.-J.; Cuppen, H. M.; Clements, A. R.; Jin, M.; Garrod, R. T.; Qasim, D.; Kofman, V.; van Dishoeck,

E. F.; Linnartz, H. A non-energetic mechanism for glycine formation in the interstellar medium. *Nature Astronomy* **2021**, *5*, 197–205.

(97) Woon, D. E. Ab Initio Quantum Chemical Studies of Reactions in Astrophysical Ices: 1. Aminolysis, Hydrolysis, and Polymerization in H2CO/NH3/H2O Ices. *Icarus* **1999**, *142*, 550–556.

(98) Rimola, A.; Sodupe, M.; Ugliengo, P. Deep-space glycine formation via Strecker-type reactions activated by ice water dust mantles. A computational approach. *Physical Chemistry Chemical Physics* **2010**, *12*, 5285.

(99) Ćmikiewicz, A.; Gordon, A. J.; Berski, S. Characterisation of the reaction mechanism between ammonia and formaldehyde from the topological analysis of ELF and catastrophe theory perspective. *Structural Chemistry* **2018**, *29*, 243–255.

(100) Riffet, V.; Frison, G.; Bouchoux, G. Quantum-Chemical Modeling of the First Steps of the Strecker Synthesis: From the Gas-Phase to Water Solvation. *The Journal of Physical Chemistry A* **2018**, *122*, 1643–1657.

(101) Ali, M. A. Theoretical study on the gas phase reaction of $CH_2O + NH_3$: the formation of CH_2ONH_3 , NH_2CH_2OH, or $CH_2NH + H_2O$. *Physical Chemistry Chemical Physics* **2019**, *21*, 19242–19251.

(102) Courmier, D.; Gardebien, F.; Minot, C.; St-Amant, A. A computational study of the water-catalyzed formation of NH2CH2OH. *Chemical Physics Letters* **2005**, *405*, 357–363.

(103) Chen, L.; Woon, D. E. A Theoretical Investigation of the Plausibility of Reactions between Ammonia and Carbonyl Species (Formaldehyde, Acetaldehyde, and Acetone) in Interstellar Ice Analogs at Ultracold Temperatures. *The Journal of Physical Chemistry A* **2011**, *115*, 5166–5183.

(104) Woon, D. E. Ab Initio Quantum Chemical Studies of Reactions in Astrophysical Ices 3. Reactions of HOCH $_2$ NH $_2$ Formed in H $_2$ CO/NH $_3$ /H $_2$ O Ices. *The Journal of Physical Chemistry A* **2001**, *105*, 9478–9481.

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

(106) Cheung, A. C.; Rank, D. M.; Townes, C. H.; Thornton, D. D.; Welch, W. J. Detection of N H 3 Molecules in the Interstellar Medium by Their Microwave Emission. *Physical Review Letters* **1968**, *21*, 1701–1705.

(107) Zuckerman, B.; Buhl, D.; Palmer, P.; Snyder, L. E. Observations of Interstellar Formaldehyde. *The Astrophysical Journal* **1970**, *160*, 485.

(108) Snyder, L. E.; Buhl, D. Observations of Radio Emission from Interstellar Hydrogen Cyanide. *The Astrophysical Journal Letters* **1971**, *163*, L47.

(109) Huebner, W. F.; Snyder, L. E.; Buhl, D. HCN radio emission from Comet Kohoutek (1973f). *Icarus* **1974**, *23*, 580–584.

(110) Saykally, R. J.; Szanto, P. G.; Anderson, T. G.; Woods, R. C. The microwave spectrum of hydrogen isocyanide. *The Astrophysical Journal Letters* **1976**, *204*, L143–L145.

(111) Belloche, A.; Menten, K. M.; Comito, C.; Müller, H. S. P.; Schilke, P.; Ott, J.; Thorwirth, S.; Hieret, C. Detection of amino acetonitrile in Sgr B2(N). *Astronomy & Astrophysics* **2008**, *482*, 179–196.

(112) Agúndez, M.; Wakelam, V. Chemistry of Dark Clouds: Databases, Networks, and Models. *Chemical Reviews* **2013**, *113*, 8710–8737.

(113) Zhu, C.; Frigge, R.; Bergantini, A.; Fortenberry, R. C.; Kaiser, R. I. Untangling the Formation of Methoxymethanol (CH3OCH2OH) and Dimethyl Peroxide (CH3OOCH3) in Star-forming Regions. *The Astrophysical Journal* **2019**, *881*, 156.

(114) Godfrey, P. D.; Brown, R. D.; Robinson, B. J.; Sinclair, M. W. Discovery of Interstellar Methanimine (Formaldimine). *Astrophysical Letters* **1973**, *13*, 119.

(115) Alonso, E. R.; Kolesniková, L.; Bia\lkowska-Jaworska, E.; Kisiel, Z.; León, I.; Guillemin, J.-C.; Alonso, J. L. Glycinamide, a Glycine Precursor, Caught in the Gas Phase: A Laser-ablation Jet-cooled Rotational Study. *The Astrophysical Journal* **2018**, *861*, 70.

(116) Bossa, J. B.; Theule, P.; Duvernay, F.; Chiavassa, T. NH $_2$ CH $_2$ OH THERMAL FORMATION IN INTERSTELLAR ICES CONTRIBUTION TO THE 5-8 m REGION TOWARD EMBEDDED PROTOSTARS. *The Astrophysical Journal* **2009**, *707*, 1524–1532.

(117) Singh, S. K.; Zhu, C.; La Jeunesse, J.; Fortenberry, R. C.; Kaiser, R. I. Experimental identification of aminomethanol (NH2CH2OH)—the key intermediate in the Strecker Synthesis. *Nature Communications* **2022**, *13*, 375.

(118) Sharma, M. K.; Chandra, S. Sobolev LVG Analysis of Aminomethanol and N-Methylhydroxylamine: Potential Spectral Lines for Their Detection in a Cosmic Object. *Astrophysics* **2021**, *64*, 388–404.

(119) Schutte, W.; Allamandola, L.; Sandford, S. An Experimental Study of the Organic Molecules Produced in Cometary and Interstellar Ice

Analogs by Thermal Formaldehyde Reactions. *Icarus* **1993**, *104*, 118–137.

(120) Rimola, A.; Skouteris, D.; Balucani, N.; Ceccarelli, C.; Enrique-Romero, J.; Taquet, V.; Ugliengo, P. Can Formamide Be Formed on Interstellar Ice? An Atomistic Perspective. *ACS Earth and Space Chemistry* **2018**, *2*, 720–734.

(121) Duarte, F.; Vöhringer-Martinez, E.; Toro-Labbé, A. Insights on the mechanism of proton transfer reactions in amino acids. *Physical Chemistry Chemical Physics* **2011**, *13*, 7773.

(122) Lamberts, T.; Vries, X. d.; Cuppen, H. M. The formation of ice mantles on interstellar grains revisited – the effect of exothermicity. *Faraday Discussions* **2014**, *168*, 327–347.

(123) Gauld, J. W.; Audier, H.; Fossey, J.; Radom, L. Water-Catalyzed Interconversion of Conventional and Distonic Radical Cations: Methanol and Methyleneoxonium Radical Cations. *Journal of the American Chemical Society* **1996**, *118*, 6299–6300.

(124) Szabó, A.; Ostlund, N. S., *Modern quantum chemistry : introduction to advanced electronic structure theory*; Mineola (N.Y.) : Dover publications, 1996.: 1989.

(125) McQuarrie, D. A.; Simon, J. D., *Physical chemistry: a molecular approach*; University Science Books: Sausalito, Calif, 1997.

(126) Steeb, W.-H., *Hilbert Spaces, Wavelets, Generalised Functions and Modern Quantum Mechanics*; Springer Netherlands: Dordrecht, 1998.

(127) Steeb, W.-H. In *Hilbert Spaces, Wavelets, Generalised Functions and Modern Quantum Mechanics*; Springer Netherlands: Dordrecht, 1998, pp 69–76.

(128) Jensen, F., *Introduction to computational chemistry*, Third edition; John Wiley & Sons: Chichester, UK ; Hoboken, NJ, 2017.

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

(130) Purvis, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *The Journal of Chemical Physics* **1982**, *76*, 1910–1918.

(131) Bartlett, R. J.; Dykstra, C. E.; Paldus, J. In *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*, Dykstra, C. E., Ed.; NATO ASI Series; Springer Netherlands: Dordrecht, 1984, pp 127–159.

(132) Bartlett, R. J. Coupled-cluster approach to molecular structure and spectra: a step toward predictive quantum chemistry. *The Journal of Physical Chemistry* **1989**, *93*, 1697–1708.

(133) Cramer, C. J. Essentials of Computational Chemistry: Theories and Models, 2nd Edition. **2002**.

(134) *Methods of Electronic Structure Theory*; Schaefer, H. F., Ed.; Springer US: Boston, MA, 1977.

(135) Brueckner, K. A. Many-Body Problem for Strongly Interacting Particles. II. Linked Cluster Expansion. *Physical Review* **1955**, *100*, 36–45.

(136) Bartlett, R. J.; Purvis, G. D. Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitations for the correlation problem. *International Journal of Quantum Chemistry* **1978**, *14*, 561–581.

(137) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Physical Review* **1934**, *46*, 618–622.

(138) Sinanoğlu, O. Many-Electron Theory of Atoms and Molecules. II. *The Journal of Chemical Physics* **1962**, *36*, 3198–3208.

(139) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters* **1989**, *157*, 479–483.

(140) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *The Journal of Chemical Physics* **1997**, *106*, 9639–9646.

(141) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *The Journal of Chemical Physics* **2007**, *127*, 221106.

(142) Warden, C. E.; Smith, D. G. A.; Burns, L. A.; Bozkaya, U.; Sherrill, C. D. Efficient and automated computation of accurate molecular geometries using focal-point approximations to large-basis coupled-cluster theory. *The Journal of Chemical Physics* **2020**, *152*, 124109.

(143) Koch, W.; Holthausen, M. C. A Chemist's Guide to Density Functional Theory, 2nd Edition | Wiley, en-us, 2001.

(144) Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual Density Functional Theory. *Chemical Reviews* **2003**, *103*, 1793–1874.

(145) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Physical Review* **1964**, *136*, B864–B871.

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

(147) Lieb, E. H.; Oxford, S. Improved lower bound on the indirect Coulomb energy. *International Journal of Quantum Chemistry* **1981**, *19*, 427–439.

(148) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A* **1988**, *38*, 3098–3100.

(149) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*, 3865–3868.

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

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

(152) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B* **1992**, *46*, 6671–6687.

(153) Becke, A. D.; Roussel, M. R. Exchange holes in inhomogeneous systems: A coordinate-space model. *Physical Review A* **1989**, *39*, 3761–3767.

(154) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids. *Physical Review Letters* **2003**, *91*, 146401.

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

(156) Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. *The Journal of Chemical Physics* **1999**, *110*, 5029–5036.

(157) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *The Journal of Chemical Physics* **2003**, *119*, 12129–12137.

(158) Vydrov, O. A.; Scuseria, G. E. Assessment of a long-range corrected hybrid functional. *The Journal of Chemical Physics* **2006**, *125*, 234109.

(159) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics* **2008**, *10*, 6615–6620.

(160) Alder, B. J.; Wainwright, T. E. Phase Transition for a Hard Sphere System. *The Journal of Chemical Physics* **1957**, *27*, 1208–1209.

(161) Rick, S. W.; Stuart, S. J. In *Reviews in Computational Chemistry, Volume 18*, Lipkowitz, K. B., Boyd, D. B., Eds.; John Wiley & Sons, Inc.: Hoboken, New Jersey, USA, 2002; Vol. 18, pp 89–146.

(162) Iftimie, R.; Minary, P.; Tuckerman, M. E. *Ab initio* molecular dynamics: Concepts, recent developments, and future trends. *Proceedings of the National Academy of Sciences* **2005**, *102*, 6654–6659.

(163) Warshel, A.; Weiss, R. M. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *Journal of the American Chemical Society* **1980**, *102*, 6218–6226.

(164) Tuckerman, M. E. *Ab initio* molecular dynamics: basic concepts, current trends and novel applications. *Journal of Physics: Condensed Matter* **2002**, *14*, R1297–R1355.

(165) Kühne, T. D. Ab-Initio Molecular Dynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 391–406.

(166) Paquet, E.; Viktor, H. L. Computational Methods for Ab Initio Molecular Dynamics, en, Review Article, 2018.

(167) Born, M.; Fock, V. Beweis des Adiabatensatzes. *Zeitschrift fur Physik* **1928**, *51*, 165–180.

(168) Zaverkin, V.; Kästner, J. Gaussian Moments as Physically Inspired Molecular Descriptors for Accurate and Scalable Machine Learning Potentials. *Journal of Chemical Theory and Computation* **2020**, *16*, 5410–5421.

(169) Zaverkin, V.; Molpeceres, G.; Kästner, J. Neural-network assisted study of nitrogen atom dynamics on amorphous solid water – II. Diffusion. *Monthly Notices of the Royal Astronomical Society* **2022**, *510*, 3063–3070.

(170) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Physical Review Letters* **2007**, *98*, 146401.

(171) Jónsson, H.; Mills, G.; Jacobsen, K. W. In *Classical and Quantum Dynamics in Condensed Phase Simulations*, WORLD SCIENTIFIC: LERICI, Villa Marigola, 1998, pp 385–404.

(172) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics* **2000**, *113*, 9901–9904.

(173) Nance, J.; Jakubikova, E.; Kelley, C. T. Reaction Path Following with Sparse Interpolation. *Journal of Chemical Theory and Computation* **2014**, *10*, 2942–2949.

(174) Vogt-Geisse, S.; Toro-Labbé, A. The mechanism of the interstellar isomerization reaction HOC+→HCO+ catalyzed by H2: New Insights from the reaction electronic flux. *The Journal of Chemical Physics* **2009**, *130*, 244308.

(175) Fukui, K. The path of chemical reactions - the IRC approach. *Accounts of Chemical Research* **1981**.

(176) Toro-Labbé, A. Characterization of Chemical Reactions from the Profiles of Energy, Chemical Potential, and Hardness. *The Journal of Physical Chemistry A* **1999**, *103*, 4398–4403.

(177) Kraka, E.; Cremer, D. Computational Analysis of the Mechanism of Chemical Reactions in Terms of Reaction Phases: Hidden Intermediates and Hidden Transition States. *Accounts of Chemical Research* **2010**, *43*, 591–601.

(178) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics* **1970**, *19*, 553–566.

(179) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chemical Reviews* **1994**, *94*, 1887–1930.

(180) Gonthier, J. F.; Sherrill, C. D. Density-fitted open-shell symmetry-adapted perturbation theory and application to -stacking in benzene dimer cation and ionized DNA base pair steps. *The Journal of Chemical Physics* **2016**, *145*, 134106.

(181) Weinhold, F.; Wilson, E. B. Reduced Density Matrices of Atoms and Molecules. I. The 2 Matrix of Double-Occupancy, Configuration-Interaction Wavefunctions for Singlet States. *The Journal of Chemical Physics* **1967**, *46*, 2752–2758.

(182) Murray, J. S.; Politzer, P. The electrostatic potential: an overview. *WIREs Computational Molecular Science* **2011**, *1*, 153–163.

(183) Smith, D. G. A.; Altarawy, D.; Burns, L. A.; Welborn, M.; Naden, L. N.; Ward, L.; Ellis, S.; Pritchard, B. P.; Crawford, T. D. The MolSSI QCArchive project: An open-source platform to compute, organize, and share quantum chemistry data. *WIREs Computational Molecular Science* **2020**, *n/a*, e1491.

(184) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010**, *132*, 154104.

(185) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *Journal of Chemical Theory and Computation* **2009**, *5*, 2619–2628.

(186) Titov, A. V.; Ufimtsev, I. S.; Luehr, N.; Martinez, T. J. Generating Efficient Quantum Chemistry Codes for Novel Architectures. *Journal of Chemical Theory and Computation* **2013**, *9*, 213–221.

(187) Werner, H.-J. et al. The Molpro quantum chemistry package. *The Journal of Chemical Physics* **2020**, *152*, 144107.

(188) Parrish, R. M. et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Inter-

operability. *Journal of Chemical Theory and Computation* **2017**, *13*, 3185–3197.

(189) Hjorth Larsen, A. et al. The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter* **2017**, *29*, 273002.

(190) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *The Journal of Chemical Physics* **2020**, *152*, 224108.

(191) Frisch, M. J. et al. Gaussian~16 Revision C.01, 2016.

(192) Vogt-Geisse, S. Kudi: A free open-source python library for the analysis of properties along reaction paths. *Journal of Molecular Modeling* **2016**, *22*, 110.

(193) Glendening, E. D.; Landis, C. R.; Weinhold, F. *NBO 6.0* : Natural bond orbital analysis program. *Journal of Computational Chemistry* **2013**, *34*, 1429–1437.

(194) MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry B* **1998**, *102*, 3586–3616.

(195) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr. *Chemical Physics Letters* **1989**, *157*, 200–206.

(196) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*.

(197) Szalewicz, K.; Leforestier, C.; van der Avoird, A. Towards the complete understanding of water by a first-principles computational approach. *Chemical Physics Letters* **2009**, *482*, 1–14.

(198) Chen, M.; Ko, H.-Y.; Remsing, R. C.; Calegari Andrade, M. F.; Santra, B.; Sun, Z.; Selloni, A.; Car, R.; Klein, M. L.; Perdew, J. P.; Wu, X. Ab initio theory and modeling of water. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114*, 10846–10851.

(199) Schneider, T.; Stoll, E. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Physical Review B* **1978**, *17*, 1302–1322.

(200) Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography* **1976**, *32*, 922–923.

(201) Brown, M. E.; Calvin, W. M. Evidence for Crystalline Water and Ammonia Ices on Pluto's Satellite Charon. *Science* **2000**, *287*, 107–109.

(202) Emery, J. P.; Burr, D. M.; Cruikshank, D. P.; Brown, R. H.; Dalton, J. B. Near-infrared (0.8–4.0 m) spectroscopy of Mimas, Enceladus, Tethys, and Rhea. *Astronomy & Astrophysics* **2005**, *435*, 353–362.

(203) Cook, J. C.; Desch, S. J.; Roush, T. L.; Trujillo, C. A.; Geballe, T. R. Near-Infrared Spectroscopy of Charon: Possible Evidence for Cryovolcanism on Kuiper Belt Objects. *The Astrophysical Journal* **2007**, *663*, 1406–1419.

(204) Lorenz, R. D.; Shandera, S. E. Physical properties of ammonia-rich ice: Application to Titan. *Geophysical Research Letters* **2001**, *28*, 215–218.

(205) Moore, M.; Ferrante, R.; Hudson, R.; Stone, J. Ammonia–water ice laboratory studies relevant to outer Solar System surfaces. *Icarus* **2007**, *190*, 260–273.

(206) Suter, M. T.; Andersson, P. U.; Pettersson, J. B. C. Formation of water–ammonia ice on graphite studied by elastic helium scattering. *Chemical Physics Letters* **2007**, *445*, 208–212.

(207) Zheng, W.; Jewitt, D.; Kaiser, R. I. INFRARED SPECTRA OF AMMONIA-WATER ICES. *The Astrophysical Journal Supplement Series* **2009**, *181*, 53–61.

(208) Giuliano, B. M.; Martín-Doménech, R.; Escribano, R. M.; Manzano-Santamaría, J.; Muñoz Caro, G. M. Interstellar ice analogs: H $_2$ O ice

mixtures with CH $_3$ OH and NH $_3$ in the far-IR region. *Astronomy & Astrophysics* **2016**, *592*, A81.

(209) Liu, C. et al. Topologically frustrated ionisation in a water-ammonia ice mixture. *Nature Communications* **2017**, *8*, 1065.

(210) Robinson, V. N.; Wang, Y.; Ma, Y.; Hermann, A. Stabilization of ammonia-rich hydrate inside icy planets. *Proceedings of the National Academy of Sciences* **2017**.

(211) Al-Halabi, A.; Van Dishoeck, E. F. Hydrogen adsorption and diffusion on amorphous solid water ice. *Monthly Notices of the Royal Astronomical Society* **2007**, *382*, 1648–1656.

(212) Nagasawa, T.; Sato, R.; Hasegawa, T.; Numadate, N.; Shioya, N.; Shimoaka, T.; Hasegawa, T.; Hama, T. Absolute Absorption Cross Section and Orientation of Dangling OH Bonds in Water Ice. *The Astrophysical Journal Letters* **2021**, *923*, L3.

(213) Zhao, Y.; Truhlar, D. G. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *The Journal of Physical Chemistry. A* **2005**, *109*, 5656–5667.

(214) Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *Journal of Computational Chemistry* **2013**, *34*, 1672–1685.

(215) Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *The Journal of Chemical Physics* **2015**, *143*, 054107.

(216) Vydrov, O. A.; Scuseria, G. E.; Perdew, J. P. Tests of functionals for systems with fractional electron number. *The Journal of Chemical Physics* **2007**, *126*, 154109.

(217) Vogt-Geisse S.; Bovolenta, G. M. BEEP, `www.github.com/QCMM/beep`.

(218) Neufeld, D. A.; Wolfire, M. G.; Schilke, P. The Chemistry of Fluorine-bearing Molecules in Diffuse and Dense Interstellar Gas Clouds. *The Astrophysical Journal* **2005**, *628*, 260.

(219) Van der Wiel, M. H. D.; Naylor, D. A.; Makiwa, G.; Satta, M.; Abergel, A. Three-dimensional distribution of hydrogen fluoride gas toward NGC 6334 I and I(N). *Astronomy & Astrophysics* **2016**, *593*, A37.

(220) Jeffrey, G. A., *An introduction to hydrogen bonding*; Oxford University Press: 1997.

(221) Chakraborty, D.; Chandra, A. Hydrogen bonded structure and dynamics of liquid-vapor interface of water-ammonia mixture: An *ab initio* molecular dynamics study. *The Journal of Chemical Physics* **2011**, *135*, 114510.

(222) Minissale, M. et al. Thermal Desorption of Interstellar Ices: A Review on the Controlling Parameters and Their Implications from Snowlines to Chemical Complexity. *ACS Earth and Space Chemistry* **2022**, *6*, 597–630.

(223) Bovolenta, G. M.; Vogt-Geisse, S.; Bovino, S.; Grassi, T. Binding Energy Evaluation Platform: A Database of Quantum Chemical Binding Energy Distributions for the Astrochemical Community. *The Astrophysical Journal Supplement Series* **2022**, *262*, 17.

(224) McMurry, J., *Organic Chemistry*, 7th ed.; Thomson Brooks Cole: 2008.

(225) Baiano, C.; Lupi, J.; Barone, V.; Tasinato, N. Gliding on Ice in Search of Accurate and Cost-Effective Computational Methods for Astrochemistry on Grains: The Puzzling Case of the HCN Isomerization. *Journal of Chemical Theory and Computation* **2022**, *18*, 3111–3121.

(226) Chen, S.-J.; Chen, C.; Hong, Y.-S. Theoretical influence of third molecule on reaction channels of weakly bound complex CO2? HF systems. *International Journal of Quantum Chemistry* **2006**, *106*, 1640–1652.

(227) Hasegawa, T. I.; Herbst, E. New gas-grain chemical models of quiescent dense interstellar clouds - The effects of H2 tunnelling reactions and cosmic ray induced desorption. *Monthly Notices of the Royal Astronomical Society* **1993**, *261*, 83–102.

(228) Goumans, T. P. M. Isotope effects for formaldehyde plus hydrogen addition and abstraction reactions: rate calculations including tunnelling. *Monthly Notices of the Royal Astronomical Society* **2011**, *413*, 2615–2620.

(229) Halfen, D. T.; Ilyushin, V.; Ziurys, L. M. FORMATION OF PEPTIDE BONDS IN SPACE: A COMPREHENSIVE STUDY OF FORMAMIDE AND ACETAMIDE IN Sgr B2(N). *The Astrophysical Journal* **2011**, *743*, 60.

(230) Suzuki, T.; Ohishi, M.; Hirota, T.; Saito, M.; Majumdar, L.; Wakelam, V. SURVEY OBSERVATIONS OF A POSSIBLE GLYCINE PRECURSOR, METHANIMINE (CH2NH). *The Astrophysical Journal* **2016**, *825*, 79.

(231) Rimola, A.; Ferrero, S.; Germain, A.; Corno, M.; Ugliengo, P. Computational Surface Modelling of Ices and Minerals of Interstellar Interest—Insights and Perspectives. *Minerals* **2021**, *11*, 26.

(232) Miotello, A.; Kamp, I.; Birnstiel, T.; Cleeves, L. I.; Kataoka, A. Setting the Stage for Planet Formation: Measurements and Implications of the Fundamental Disk Properties, en, 2022.

(233) Zaverkin, V.; Holzmüller, D.; Steinwart, I.; Kästner, J. Fast and Sample-Efficient Interatomic Neural Network Potentials for Molecules and

Materials Based on Gaussian Moments. *Journal of Chemical Theory and Computation* **2021**, *17*, 6658–6670.

(234) Zhao, Y.; Truhlar, D. G. Hybrid Meta Density Functional Theory Methods for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions: The MPW1B95 and MPWB1K Models and Comparative Assessments for Hydrogen Bonding and van der Waals Interactions. *The Journal of Physical Chemistry A* **2004**, *108*, 6908–6918.

(235) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation* **2019**, *15*, 1652–1671.

(236) Gale, J. D.; LeBlanc, L. M.; Spackman, P. R.; Silvestri, A.; Raiteri, P. A Universal Force Field for Materials, Periodic GFN-FF: Implementation and Examination. *Journal of Chemical Theory and Computation* **2021**, *17*, 7827–7849.

(237) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **2007**, *9*, 90–95.

Appendix

A. Astrophysical framework

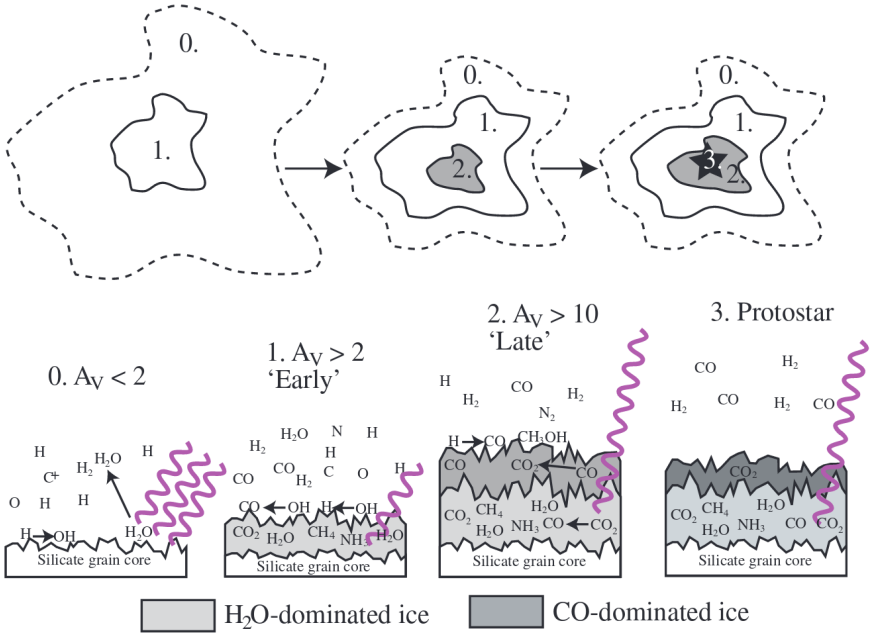
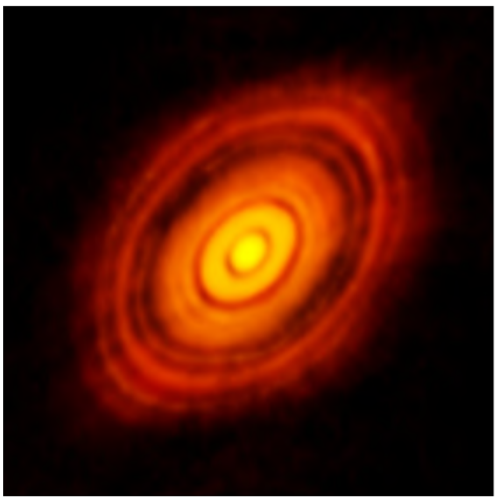
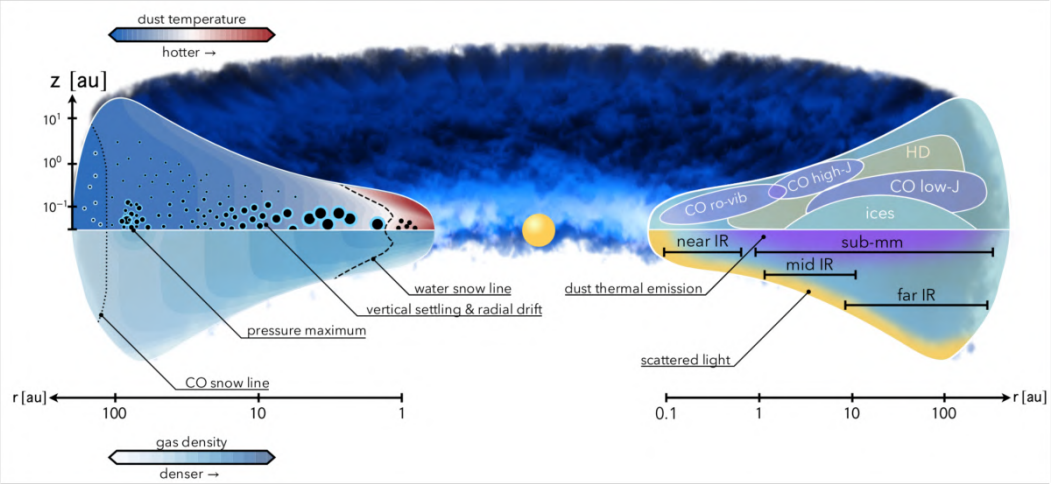


Figure A.1: Key ice reactions during the different stages of cloud and star formation, where 0. corresponds to cloud edges where UV photons can penetrate, 1. the early stages of dense cloud formation, 2. the later formation of cloud cores, and 3. the protostellar envelope. Ice formation begins in 1. with hydrogenation of atoms resulting in an H_2O-dominated ice with a high CO_2 concentration. At later times CO freezes out catastrophically, resulting in a second layer where CO_2 formation continues and CH_3OH formation begins. In the protostellar stage desorption of CO starts at 20 K and H_2O:CO_2 segregation at 30 K. Adapted from Öberg et al[11].



a)



b)

Figure A.2: a) ALMA sharpest observation of a protoplanetary disk surrounding the young star HL Tauri. The dark ring midway through the disk is the water snowline, the point from the star where the temperature and pressure are low enough for ice to form. Credit: ALMA (ESO/NAOJ/NRAO); b) Left: Illustration of the dust temperature and gas density structure in a protoplanetary disks. A sketched distribution of dust particles is shown by the black circles: bare grains are present within the water snowline (dashed curve), H_2O-coated grains are contoured in blue, and CO-coated grains, outside the CO snowline (dotted curve), are contoured in white. Right: the top panel shows a simplified representation of the emission regions of the main simple molecules, while the main emission regions are highlighted in purple and yellow in the bottom panel. Adapted from Miotello et al.[232]

B. Machine Learning Potential

B.1 Training procedure

A *ad hoc* Machine Learning Potential (MLP) is trained for the astrophysically relevant system: formyl radical (HCO) adsorbed on ASW, using the GM-NN potential[168,233]. The training set configurations consist of 9226 structures, generated *via* MD simulations at various temperatures on water clusters of difference sizes (first group, sub-total 1700) as well as equilibrium structures where the target molecule interact with the water clusters (second group: 7526). Table B.1 reports the details about training set composition and propagation methods used. Energy and forces of DFT quality are subsequently computed for the total set of configurations extracted from the MD trajectories. The DFT method used is MPWB1K-D3BJ/def2-TZVP[234], selected based on the results of geometry and energy benchmarks carried out in a previous binding energy study[223]. BSSE error has been accounted for using the gCP factor, which incorporates an approximation of the counterpoise correction to the energy calculation, as provided by Orca software. D3BJ dispersion correction has been applied to all DFT energies.

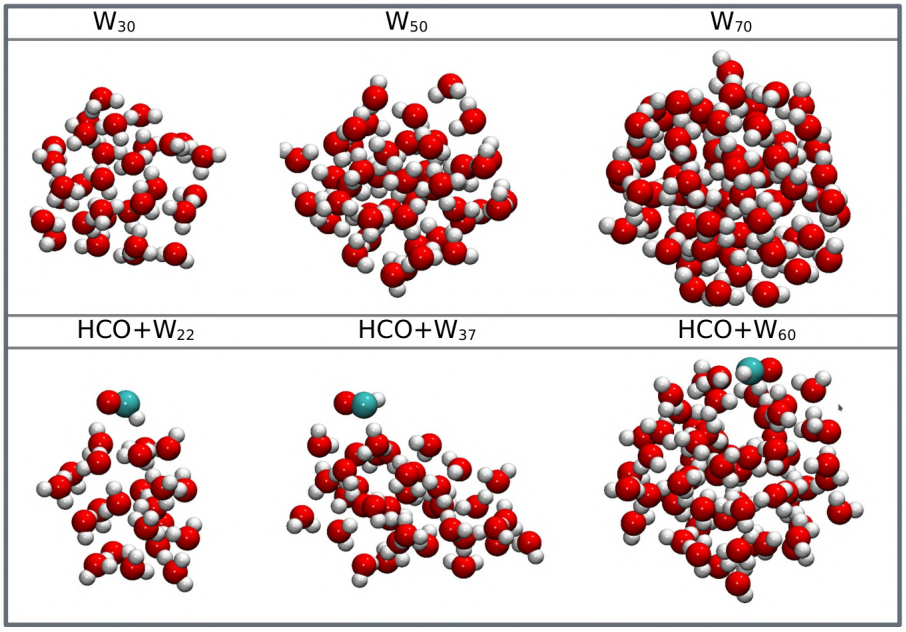


Figure B.1: Training set composition.

.

An ensemble of three MLP models has been trained for 1000 epochs, using the same training data but with different randomly initialized parameters. As long as the deviation between the prediction of the models is small, the structure is well-represented within the training set and the MLP PES generalizes well. All relevant hyper-parameters used for the training procedure can be found in the original work[233].

Table B.1: Structures included in the training data. The last column refers to the number of structures extracted for refinement for each system. The propagation method is only used for sampling geometries, while energy and gradients are computed at MPWB1K-D3BJ/def2-TZVP level of theory.

System	Propagation Method (NVT)	T(K)	Number of points
W_{30}	GFN2-xTB[235]	50/300/500	200/400/400
W_{50}	GFN2-xTB	50/300/500	167/167/167
W_{70}	GFN2-xTB	50/300/500	50/50/100
			sub-total: 1700
HCO + W_{22}	GFN-FF[236]	100/300	150/150
	metadynamics - GFN-FF		150
HCO + W_{37}	GM-NN[a]	100	500
	GFN-FF	300	500
	metadynamics - GFN-FF		500
HCO + W_{60}	GM-NN[a]	100	100
	GFN-FF	300	100
	metadynamics - GFN-FF		100
			sub-total: 7526
			Total: 9226

[a] We used a previously trained (intermediate) GM-NN.

B.2 Quality tests in reproducing the BE

Several tests have been performed to evaluate the performance of the MLP. HCO – W_{22} set-of-clusters is used as model system. The first test is aimed to assess the overall quality of the MLP, in reproducing binding sites geometry and DFT BEs. The group of $\sim$ 250 binding sites of the system produced in a previous work[223] has been re-optimized using the MLP, followed by MLP BEs estimation. Figure B.2 shows the comparison between the original DFT BE distribution and the MLP one. The agreement between the results is below 0.5 kcal/mol in Mean BE, showing a superlative performance of the MPL. Moreover, a second test has been carried in order to assess MLP

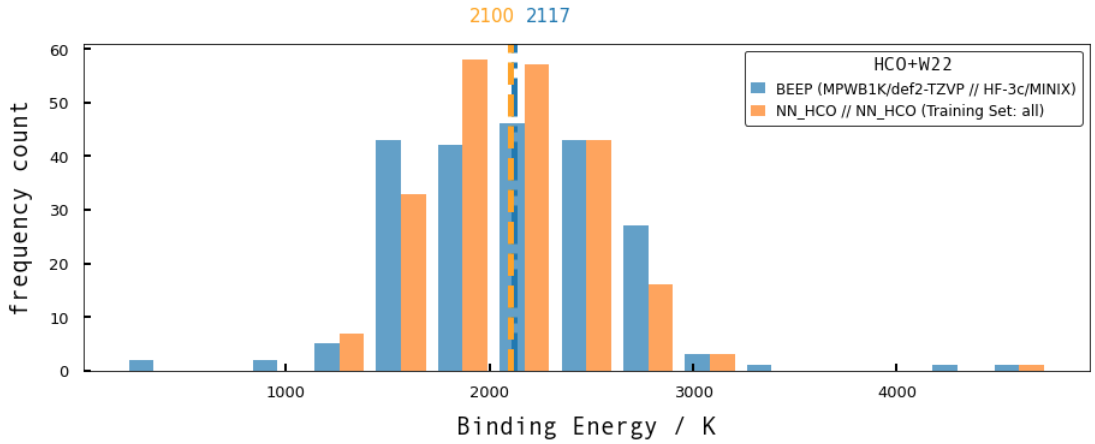


Figure B.2: System: ∼250 structures of HCO adsorbed on a 22-water molecules set-of-clusters. Comparison between DFT BE distribution obtained using BEEP methodology and infrastructure (blue) and the distribution computed on the same set of binding sites using the trained MLP (orange). Dashed lines represent the average BE for each set.

ability in specifically reproducing the BE, estimating MLP BEs for a subset of DFT structures extracted from the original distribution. The mean absolute error (MAE) in BE for the subset, taken as the average of the sum of the error of each single structure, is < 0.5 kcal/mol, confirming the reliability of the MLP.

B.3 Computational pipeline for BE calculation for the periodic surfaces

We selected 5 surfaces per density, resulting in a total of 25 periodic 500-ASW surfaces for each of the two system. We then proceeded to sample the surfaces with the target molecule. The sampling algorithm is composed of 5

steps. (1) Generate a grid on the ice surface (we assumed a step of 3Å to be a reasonable diffusion hopping distance for the molecule), we introduced noise on the grid points by relaxing the grid coordinate by $\pm$ 1/4 of the step size in both X and Y directions. (2) Place the molecule in corresponding of a grid point and adjust its position in order for the center of mass of the molecule to be placed at a specific distance from the nearest neighboring surface atom (we performed three rounds of sampling with initial distances of 2.50, 2.75 and 3.00 Å from the surface grid point). Furthermore, apply a final random rotation to the molecule to guarantee the largest variety of initial orientations toward the surface. (3) Fully optimize the so produced binding sites candidates using the respective MPL. We noted that the final number of binding site vary for each system, since the cell volume is established based on the initial density. (4) Filtered the minima belonging to an initial surface model (resulted from the three sampling rounds) grouping the structures according to geometrical criteria (similarity threshold of root-mean-square deviation of atomic positions (RMSD) $\leq$ 0.09 Å). The sampling procedure uses QCElemental package. The trained MLP have been used to optimize the binding site candidates and calculate their BE, obtaining a BEd for each density model. In order to include the Zero-Point Vibrational Energy (ZPVE) contribution to the BEd, we also use the MPL to compute the Hessian matrix for selected structures of model system (X-W22) at the equilibrium geometry and derive a linear model that we used to correct all the BE values.

C. Benchmark Results

Table C.1: Summary of the results of binding sites geometry and energy benchmarks for $W_{2-3}-X$ (W_2-X for radicals) and W_4-X (W_3-X for radicals) systems, respectively. The first column reports the molecules. Columns 2-3 report the performance of the best DFT functional for each group, and of HF-3c. Only structures that converged (n) to the reference minima (N) were considered for the benchmark. The forth column reports reference energies calculated at CCSD(T)/CBS level of theory. The fifth column reports the Mean Absolute Error (MAE) of the best DFT functional for each group. All DFT geometries and energies were computed using a def2-TZVP basis set and including D3BJ dispersion correction. HF-3c method is coupled with MINIX basis set.

	RMSD / Å		**BEs / K**	**MAE / K**
Group D	**B3LYP(n/N)**	**HF-3c(n/N)**	**CCSD(T)/CBS**	**ω-PBE**
H_2	0.11 (3/6)	0.14 (5/6)	320, 116	19
CO	0.12 (7/7)	0.22 (5/7)	950, 870, 791	10
CH_4	0.07 (2/2)	0.14 (2/2)	712	74
CH_3	0.09 (1/2)	0.12 (1/2)	821, 824	45
N_2	0.13 (3/3)	0.26 (3/3)		
Average	0.10 (16/20)	0.18 (16/20)		37
Group H	**PWB6K(n/N)**	**HF-3c(n/N)**	**CCSD(T)/CBS**	**ω-PBE**
NH_3	0.06 (4/7)	0.14 (5/7)	3632, 3516, 3562	79
CH_3OH	0.08 (8/8)	0.13 (6/8)	3922, 4111, 4005	119
HCOOH	0.06 (11/13)	0.17 (10/13)		
H_2CO	0.06 (5/6)	0.15 (4/6)	2600, 1197, 1181	338
HF	0.04 (3/4)	0.06 (2/4)	5956, 5380, 4158	83
HCl	0.07 (6/6)	0.18 (2/6)	3445, 2923, 956	146
HCO	0.05 (3/3)	0.07 (1/3)	2224, 1684	56
HNC	0.08 (4/5)	0.30 (3/5)	4211,3953	305
HCN	0.06 (4/5)	0.20 (3/5)		
Average	0.06 (48/57)	0.15 (36/57)		160

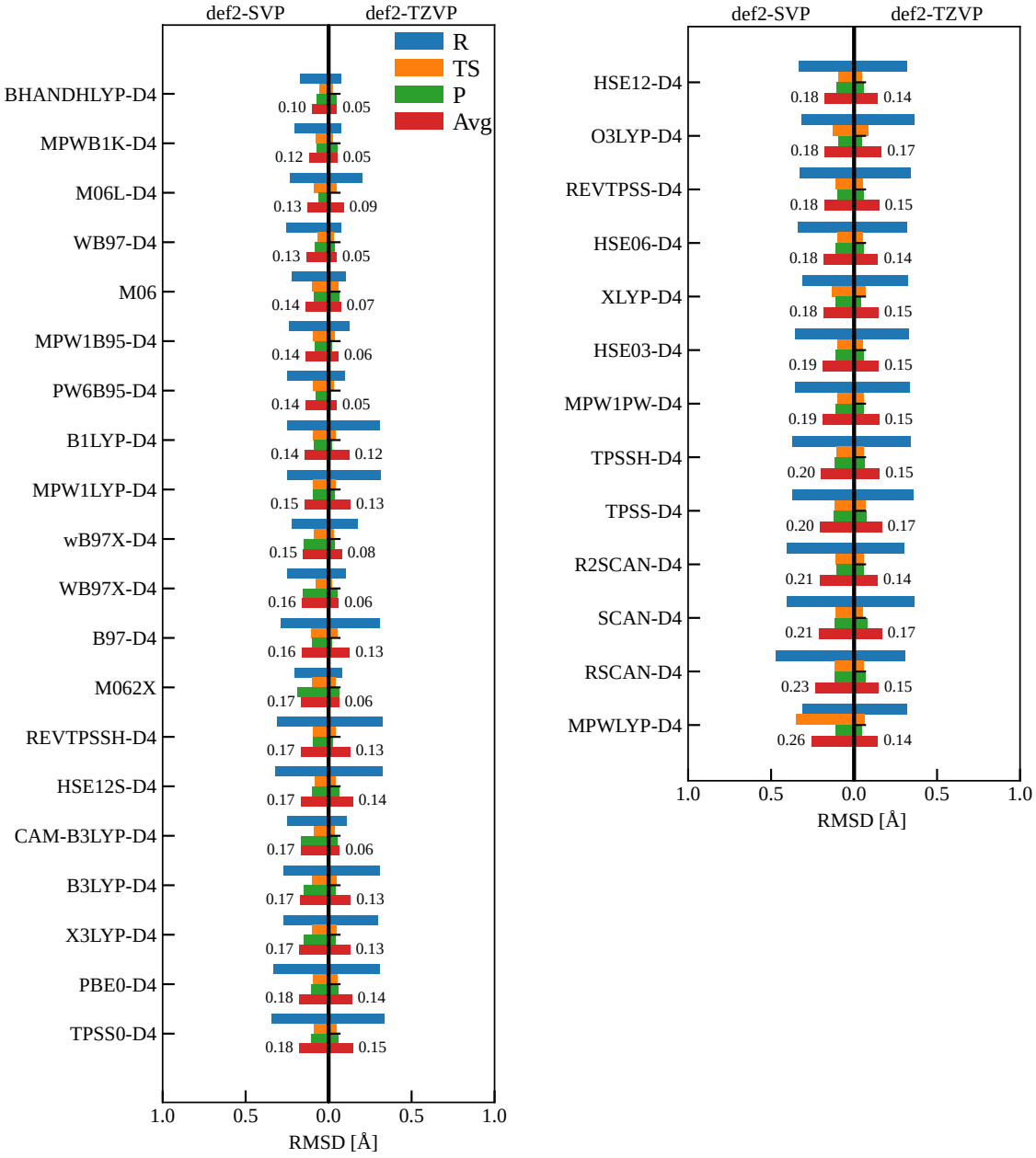


Figure C.1: RMSD for different xc-functionals used in the geometry optimization for the S1 reaction on the W_2 ice model. The reference geometry is DF-CCSD(T)-F12/cc-pVDZ (See Section 3.2.2)

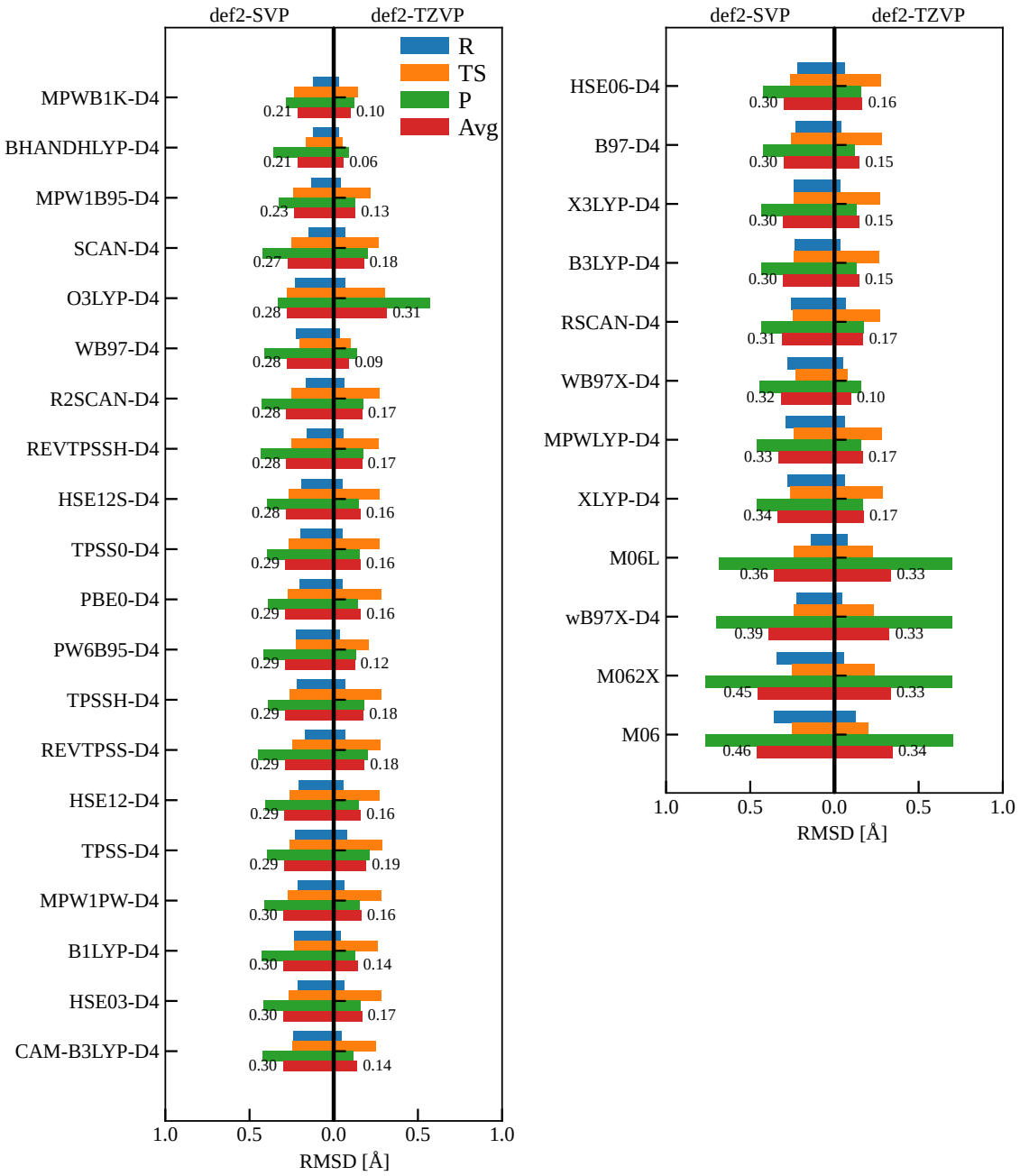


Figure C.2: RMSD for different xc-functionals used in the geometry optimization for the S2 reaction on the W_2 ice model. The reference geometry is DF-CCSD(T)-F12/cc-pVDZ (See Section 3.2.2)

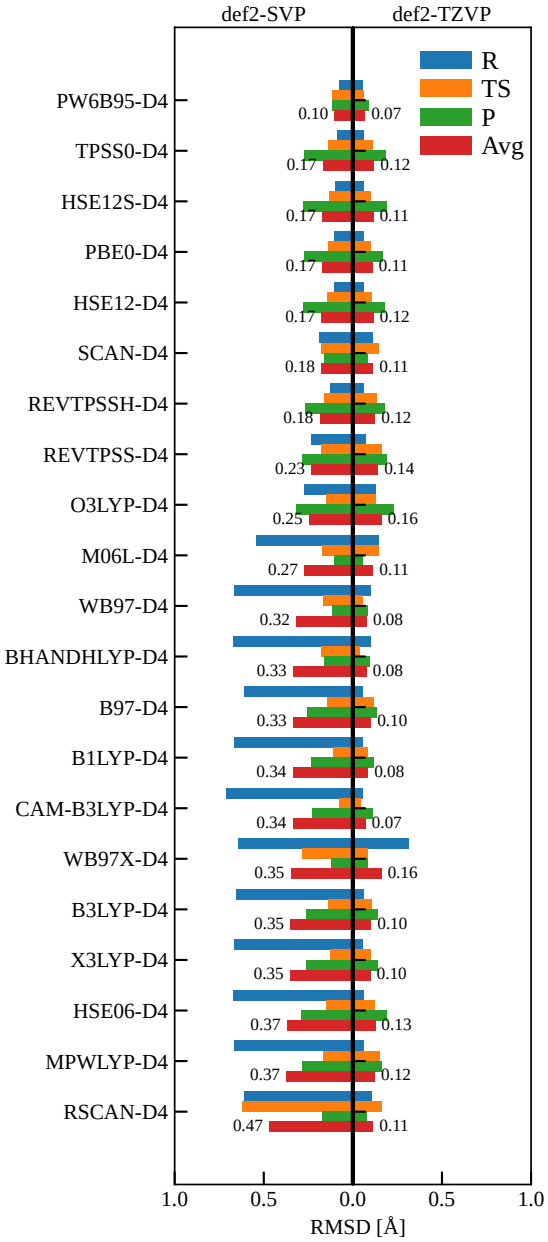


Figure C.3: RMSD for different xc-functionals used in the geometry optimization for the S3 reaction on the W_2 ice model. The reference geometry is DF-CCSD(T)-F12/cc-pVDZ (See Section 3.2.2)

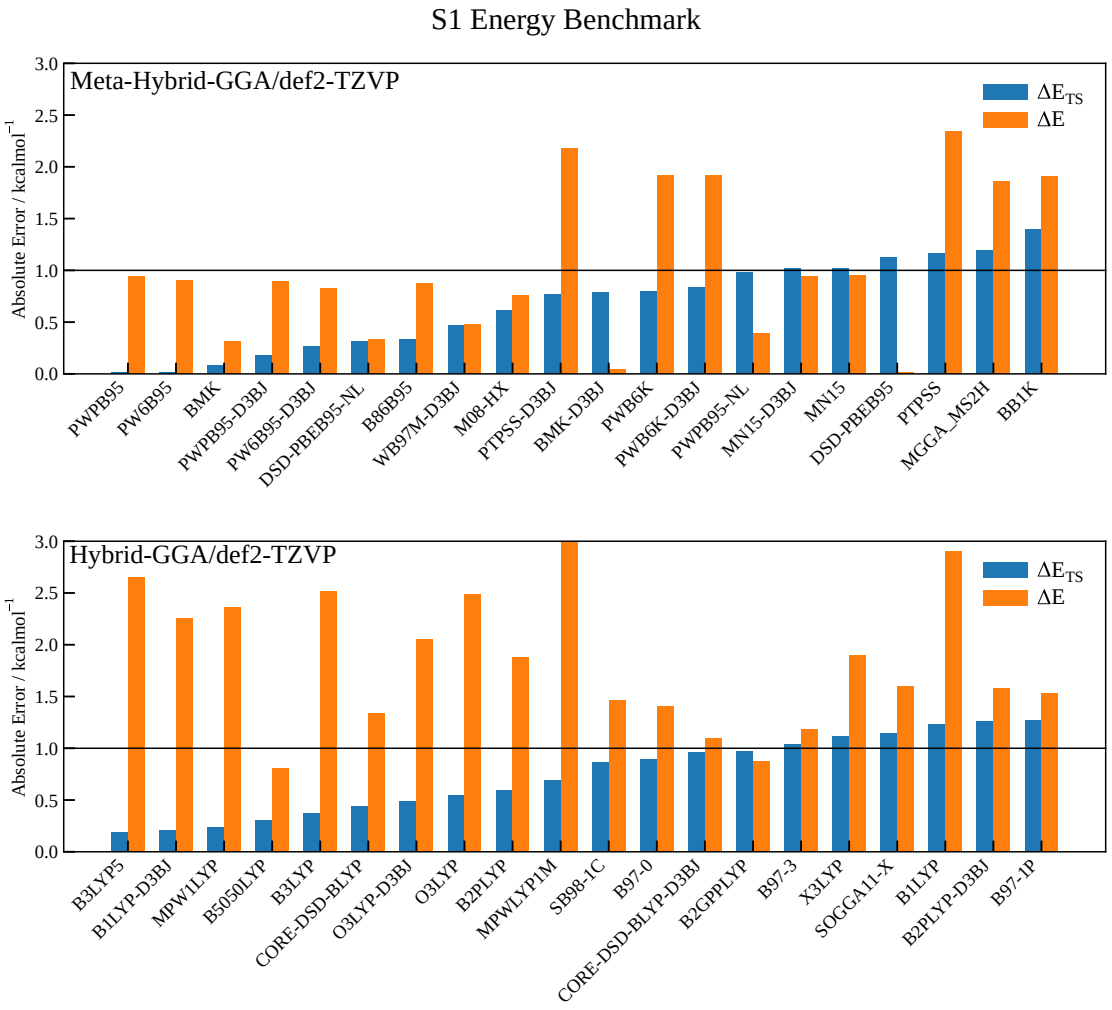


Figure C.4: Energy benchmark results for S1 on the DF-CCSD(T)-F12/cc-pVDZ geometry (See Section 3.2.2)

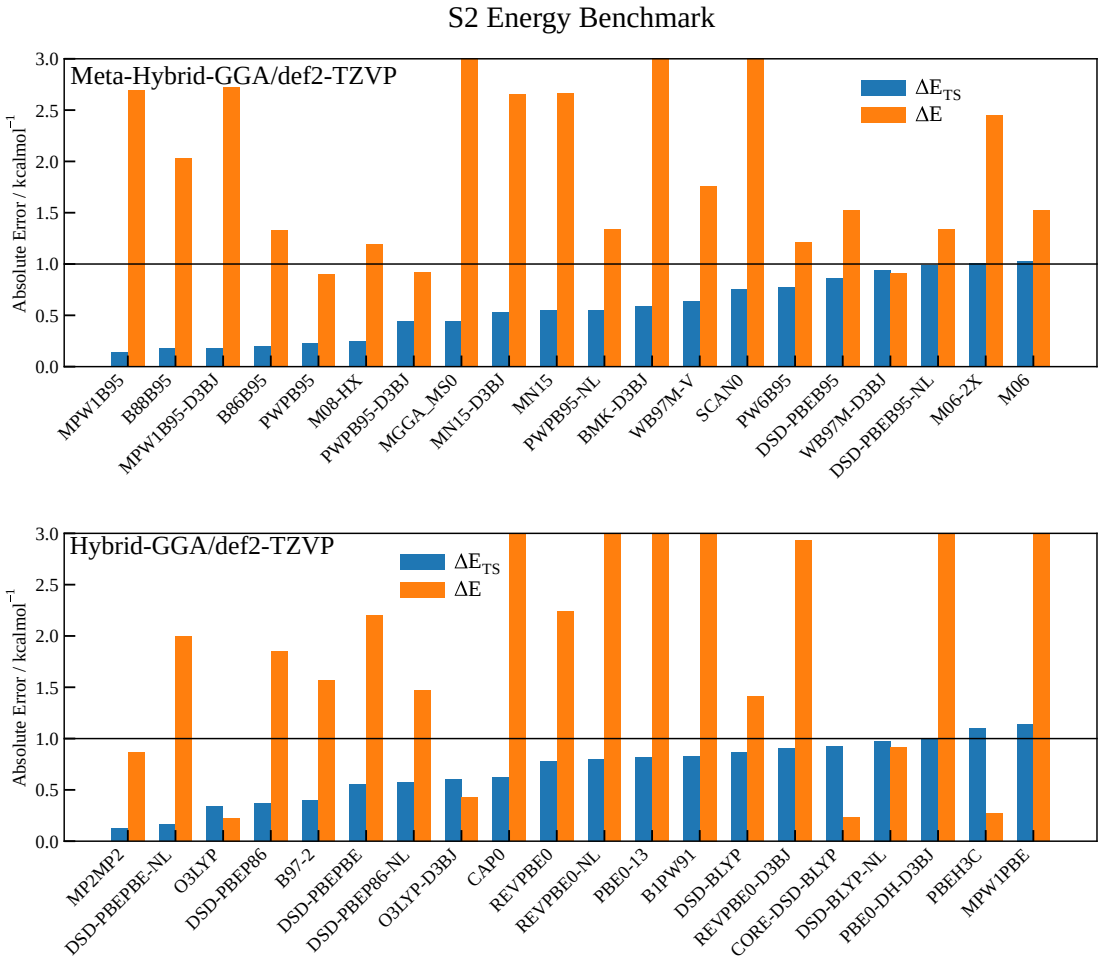


Figure C.5: Energy benchmark results for S2 on the DF-CCSD(T)-F12/cc-pVDZ geometry (See Section 3.2)

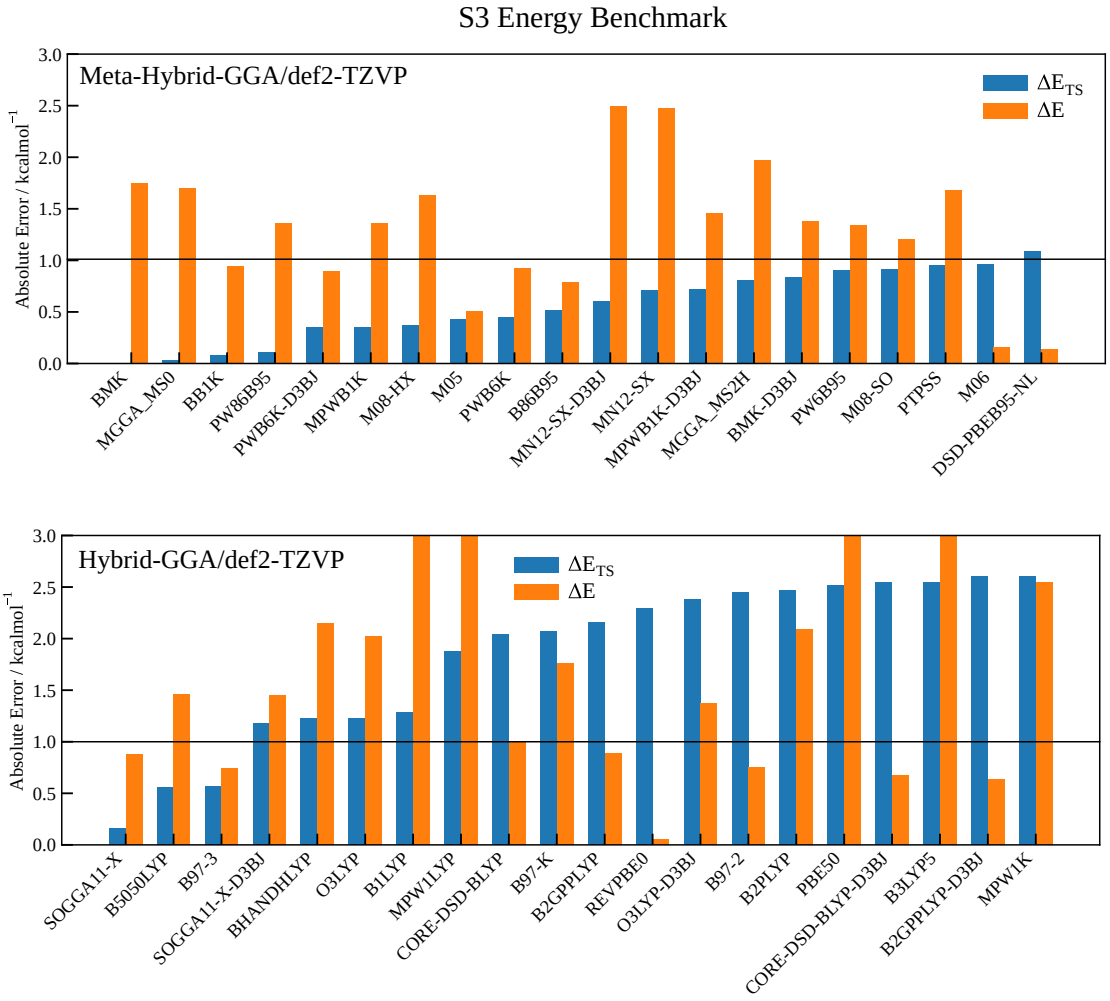


Figure C.6: Energy benchmark results for S3 on the DF-CCSD(T)-F12/cc-pVDZ geometry (See Section 3.2.2)

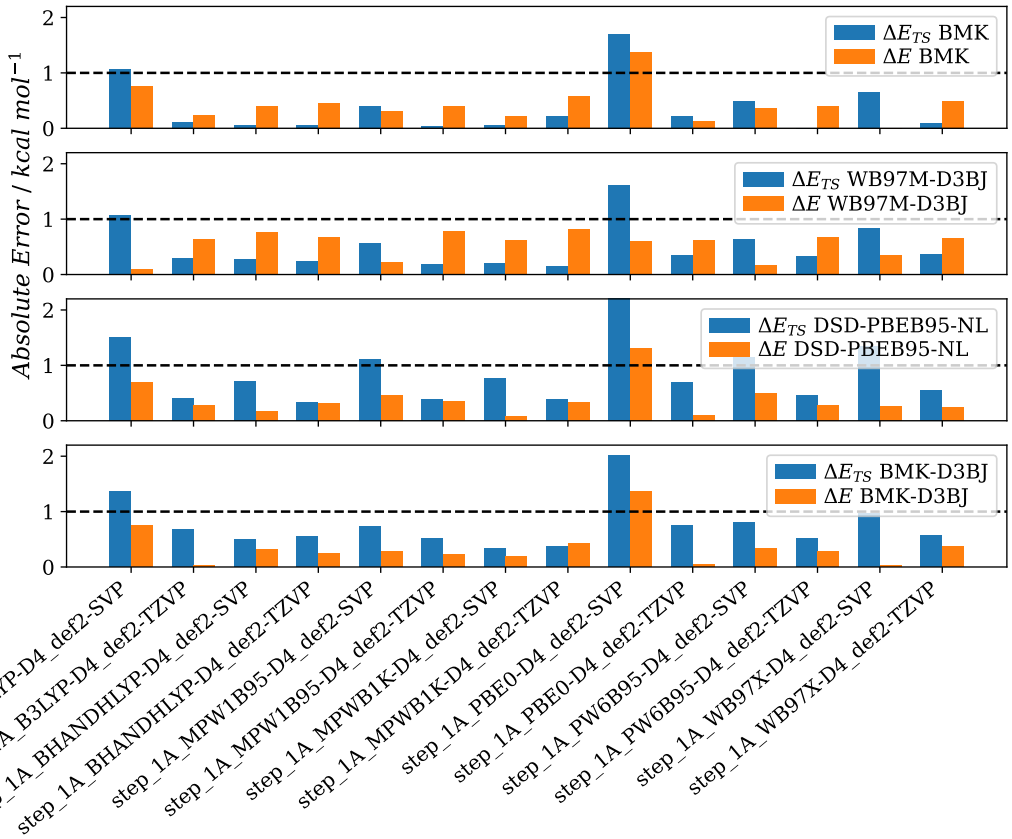


Figure C.7: Condensed benchmark results for the S1 reaction on a water dimer discussed in Section 3.2.2. Energies with four different xc-functionals (figure inset) were computed on the geometries that displayed the smallest RMSD values with respect to the CCSD(T)-F12/cc-pVDZ-F12 reference (Horizontal axis). Several models chemistry fall below the 1 kcal/mol chemical accuracy threshold.

D. Computational Details

D.1 Gaussian fitting procedure

To fit the BE distribution data with a Gaussian function, we employed a bootstrap method. We first divide our sample in equally-spaced bins, so that each bin contains N_i samples, with a Poisson error $\sqrt{N_i}$. We then produce 10^4 distributions analogue to the original data, randomizing the points assuming a Gaussian error of $\sqrt{N_i}$ around the mean N_i and we fit each distribution with

$$f(x) = a \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{D.1}$$

where a, μ, and σ are free parameters. The binned distribution of each parameter after the 10^4 iterations is also a Gaussian, where the average is the value we assume for the given parameter and the dispersion is the associated error.

D.2 Tri-Surface plots generation

A Triangular 3D surface (Tri-Surface) Plot is a type of surface plot, created by triangulation of compact surfaces using finite number of triangles in a manner that each point on the surface is a triangle. The intersection of any two triangles results in void or a common edge or vertex. The triangulation took into account solely atoms that belongs to the surface of the periodic ice models, based on their Z value. All surface plots are created using $ax.plot\_trisurf()$ function of matplotlib library[237].

D.3 Dispersion correction

The following Figure D.1 reports the comparison between the histograms of the BE distributions computed in this work, with and without including dispersion correction (D3BJ). The low impact of this contribution on the BE is reflected in a small shift of the distributions for most of the molecules in Group H, lower panel. On the other hand, the D3BJ correction is essential for Group D molecules, upper panel, as it shifts the BE distributions into the bound regime.

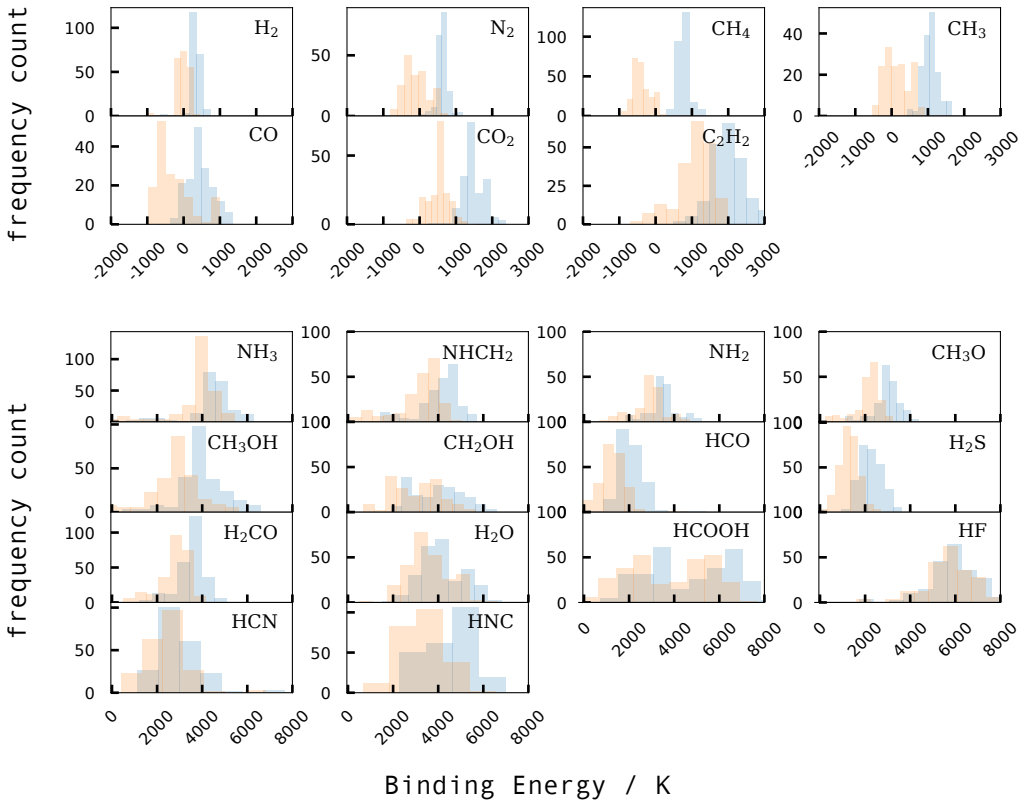


Figure D.1: Binding energy distributions computed using the best performing DFT functional from the energy benchmark for each molecule, with (blue) or without (orange) including D3BJ correction. Upper panel: Group D; lower panel: Group H.

E. Tables and Figures Reactivity

Table E.1: Bond distances in angstrom, angle in degrees, stage 1, small system (See Section **??**

System		B(C-N)	B(N-H1)	B(O-H2)	A(O-C-N)
	R	2.77	1.01	2.97	104.7
S_1-W_0	TS	1.57	1.16	1.41	97.2
	P	1.44	2.37	0.96	110.6
	R	2.63	1.02	1.88	102.1
S_1-W_1	TS	1.56	1.13	1.25	107.9
	P	1.46	1.91	0.97	110.7
	R	2.52	1.02	1.76	107.9
S_1-W_2	TS	1.54	1.11	1.14	109.3
	P	1.46	1.81	0.97	110.8

Table E.2: Energy barriers ($\Delta E^{\ddagger}$) and reaction energies (ΔE^{o}) computed at ωB97-M/def2-TZVP level of theory, using BHANDHLYP-D4/def2-SVP geometries, for $S4a-W_{1-2}$. Column two, W_X, indicates the number of water molecules involved in the proton relay. Values in kcal mol^{-1}.

System	W_X	$\Delta E^{\ddagger}$	ΔE^{o}
$S4a-W_1$	1	38.1	-5.3
$S4a-W_2$	2	44.3	-0.1

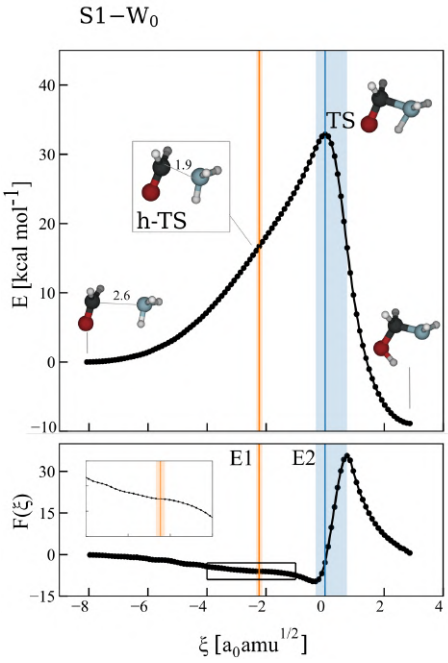
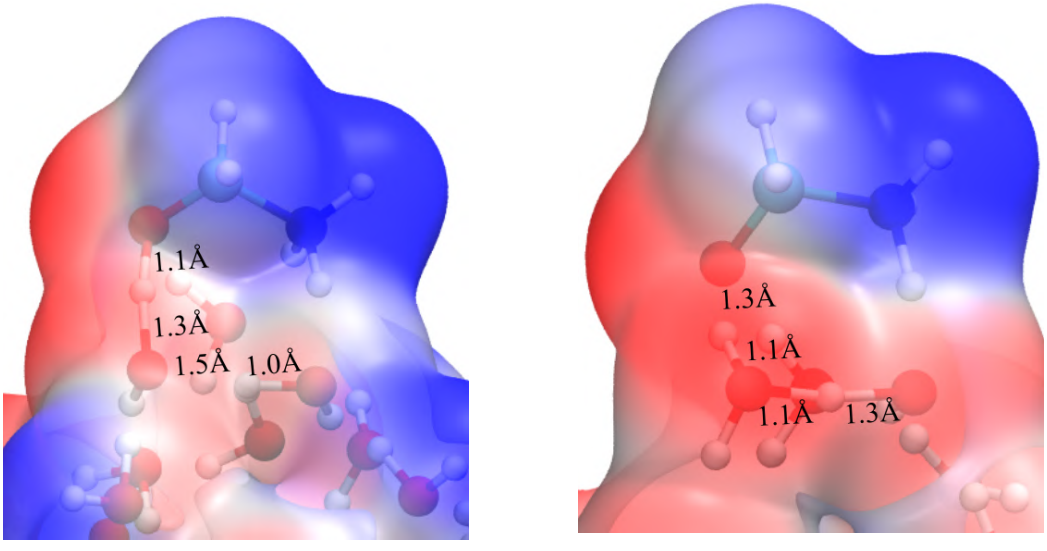


Figure E.1: Energy (upper panels) and reaction force profiles (lower panels) for S1 − W_0, using BMK/def2-TZVP//BHANDHLYP-D3BJ/def2-SVP levels of theory. Blue and orange lines represent TS and h-TS. TS and h-TS regions are displayed as orange and blue shadowed areas.



a) b)

Figure E.2: Comparison between MEP of two S1-C3 structures (TS). Example of HB coordination anomaly.

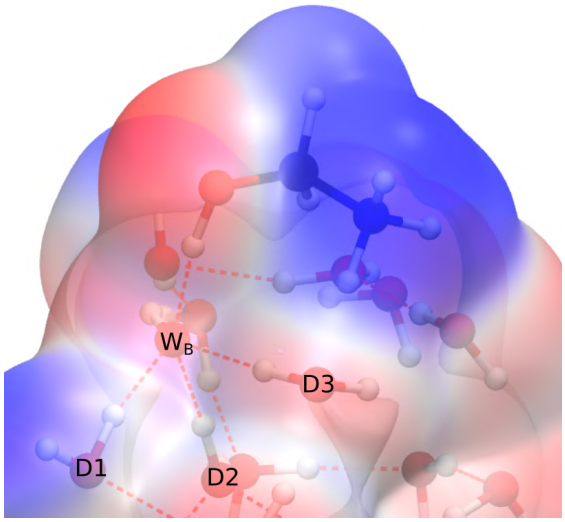


Figure E.3: MEP of the second hidden intermediate in S1-C3-C.

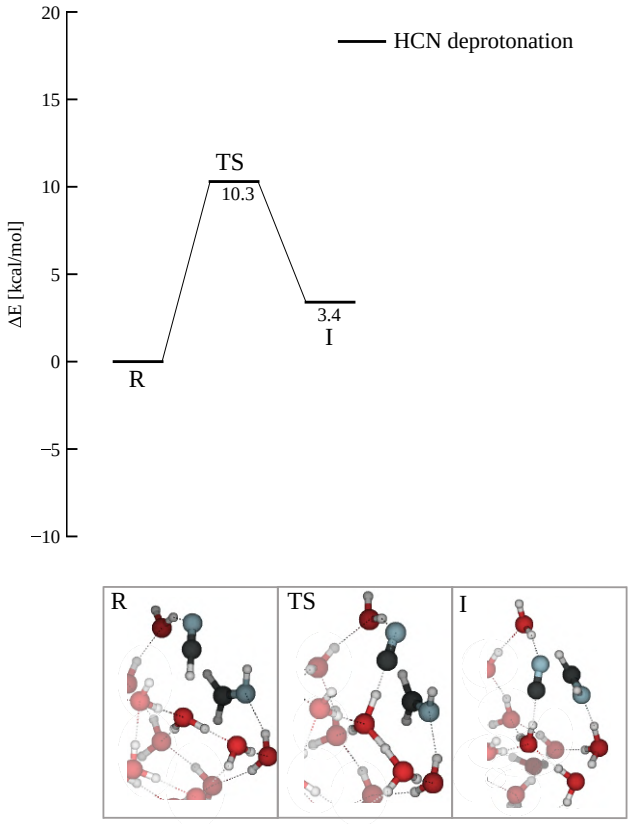


Figure E.4: Energy diagram for HCN deprotonation in a pair dangling-H site, computed at ωB97-M/def2-TZVP level of theory, using PW6B95-D4/def2-SVP geometries.

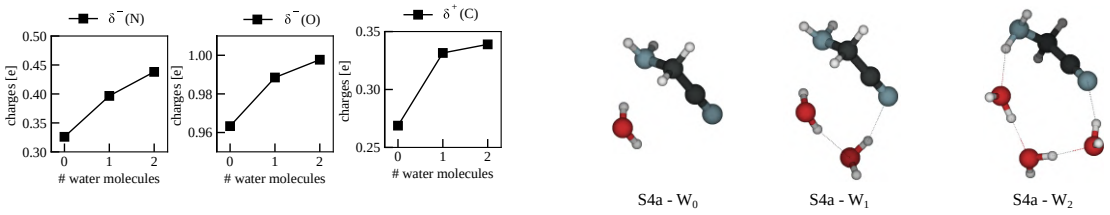


Figure E.5: Change in NAO charges along the reaction coordinate for S4a, small systems.

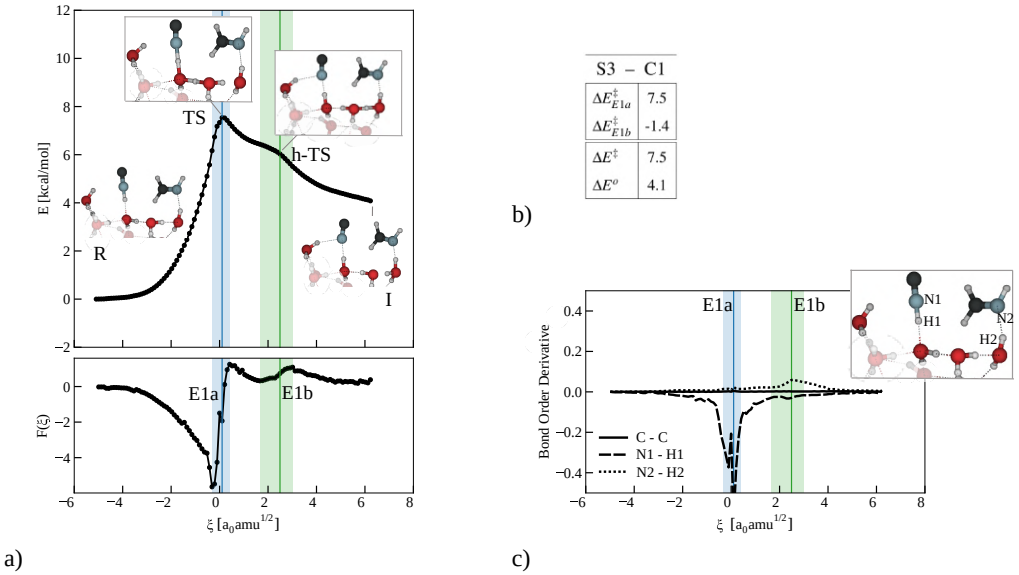


Figure E.6: a) Energy (upper panel) and reaction force profiles (lower panel) for S3-C1, using BMK/def2-TZVP//PW6B95-D3BJ/def2-SVP levels of theory. Blue and green lines represent TS and h-TS respectively. TS and h-TS regions are displayed as blue and green shadowed areas. R, TS, h-TS and P structures have been reported as extracted from the energy profile. b) Table that reports TS energy ($\Delta E^{\ddagger}$) and reaction energy (ΔE^{o}) extracted from the IRC, as well as the partition of the barrier: $\Delta E^{\ddagger}_{1a}$ is relative to the TS, while $\Delta E^{\ddagger}_{1b}$ (h-TS) has been included for completeness. c) Bond order derivative for the main bond distances involved.

F. Pubblications

1. Bovolenta, G. M.; Vogt-Geisse, S.; Bovino, S.; Grassi, T. Binding Energy Evaluation Platform: A Database of Quantum Chemical Binding Energy Distributions for the Astrochemical Community. *ApJS* **2022**, 262, 17.

2. Bovolenta, G.; Bovino, S.; Vöhringer-Martinez, E.; Saez, D. A.; Grassi, T.; Vogt-Geisse, S. High level ab initio binding energy distribution of molecules on interstellar ices: Hydrogen fluoride. *Mol. Astrophys* **2020**, 100095.

3. Grassi, T.; Bovino, S.; Caselli, P.; Bovolenta, G.; Vogt-Geisse, S.; Ercolano, B. A Novel Framework for Studying the Impact of Binding Energy Distributions on the Chemistry of Dust Grains. *A&A* **2020**, 643, A155.

G. Conferences

Jun 2023 "ICQC2023, International Congress of Quantum Chemistry", Bratislava, Slovakia.

Jun 2023 "ACO final conference: Chemical Processes in Solar-type Star-Forming Regions", Toulouse, France

Oct 2022 "QCMM 2022. Quantum Chemistry and Molecular Modeling Symposium 2022", *Universidad de Concepcion*, Chile

Sep 2022 "The life cycle of cosmic PAHs 2022", *Aharus University*, Denmark

Sep 2021 "Chemical Processes in Solar-Type Star Forming Regions", *Università di Torino*, Italy

Dec 2020 "Workshop Químico-Biológico Generación 2019", *Universidad de Concepcion*, Chile