Learning quantum chemistry with ORCA on nanoHUB

Nicolas Onofrio[1*] and Alejandro Strachan[1]

Contents

*nonofrio@purdue.edu

[1]School of Materials Engineering and Birck Nanotechnology Center Purdue University, West Lafayette, IN 47906 USA

Keywords: Density functional theory, Hartree-Fock theory, Electronic correlation, Hybridization, Normal mode analysis

1 Introduction

This document introduces the features of the ORCAtool on nanoHUB. We first present an overview of the tool then we show how to use the tool for academic purpose along a series of examples. Some additional informations can be found in the Appendix at the end of the document.

2 Overview of the tool

ORCA [1] is an ab initio, density functional theory, semi-empirical SCF-MO package written by Frank Neese[1]. The principal levels of theory implemented in ORCA range from Hartree-Fock to DFT and multi-reference model. The actual nanoHUB tool includes HF, DFT, complete active space and coupled-clusters methods with various exchange correlation functionals and basis sets. The tool allows geometry optimizations, potential energy surface calculations, automatic ionization potential calculations and normal mode analysis. We describe the graphical user interface in the following sections.

2.1 Basic input

The basic input tab has been designed for beginner running the tool for the first time and willing to perform a quick simulation and go over the outputs. Therefore, the user only needs to select the property he wants to extract from the "task" dropdown menu and the system of interest from the "Molecule" dropdown menu (see Figure 1). The pre-built examples include the hydrogen molecule, water, hydrogen cyanide, methyl-methacrylate and some hydrocarbons (ethane and ethene). All combinations between

[1]neese [at] thch.uni-bonn.de

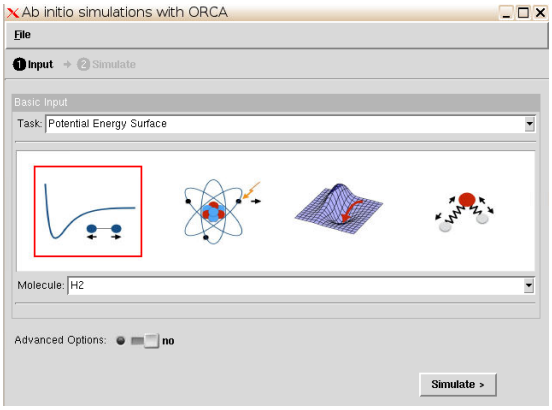

Fig. 1: Basic input tab

property and molecule will run and produce some outputs, however for a better understanding and advanced usage, details of the calculation can be displayed in separated tabs by turning on the advanced options. We will discuss these options in the next subsections.

2.2 Geometric input

The geometric input tab (Figure 2) defines the coordinates of the molecular system and its electronic state. The "Coordinates" dropdown menu on top describes the type of coordinates used to define the molecule: cartesian or internal (Z-matrix). The field below the dropdown menu defines the atomic coordinates. Cartesian coordinates are defined as usual as the element type, the x-coordinate, the y-coordinate and the z-coordinate of each atom in the molecule. Details about internal coordinates are presented in Appendix A. The last two boxes of the tab specified the spin state of the molecule and its total charge. The spin state is defined as $2S + 1$ hence the singlet state is 1 and the triplet 3. A total spin of 1/2 (as in hydrogen) corresponds to the spin state 2.

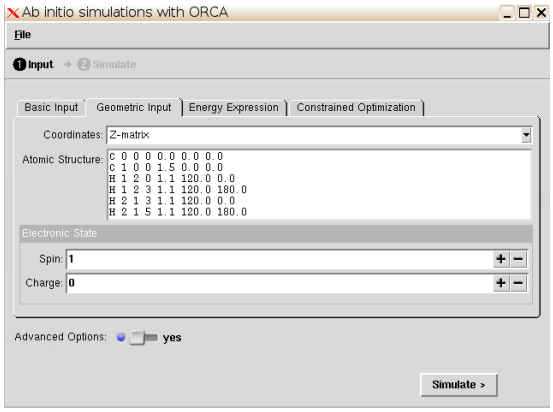

Fig. 2: Geometric input tab

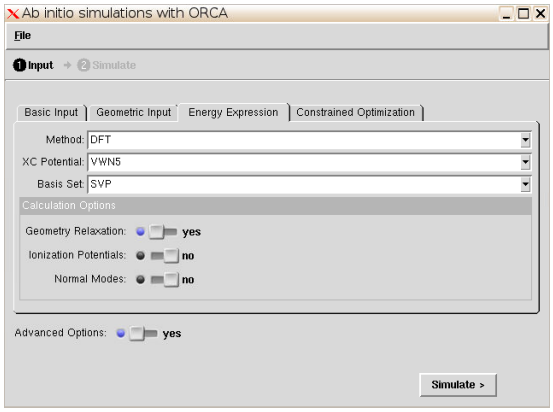

Fig. 3: Energy tab

2.3 Energy expression

The energy tab tells us about the level of theory to use and the options for the calculation. The user can choose between various methods, exchange correlation potentials (for DFT) and basis sets. If the user decides to perform a complete active space calculation, he can specified the number of orbitals and electrons to be included. The "Calculation options" group includes geometry relaxation, the calculation of ionization potentials and normal mode analysis.

- If "ionization potential" is selected, the code will compute three energies: the molecule and the molecule plus and minus 1 electron. The output will display histograms of the electron affinity and the ionization energy
- If "normal mode" is selected, the code will automatically compute the vibrational frequency and the decomposition of the normal mode analysis. The output will display the IR spectrum and the molecular sequences of each mode

2.4 Constrained optimization

This last tab contains the definition of constrains during a geometry optimization (scan parameters in future version). Therefore, in order to have access to these settings, the user must have selected "geometry relaxation" in the energy tab. Moreover, the constrains can only be defined from internal coordinates hence the "Z-matrix" choice has to be set in the geometry tab. Two types of constrains can be added to a geometry relaxation:

- Up to 4 rigid constrains can be set to the molecule by changing the integer in the "number of constrain" box. The user must specify in the "constrain #" box which bond/angle or dihedral he

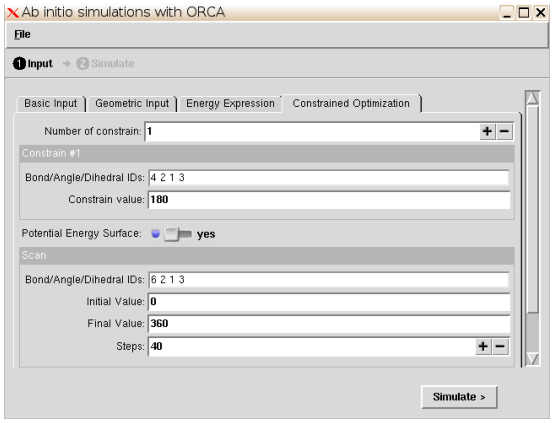

Fig. 4: Constrain tab

desires to constrain and the corresponding value. The definition of the constrain can be a series of two, three or four integer depending if the constrain applies to a bond, an angle or a dihedral angle, respectively. The atom-IDs start at 1 for the first atom to N the last atom in the Z-matrix definition.

- If "potential energy surface" is turned on, the code will perform a relaxed scan along the bond/angle/dihedral defined in a similar way as the rigid constrain. Additionally, the user needs to input the initial and final value of the constrain and the number of steps to perform the scan.

2.5 Outputs

The outputs depend on the property selected initially for the calculation. Here are some details about the outputs for each of the property computed:

- If "PES" is selected, the output contains the potential energy surface, the molecular structure along the scan
- If "Ionization potentials" is selected, the output displays histograms representing the electron affinity and the ionization energy
- If "Geometry optimization" is selected, the output includes the energy as a function of the optimization steps and the molecular structure along the relaxation
- If "Normal mode analysis" is selected, the output contains the IR spectrum and an animation of molecular structures vibrational frequencies

For all cases, the output log is included in the outputs.

3 Bonding and hybridization in hydrocarbons

The carbon element can form a large number of compounds, mainly due to the various types of bonds it can create. The geometry and conformational stability of carbon-based compounds can be easily understand, in simple systems, from the theory of hybridization of the carbon orbitals. In this first section, we will introduce the theory of hybridization and show how from first principle calculations the hybridization has an impact on the conformation of simple molecules.

3.1 Hybridization theory

We know from fundamental chemistry that the carbon element will form an excited state with the valence configuration $2s^1 2p^3$ in order to maximize its number of bonds. From the Lewis structure, we understand how methane will form however, the tetrahedral geometry comes directly from the hybridization of the valence orbitals. We show on Figure 5 a representation of the hybrid orbitals in carbon.

In methane, the $2s$ orbital combines with the $3 \times 2p$ to form four hybrid sp^3 orbitals. In order to lower the energy of the molecule (i.e. lower the interaction between electrons), the orbitals adopt a tetragonal conformation with HCH angles $\sim$109.4°. In ethane, the C-C bond comes from the overlap between two sp^3 orbitals and the bond is of σ-type. The σ bond is formed from orbitals that overlap along the axis of the bond leading to free rotation around the bond.

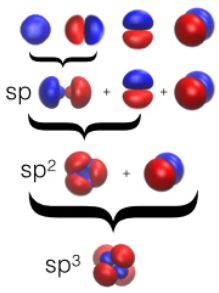

Fig. 5: Hybridization of carbon orbitals

In ethene, each $2s$ orbital of the carbon atoms hybridize with $2 \times 2p$ to form three hybrids sp^2 plus a remaining p orbital orthogonal to the plane of the hybrid sp^2. The direct overlap between the sp^2 orbitals form a σ bond and the angle between the HCH bonds is $\sim$120°. Interestingly, the remaining p orbitals that do not participate in the hybridization are at the origin of the π bond in ethene. The π bond is perpendicular to the plane of the molecule and determines the stiffness of the ethene, preventing the free rotation around the central bond.

Finally, in ethyne, each $2s$ orbital of the carbon atoms combine with $1 \times 2p$ to form two hybrid sp plus two remaining p orbitals orthogonal to the hybrid orbital and orthogonal to each other. In ethyne, a σ bond is formed between the hybrid sp and two π bonds are established perpendicular to the bond and perpendicular to each other. This last molecule exhibits a very strong and stiff double π bond also called of δ-type.

The VSEPR theory assumes that each atom in a molecule will achieve a geometry that minimizes the repulsion between electrons in the valence shell of that atom. Ethane will present two types of

conformations: (i) eclipsed when the torsion angle between HC1-C2H atoms is equal to 0, 120 and 240°; (ii) staggered when the torsion angle between HC1-C2H is equal to 60, 180 and 300°.

To sum up, we expect the following trend of bond stiffness: $\sigma < \pi < \delta$ and therefore, the dissociation energies: $E_\sigma^{\text{diss}} < E_\pi^{\text{diss}} < E_\delta^{\text{diss}}$. Moreover, the rotation around a σ bond will be easier than the rotation around a π bond and a δ bond. Note that these rules are based only on the simple analysis of the atomic orbitals and might not be generalizable to all molecules.

3.2 Stiffness of the carbon bond

We now propose to study the stiffness of the carbon-carbon bond in ethane and ethene with ORCA tool on nanoHUB.

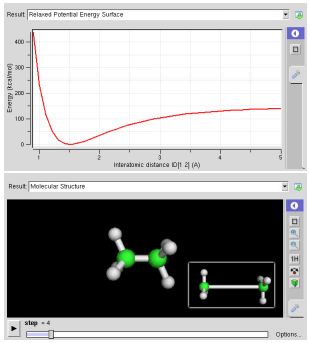

Fig. 6: Dissociation of the σ bond in ethane

You will find in the examples of the tool both hydrocarbon molecules defined in internal coordinates. We will use DFT to solve the energy problem with the exchange correlation potential B3LYP and the basis set SVP. The stiffness of the carbon-carbon bond can be estimated upon the dissociation of the bond. We performed PES of the C-C bond from 0.9 to 5.0 Å for ethene (Figure 6) and ethane (Figure 7). As expected, the dissociation energy for ethane is lower than ethene.

Interestingly, by looking at the molecular structure upon dissociation, we can see that the conformation around the carbon atoms is changing from tetrahedral to planar, in other words, the hybridization is changing from sp^3 to sp^2. In the case of ethene, at the dissociation limit, the carbon atoms make only two bonds and go back to the ground state configuration $2s^2 2p^2$, adopting a bended conformation similar to water (tetrahedral with two lone pairs of electrons).

Finally, we notice that we overestimate the dissociation energy by almost a factor of 2. We found for the dissociation energies computed by DFT $E_{\text{ethane}}^{\text{diss}} \sim 140$ kcal/mol and $E_{\text{ethene}}^{\text{diss}} \sim 240$ kcal/mol compared to the experimental values of $E_{\text{ethane}}^{\text{diss}} \sim 80$ kcal/mol and $E_{\text{ethene}}^{\text{diss}} \sim 150$ kcal/mol. This overestimation will be discussed in the last section of this document. One can pursue the problem by computing the dissociation curve of ethyne and show that the dissociation energy is even higher for a triple bond.

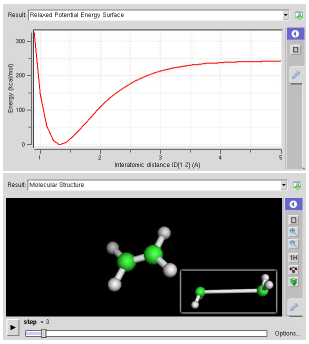

Fig. 7: Dissociation of the π bond in ethene

3.3 Torsion around the carbon bond

We have confirmed that the π bond in ethene is stiffer than the σ bond in ethane, let's now study the torsional conformation of these molecules around the central C-C bond. From the definition of the atomic coordinates in the Z-matrix form, we can perform such PES by defining the scan around the dihedral angle HCCH. Additionally for ethene, we need to impose the plane of the sp^2 orbitals to remain planar during the scan.

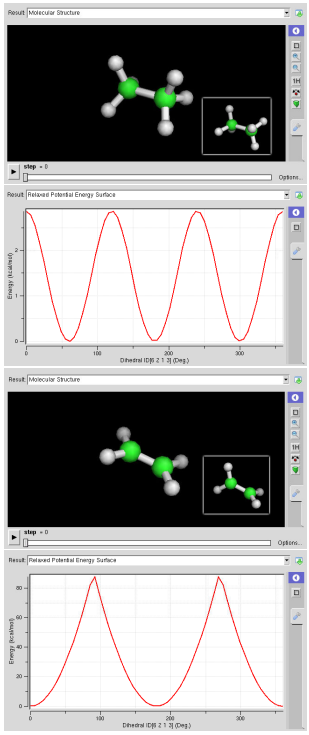

Fig. 8: Torsion around the C-C bond in ethane (top two plots) and ethene (lower two plots)

We show on Figure 8 the resulting PES for ethane and ethene. We found the PES in agreement to the VSEPR model with minima when the configuration is staggered and maxima when eclipsed. The height of the energy barriers between conformations tell us about the nature of the bond between carbons. We found barriers between the staggered and eclipse conformations equal to $E^{\text{tor}} \sim 3$ kcal/mol for ethane and $E^{\text{tor}} \sim 90$ kcal/mol for ethene. Interestingly, the rotational barriers are in good agreement with experiment. This tells us that DFT describes well molecular conformations around the equilibrium and steric effects but fails to describe properly the dissociation limit (more on this in the last Section). The steric effect between terminal hydrogens can also be confirmed by looking at the evolution of the C-C bond length during torsion. When the hydrogen atoms are eclipse, the steric effect elongates the C-C bond and when the hydrogens are staggered the C-C bond is shorter.

One can extend the study with homologs silicon and germanium molecules and confirm (or not) the validity of the hybridization and VSEPR rules, or analyze the PES of the previous molecule by swaping the hydrogen substituents with other elements (see for example Figure 9). We will see in the next section that the hybridization also plays an important role during chemical reactions.

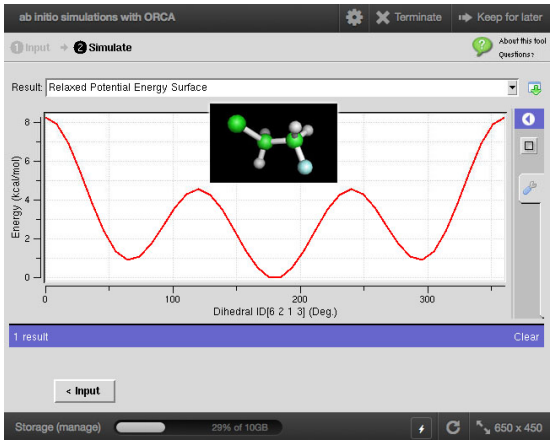

Fig. 9: Torsion around the Cl-C-C-F angle in 1-chloro 2-fluorethane

4 Chemical reaction mechanism

The nucleophilic substitution is a fundamental class of chemical reaction in which a nucleophile attacks an atom or group of atoms to replace the leaving group. This class of chemical reaction has major importance in organic chemistry because of the wide variety of products that can be synthesized including alcohols, ethers, thiols, sulfides, etc. We will describe a subgroup of nucleophilic substitution called SN2 and propose some simulations in order to demonstrate the mechanism from first principle calculations.

4.1 Theory of nucleophilic substitution

The nucleophilic substitution (SN) occurs with two different mechanism leading to two types: SN1 and SN2. We will only focus on SN2 in this document and we show on Figure 10 the general mechanism of the reaction and the corresponding potential energy surface. In the SN2 reaction, the addition of the nucleophile and the elimination of leaving group take place simultaneously. The general form of the PES is characterized by two minimia associated with the reactant and product complexes (RC and PC). The backside attack of the nucleophile can be sterically hindered by voluminous substituents. In the example shown on Figure 10 the reaction occurs through a transition state (TS) where the carbon atom is pentacoordinate. Interestingly, the carbon changes conformation from sp^3 in the R state to sp^2 in the TS state.

The height of the TS barrier depends on the nature of the halide substituents. Electronegative substituents will increase the charge of the central atom and lower the energy barrier. If the substrate under the nucleophilic attack is chiral, this often leads to inversion of configuration (R $\leftrightarrow$ L).

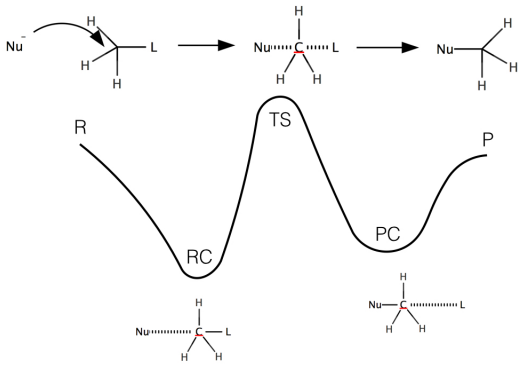

Fig. 10: Schematic representation of the nucleophilic substitution SN2

4.2 Benchmark study of SN2

The energy landscape of the SN2 reaction can be simulated using relaxed potential energy surface calculations (see Appendix A) coupled with an optimization of the TS state geometry. The TS has only one direction of downward curvature which corresponds to a first order saddle point. In a more mathematical approach, the TS has one negative Hessian eigenvalue or one imaginary frequency. The TS finder in ORCA couple with a PES will optimize the geometry at the saddle point of the energy surface and stop the scan when it encounters the maximum of the energy surface. Therefore, the tool will report the whole PES up to the saddle point (the maximum in energy) plus next point. The energy and PES parameter of the TS can be find in the energy plot in the output of the scan as well as the geometry of the TS.

Swart et al. report an extensive study of SN2 reactions from first principle calculations [2] and we will try to reproduce some of their work using the ORCA tool. In the supplementary materials of the Ref. [2] the energy of the RC, PC and TS are reported as well as internal coordinates for the atomic structures. We will start as an example with the nucleophile attack of hydride on methane. The setting for such simulation are shown in Figure 11 and the output on Figure 12.

From the PES, we verify the shape of the PES up to the TS, the minimum for the RC and the

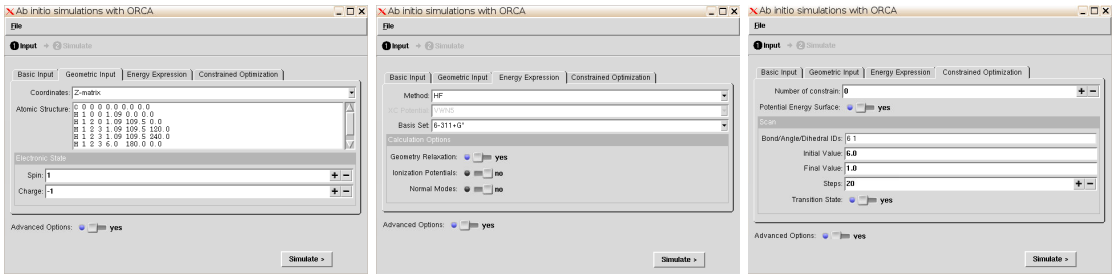

Fig. 11: Setting for the nucleophile attack of hydride on methane

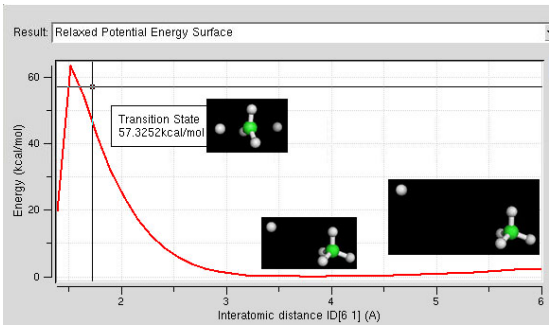

Fig. 12: Output PES and molecular structure for the nucleophile attack of hydride on methane

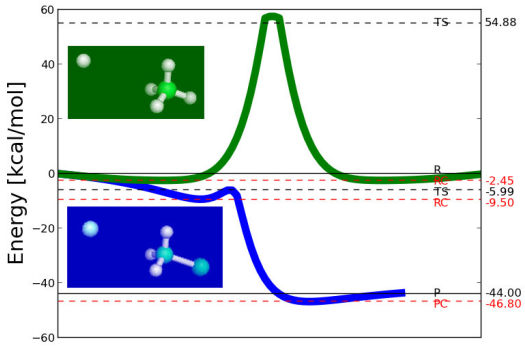

Fig. 13: Potential energy surface for the nucleophile attack of H^- + CH_4 ↔ CH_4 + H^- (green) and F^- + CH_3Cl ↔ CH_3F + Cl^- (blue) computed at the HF/6-311+G* level

sp^2 configuration of the central carbon at the TS. Because of the symmetry in the TS, this PES can be mirrored and the resulting potential energy surface of the hydride attack on ethane can be post-processed and we show a representation on Figure 13 (green). In some cases, ORCA will fail at finding the TS state mainly because of the maximum of the PES is not the maximum of the barrier (see Figure 13, blue). The tool will still report the last energy as the TS energy so it is important to look at the end of the log file to see if the calculation is completed. For the case reported on Figure 13 (blue), we found the TS starting from the addition of chlorine on fluoro methane (from right to left). For difficult cases where you can't scan the reaction backward you will have to sample the PES in two separated scans and include the TS search in the part of the scan where the TS is at the higher energy point of the PES. This occurs when the R is higher than the TS as in F^-+CH_4 ↔ CH_4+F^-. Finally, we report the energies of the TS for a variety of SN2 reaction computed at several level of theory in Table 1. If we consider the Ref. [2] as the higher level of theory, we can see that the various model reported in the Table have large uncertainties mainly due to the error in the treatment of the electronic correlation which will be discussed in the last section of this document.

Table 1: Energies (in kcal/mol) of the reactant (R), product (P), transition state (TS) and reactant/product complexes (RC and RP) for various SN2 reactions. All calculations have been performed with the 6-311+G* basis set.

SN2 reaction	method	R	RC	TS	PC	P
$H^-+CH_4 \leftrightarrow CH_4+H^-$	Ref.[a]	0.0	-2.87	48.33	-2.87	0.0
	HF	0.0	-2.45	54.88	-2.45	0.0
	B3LYP	0.0	-3.02	38.93	-3.02	0.0
$F^-+CH_3F \leftrightarrow CH_3F+F^-$	Ref.[a]	0.0	-14.62	-1.34	-14.62	0.0
	HF	0.0	-9.19	10.15	-9.19	0.0
	B3LYP	0.0	-10.44	-1.51[b]	-10.44	0.0
$Cl^-+CH_3Cl \leftrightarrow CH_3Cl+Cl^-$	Ref.[a]	0.0	-10.84	3.31	-10.84	0.0
	HF	0.0	-5.32	10.52	-5.32	0.0
	B3LYP	0.0	-6.22	2.14	-6.22	0.0
$Br^-+CH_3Br \leftrightarrow CH_3Br+Br^-$	Ref.[a]	0.0	-10.02	1.57	-10.02	0.0
	HF	0.0	-4.60	8.63	-4.60	0.0
	B3LYP	0.0	-5.62	0.74	-5.62	0.0
$F^-+CH_3Cl \leftrightarrow CH_3F+Cl^-$	Ref.[a]	0.0	-16.56	-13.95	-43.44	-33.54
	HF	0.0	-9.5	-5.99	-46.8	-44.00

[a]Bickelhaupt et al. J. Comput. Chem. 28, 1551-1560, 2007
[b]TS not optimized

5 Electron affinity and ionization energy

6 Normal modes analysis

7 The electronic correlation

In this more advanced section we will discuss the proper treatment of the electronic correlation with ab initio methods. The limits of Hartree-Fock theory and a quick review on the configuration interaction will be introduced. We will then propose to solved the H_2 molecule by hand and with ORCA tool.

7.1 Limits of Hartree-Fock theory

In the Hartree-Fock theory, we made the approximation that each electron is moving in the average field of the nuclei and the $N-1$ other electrons. We then optimized the orbitals self-consistency by variational constrain. The main error in this approximation comes from the fact that the motion of each electron is not instantaneously correlated with others. This electronic correlation is called *dynamical correlation*. Let's build the H_2 molecule in a minimal basis of $1s$ orbitals a and b. From the LCAO principle we can develop the molecular orbitals $g = (a+b)/\sqrt{2}$ and $u = (a-b)/\sqrt{2}$. The Hartree-Fock determinant Φ_0

solution of the problem can be written as:

$$\Phi_0 = |g\bar{g}| = \frac{1}{2}\left(|a\bar{a}| + |b\bar{b}| + |a\bar{b}| + |b\bar{a}|\right) \tag{1}$$

The two first terms in the right-hand side correspond to ionic contributions while the remaining are covalent. We can write the wave function $\Phi_0 = (\Psi_I + \Psi_C)/\sqrt{2}$. The ionic and covalent contributions have the same weight and during the dissociation, these weight will not evolve. This last statement contradicts the idea that for an infinite distance, each electron sits on its original site and the ionic contribution vanishes to zero. The dissociation energy is therefore badly evaluated by the Hartree-Fock method and this type of electronic correlation is called *static correlation*.

If we consider the H_2 molecule at the equilibrium distance, the HF wave function describes well the system and the small correction to the exact energy comes from the dynamical correlation. At the dissociation limit, the HF wave function is intrinsically bad because of a lack of static correlation. We can define the correlation energy:

$$E_{corr} = E_{exact} - E_{HF} \tag{2}$$

We then try to evaluate this energy using *post Hartree-Fock* methods. The most used methods are the *configuration interaction* and the *perturbation theory*.

7.2 The configuration interaction

The idea here is to develop the exact wave function in a series of Slater determinants weighted by the c_i coefficients:

$$\Psi = c_0\Phi_{HF} + c_1\Phi_1 + c_2\Phi_2 + \dots \tag{3}$$

The c_i coefficients represent the weight of each configuration in the total wave function and fulfilled the normalization condition. Here we only know the HF determinant which is considers as the main contribution in the development. If we consider a basis of atomic orbitals in a N-electronic system, we can build the configurations $|S\rangle$, $|D\rangle$, $|T\rangle$, etc. by successive mono-, di- and tri-excitations and develop

the wave function on this configurational space. Because the full-CI isn't accessible in practical, we need to truncate the determinant basis:

$$\Psi = c_0\Phi_0 + \sum c_S \left|S\right\rangle + \sum c_D \left|D\right\rangle + \sum c_T \left|T\right\rangle + \dots \tag{4}$$

The problem is now to find the c_i coefficients and by projection of the different states on the Shrödinger equation we get the CI matrix to solve. We sometime reduce the space to an active space (CAS: complete active space) and it is the analysis of the problem that allows to the relevant inclusion of the number of orbitals and electron in the CI space.

7.3 H_2 in a minimal basis

Intuitive approach

Let's write the H_2 WF as a linear combination of the ground state and the double excited determinants:

$$\Phi_{CI} = |g\bar{g}| + c|u\bar{u}| \tag{5}$$

if $c = 1$, $\Phi_{CI} = |a\bar{a}| + |b\bar{b}|$, the WF is pure ionic and if $c = -1$, $\Phi_{CI} = |a\bar{b}| + |b\bar{a}|$, the WF is pure covalent. Therefore at the dissociation the second WF ($c = -1$) seems to be a good solution of the CI problem.

Rigourus approach

Usually, CI occurs when states with the same symmetry are put together. If we consider the problem of H_2, we found multiple singlet states and the fundamental state can be stabilized by another state via CI. From Brillouin theorem, the mono-excited determinants are not coupled with the ground state hence only the determinant bi-excited $|g\bar{g}|$ will be coupled with the ground state. The CI matrix can be written as follow:

$$\begin{pmatrix} \left\langle \Psi_0 \right| H \left| \Psi_0 \right\rangle & \left\langle \Psi_0 \right| H \left| \Psi_{g\bar{g}}^{u\bar{u}} \right\rangle \\ \left\langle \Psi_0 \right| H \left| \Psi_{g\bar{g}}^{u\bar{u}} \right\rangle & \left\langle \Psi_{g\bar{g}}^{u\bar{u}} \right| H \left| \Psi_{g\bar{g}}^{u\bar{u}} \right\rangle \end{pmatrix} \tag{6}$$

The wave function can be written as:

$$\Phi_0 = C_0 \left|\Psi_0\right\rangle + C_{g\bar{g}}^{u\bar{u}} \left|\Psi_{g\bar{g}}^{u\bar{u}}\right\rangle \tag{7}$$

The exact value of the coefficients in the previous equation which describe the wave function Φ_0 and the value of the exact energy $\langle\Phi_0| H |\Phi_0\rangle$ can be found by diagonalizing the CI matrix. We already know the matrix elements $\langle\Psi_0| H |\Psi_0\rangle = E_0 = 2h_{gg} + J_{gg}$ and $\langle\Psi_{g\bar{g}}^{u\bar{u}}| H |\Psi_{g\bar{g}}^{u\bar{u}}\rangle = 2h_{uu} + J_{uu}$. The cross term $\langle\Psi_0| H |\Psi_{g\bar{g}}^{u\bar{u}}\rangle = \langle g\bar{g}| |u\bar{u}\rangle = K_{gu}$ then the CI matrix can be written as:

$$\begin{pmatrix} E_0 & K_{gu} \\ K_{gu} & 2h_{uu} + J_{uu} \end{pmatrix} \tag{8}$$

We use an intermediary normalization: $\Phi_0 = |\Psi_0\rangle + c\,|\Psi_{g\bar{g}}^{u\bar{u}}\rangle$ with $E = E_0 + E_{corr}$ and $2\Delta = 2h_{uu} + J_{uu} - E_0$. We need to solve the following matrix:

$$\begin{pmatrix} 0 & K_{gu} \\ K_{gu} & 2\Delta \end{pmatrix} \begin{pmatrix} 1 \\ c \end{pmatrix} = E_{corr} \begin{pmatrix} 1 \\ c \end{pmatrix} \tag{9}$$

which give for the energy and the coefficients:

$$E_{corr} = \Delta - \sqrt{\Delta^2 + K_{gu}^2} \qquad\qquad c = -\frac{K_{gu}}{\Delta + \sqrt{\Delta^2 + K_{gu}^2}} \tag{10}$$

At the dissociation ($R \to \infty$) we have:

$$h_{uu} = h_{gg} \qquad\qquad J_{gg} = J_{uu} \tag{11}$$

Therefore, $\Delta = 0$ and $c = -1$!

7.4 Solving H_2 with ORCA

In this section we will approach the electronic correlation problem with an extended basis set using the ORCA tool on nanoHUB. We will perform a series of relaxed potential energy surfaces (see Appendix A) at various level of theory. Perform PESs of H_2 dissociation at the HF, DFT/B3LYP and CASSCF(2,2) level with SVP basis set. We will scan the bond between hydrogens from 0.4 to 5.0 Å within about 50 steps (more steps will produce a smoother curve). A pre-built structure of hydrogen molecule is already present in the examples of the tool, only the energy expression and constrained optimization tabs have to be modified. Snapshots of the various inputs are shown on Figure 14.

In the basic tab, the user can pick a pre-built molecule, here we choose H_2 and in order to see the details of the simulations we turn on the "Advanced Options" button. The geometry tab is unchanged and the specifications for the PES scan have to be set in the "Constrained Optimization" tab. The series of simulations proposed to study the electronic correlation can be obtained by changing the method in the "Energy" tab.

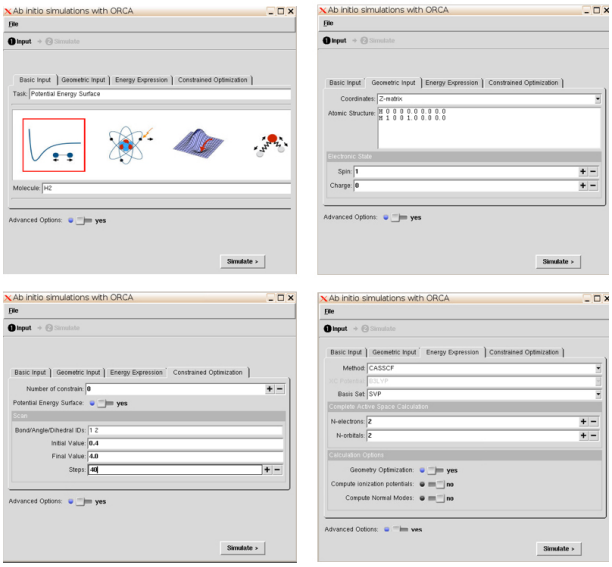

Fig. 14: Snapshots of the input parameters.

The outputs of a PES are composed by the energy curve, the molecular structure of the molecule along the PES path and the output log file. We will now discuss the energy dissociation curves computed with various methods. The energy curves are represented on Figure 15. The KS (DFT) and HF formalisms are based on a single determinant hence we expect that the KS method fails at the dissociation limit.

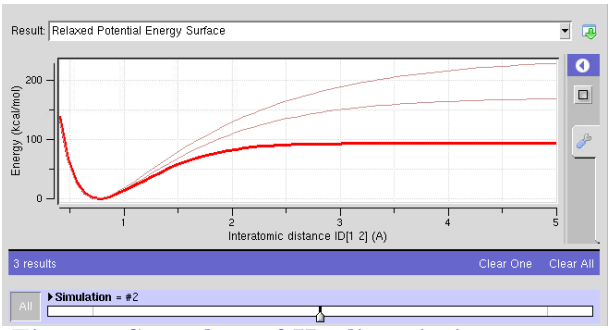

Fig. 15: Snapshot of H_2 dissociation curves.

However, the KS orbitals are different form the HF orbitals because the former include the effect of the electronic correlation (in the XC potential or functional). Therefore, it would be possible to describe the correct dissociation of the H_2 molecule if we known the exact exchange-correlation (XC) potential. However we see that even with the XC correction, the dissociation energy of H_2 is still over-estimated by DFT. As it has been discussed before, the CASSCF method is based on configuration

interaction and describes quantitatively the dissociation limit. We found a dissociation energy of 94 kcal/mol in good agreement with the experimental value of 104 kcal/mol. This value can be improved with a larger basis set and active space. It is interesting to analyze the weight of the determinants in the wave function reported in the output log. You will find that around the equilibrium distance, the CI wave function is composed only by the determinant $|g\bar{g}|$ (50% ionic and 50% covalent) while upon dissociation the configuration $|u\bar{u}|$ is equally weighted (with a negative sign of the coefficient leads to 100% covalent).

Appendix A. Relaxed potential energy surface

During a constrained geometry optimization, the full geometry is optimized but one (or many) bond(s), angle(s) or dihedral angle(s) are kept fixed during the relaxation. In order to define the bond (angle/dihedral) that will be fixed during the optimization, we need to define the initial geometry in internal coordinates. Internal coordinates are specified in form of a *Z-Matrix*. A Z-Matrix contains information about the molecular connectivity, bond lengths, bond angles and dihedral angles. In internal coordinates, we define the geometry as follow:

```
1 Atom1 0 0 0 0.0 0.0 0.0
2 Atom2 1 0 0 R1  0.0 0.0
3 Atom3 1 2 0 R2  A1  0.0
4 Atom4 1 2 3 R3  A2   D1
5 ...
6 AtomN NA NB NC RN AN DN
```

- NA: The ID of the atom making a bond with AtomN
- NB: The ID of the atom making an angle with AtomN and NA
- NC: The ID of the atom making a dihedral angle with AtomN, NA and NB. This is the angle between AtomN and atom NC when looking down the NA-NB axis.

R define the distance between atoms, A the angles and D the dihedral angles. For a simple example, here is the Z-matrix for H_2:

```
1   H  0 0 0 0.0 0.0 0.0
2   H  1 0 0 1.0 0.0 0.0
```

The constrained geometry optimization method allows to generate *potential energy surfaces* (PES) like bond dissociation curves. A PES is defined as a series of constrained geometry optimizations in which a bond (angle/dihedral) will be scanned while all others are relaxed. In ORCA, we need to define the bond (angle/dihedral) that we want to scan, the initial/final values of the bond (angle/dihedral) and the number of steps in the scan (by default will be equidistant). A screenshot of the settings for a relaxed surface scan (also called PES) with ORCA tool is showed on Figure 14.

References

[1] Frank Neese. The ORCA program system. *WILEY INTERDISCIPLINARY REVIEWS-COMPUTATIONAL MOLECULAR SCIENCE*, 2(1):73–78, JAN-FEB 2012.

[2] M Swart, M Sola, and FM Bickelhaupt. Energy landscapes of bimolecular nucleophilic substitution (SN2) reactions: A comparison of density functional theory and coupled cluster methods. *ABSTRACTS OF PAPERS OF THE AMERICAN CHEMICAL SOCIETY*, 230:U1287–U1288, AUG 28 2005. 230th National Meeting of the American-Chemical-Society, Washington, DC, AUG 28-SEP 01, 2005.