Efficient structure optimization methods for large systems and their applications to problems of heterogeneous catalysis

D I S S E R TAT I O N

zur Erlangung des akademischen Grads

doctor rerum naturalium
(Dr. rer. nat.)
im Fach Chemie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

Von

Herrn Magister Andrzej Piotr Niedziela

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. paed. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

Gutachter:
1. Prof. Dr. Dr. h.c. Joachim Sauer
2. Prof. Dr. Marek Sierka

eingereicht am:
Tag der mündlichen Prüfung: 04.12.2015

Table of Content

Table of Content 3
1. Introduction 5
2. Local Optimization 7
2.1. Potential Energy Surface 7
2.2. Coordinates 9
2.2.1. Natural Internal Coordinates 11
2.2.2. Redundant Internal Coordinates 12
2.2.3. Delocalized Internal Coordinates 14
2.3. Optimization Methods 16
2.3.1. Direct Methods 18
2.3.2. Gradient Methods 19
2.3.3. Newton Methods 20
2.3.4. Constrained Optimization 22
3. Global Optimization 25
3.1. Methods Review 26
3.2. Genetic Algorithm 32
4. Rigid Body Genetic Algorithm 39
4.1. Introduction 39
4.2. Methodology 42
4.3. Implementation 45
4.4. Rigid Body Genetic Algorithm Performance 46
5. Case Study: Methane, Ethane, Propane on MgO(001) 51
5.1. Motivation 51
5.2. MgO(001) Surface Model 56
5.3. Methods 57
5.4. Optimization Strategy 59
5.5. Thermodynamic Properties - Microscopic–Macroscopic Connection 60
5.6. Interaction Energy 62
5.7. Methane at the MgO(001) Surface 64

5.7.1. Single Molecule Coverage 65
5.7.2. Monolayer Coverage 72
5.7.2.1. Structures 72
5.7.2.2. Energies 75
5.7.3. Higher Coverage 78
5.7.3.1. Methane at Bare Surface 78
5.7.3.2. Methane at the ROT Structure 86
5.8. Ethane at the MgO(001) Surface 92
5.9. Propane at the MgO(001) Surface 96
5.10. Comparison with Experiment 101
6. Summary 107
7. Zusammenfassung 111
8. Outlook 117
9. Bibliography 121
10. Publications 133
11. Acknowledgements 135
12. Selbstständigkeitserklärung 137

1. Introduction

Structure optimization is an important, and usually first, step of most quantum chemical investigations. It is a major component of all computational chemistry studies that are concerned with the structure and reactivity of molecules. A number of structure optimization methods exists. Those methods are used to optimize equilibrium geometries, locate transition structures (TSs), and follow minimum energy paths (MEPs), which correspond to reaction paths. Quantum chemistry is concerned with optimization methods that are applicable to electronic structure calculations.[1, 2] Due to the high computational load of these calculations, structure refinement algorithms need to be fast, efficient and robust. Most major electronic structure packages have a selection of structure optimization algorithms.[3-11] The choice of the algorithm is usually a trade-off between accuracy and speed of calculations, and is bound to the physical meaning of the problem in question, the size of the model used to represent the real–life system, and a level of theory *i.e.* accuracy to which one wants to treat physical effects. Numerous reviews on structure optimization are available.[12-20] Global optimization and conformational searching present problems that are even more difficult.

Global minima are of great interest in quantum chemistry as they correspond to the configurations of atoms with the lowest total energy. These structures are the most likely to be found in the experiment, although this depends on experimental conditions, as some of the products might be kinetically, not thermodynamically, preferred structures. For a large enough system, the configuration can be difficult to determine, as the number of local energy minima rises exponentially with the size of the system. This rapid growth in the number of possible minima causes that already for small systems simple search methods may not exhaust all possible configurations and lead to biased results. Therefore global optimization methods, capable of searching for most stable configuration in an unbiased way, need to be introduced. One of such methods is the Genetic Algorithm method. However, no method guarantees to find the global minimum in finite number of steps.

This thesis presents the implementation of the rigid body optimization, within a global optimization scheme of Genetic Algorithm (GA), and its utility for global search of

hydrocarbons adsorption on the MgO (001) surface. The presented scheme allows for reduction of computational effort needed to localize global minima, and proves to perform well for weakly-bound systems. The method was used to study methane, ethane, and propane adsorption on the MgO (001) surface. The obtained adsorption energies were related to previous theoretical investigations, as well as to experimental results. The comparison yielded a good agreement.

The second chapter reviews topics connected with local optimization; importance of the Potential Energy Surface concept; the difference between coordinates systems; and comparison of optimization methods based on energy and gradient calculations. In the third chapter a number of methods for global optimization is discussed. Different techniques, using stochastic and deterministic approach are presented. Emphasis is put on evolution of genetic algorithm for clusters and periodic systems. In chapter number four, the concept of the Rigid Body Genetic Algorithm is presented. The chapter explain the main advantage of the method, i.e. optimization space reduction, and describes methodology of rigid body approximation together with implementation of this local optimization scheme within a genetic algorithm. In addition, performance of the Rigd Body Genetic Algorithm is presented. In the following chapter, chapter number five, the Rigd Body Genetic Algorithm is used to study hydrocarbons adsorption on the MgO (001) surface. Three different hydrocarbon molecules; methane, ethane, and propane, with various loadings were investigated. Obtained global minima and adsorption energies were compared with previous theoretical and experimental results. Chapter six and eight summarize the presented method and its results, and gives an outlook for possible directions in which the RBGA algorithm can be further develop.

2. Local Optimization

A fast progress in optimizations methods for *ab initio* molecular-orbital calculations over last half-century has been observed. This progress has come about primarily because of two simultaneous circumstances: the development in algorithms and growth of computational power; as well as the introduction of energy gradient techniques.[21-23] Analytical gradient-based optimization techniques are one order of magnitude faster than optimization algorithms based on energy alone. As a result, optimizing equilibrium geometries is now a routine even for fairly large systems and finding transition structures has become manageable.

2.1. Potential Energy Surface

The molecular structure is specified by the relative position of all atoms in the molecule. For a given position, and in particular electronic state, the molecule has an unequivocal energy. This energy varies as a function of electronic state and atomic coordinates, and is described by the potential energy surface (PES). Figure 1 presents a simplistic representation of a potential energy surface of two geometric variables.

The concepts of potential energy surfaces arise from the Born-Oppenheimer approximation,[24] which treats motions of nuclei and electrons separately. Since the nuclei are much heavier than electrons, they move much slower. This allows for separation of nuclei movements and electronic motions. Therefore, the electronic energy of a molecule can be obtained by determining the electronic structure for a fixed set of nuclear positions. This optimization can be performed for any nuclear configuration, giving a potential energy surface – a parametric function of the energy with respect to nuclear coordinates. Nevertheless, there are circumstances, like surface crossings, where non-adiabatic effects are important and the Born-Oppenheimer approximation is not valid anymore. In these cases classical meaning of chemical structures becomes less clear and special optimization methods may be necessary.

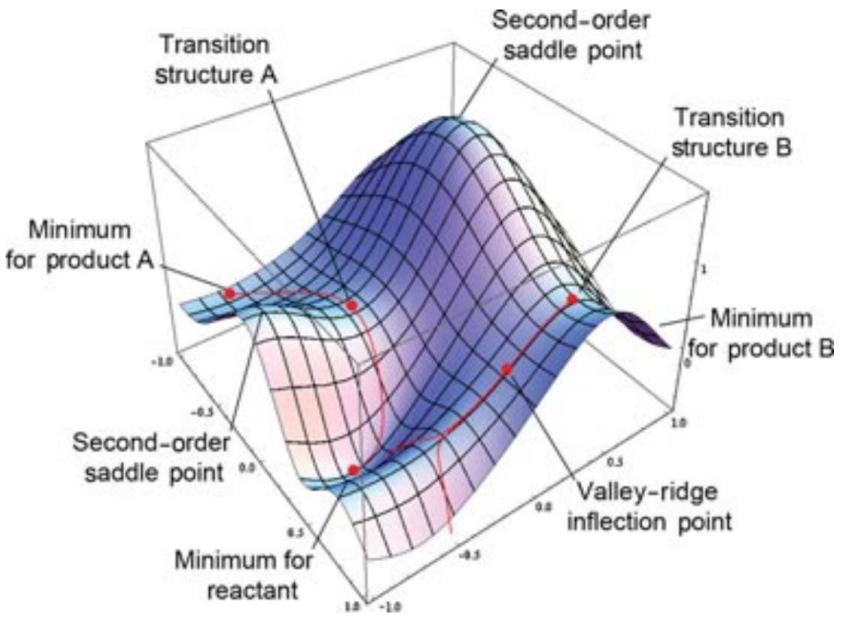

Figure 1. Model of PES showing its interesting features; minima, saddle points, and inflection points. (Reprinted Schlegel[15])

The potential energy surface is characterized by its stationary points: minima, maxima, and saddle points.[25] Each point represents different states of the molecule, and can be identified in terms of its first and second derivatives. The first derivatives of energy with respect to coordinate of each atom form a vector called gradient. The mixed second derivatives form a matrix called the Hessian. Each stationary point is characterized by vanishing gradient. Additionally, in a minimum the Hessian is positive definite (all eigenvalues of Hessian matrix are positive), in a first-order saddle point Hessian matrix has only one negative eigenvalue. Chemically speaking – minima identify stable structures, whereas first-order saddle points can be related to transition states (TS). Higher-order saddle points on the potential energy surface are of no interest for structure optimization purposes (there are relevant for electronic structure calculations). The steepest descent reaction paths (SDP) from the transition state to the minima on both sides of the saddle point form the minimum energy path (MEP) or the intrinsic reaction coordinate (IRC), when mass-weighted coordinates are used.

Taylor expansion allows to represent the potential energy surface, $E(\boldsymbol{x})$, as an infinite sum in neighbourhood of a point $\boldsymbol{x_0}$, using the step $\boldsymbol{x}$ and gradient $\boldsymbol{g^T}$ vectors, and Hessian matrix, $\boldsymbol{H}$,

$$E(\boldsymbol{x}) = E(\boldsymbol{x_0}) + \boldsymbol{g^T x} + \frac{1}{2}\boldsymbol{x^T H x} + \ldots \quad (1)$$

Most optimization methods are constructed based on this representation.

2.2. Coordinates

Before a structure optimization can be carried out, an appropriate coordinate system must be chosen. The choice of the coordinate system is as important as the choice of the optimization strategy, and can easily influence the numerical stability and convergence of the algorithm. Any complete set of coordinates can be used as a representation of a molecular structure and its potential energy surface. Nevertheless, the nature of the quadratic approximation (equation (1)) to the potential energy surface gives insight for favourable features of a good coordinate system. If the Hessian has a mixture of very large and very small eigenvalues (*i.e.* it is an ill-conditioned matrix), the optimization will be inefficient, due to the presence of very stiff coordinates and very flexible coordinates, corresponding to these eigenvalues. Introduction of a different coordinate system (like the Rigid Body coordinates) alters the spectrum of the Hessian matrix, and therefore affect the stability and performance of the optimization. Another problem that can slow down the optimization process is the strong coupling between coordinates. This occurs if the off-diagonal Hessian matrix elements are of the same magnitude as the diagonal ones. However, it is usually possible to find a different set of coordinates that reduces the coupling. Strong anharmonicity (*i.e.* all terms beyond 2nd order in equation (1)) can also have a strong influence on the optimization performance. If the composition of the Hessian matrix changes rapidly during the optimization, the quadratic expression used to model the potential energy surface is poor, and convergence will be slow. A good choice of the coordinate system can affect anharmonicity of the PES as well as the eigenvector structure of the Hessian.

The major advantage of all internal coordinate systems in structure optimization is that they significantly reduce the harmonic and anharmonic couplings between different coordinates. As a result a much smoother optimization is observed, with more rapid convergence in comparison to a more coupled set of coordinates, like the Cartesian coordinates. Natural internal coordinates are in general also preferred over the Z-matrix coordinates. The consideration of individual angles and torsions, which describe the molecular structure, can introduce strong coupling between the coordinates. This occurs in poorly constructed Z-matrices, when arbitrarily coordinates were omitted to avoid redundancy. However, internal coordinates also have disadvantages. Additional measures need to be taken to transform gradients and possibly Hessians, originally calculated in Cartesian coordinates, to the corresponding internal coordinate quantities. This involves construction of the transformation

matrix (the Wilson B-Matrix [26]), its inversion, and a number of additional matrix multiplications, like the iterative back-transformation of a displacement in internal coordinates into Cartesians coordinates, to update the geometry and calculate the energy and gradient for the next optimization cycle. The back-transformation is the single most time consuming component of the optimization cycle, and scales cubically $O(N^3)$, in terms of CPU time, with the number of primitive internal coordinates. This dependency becomes dominant as the system size increases.

A number of coordinate systems is typically available to be used for structure optimization. The cartesian coordinates are naturally the first choice. The great advantage of them is that most energy and derivative calculations are carried out in them. Unfortunately, they do not reflect the 'chemical structure' of a molecule, and therefore are not well suited for structure optimization (strong coupling between x, y, z coordinates). The internal coordinates, like bond lengths, valence and torsion angles, offer a much clearer description of the molecular structure, and are, therefore, a better choice for structure optimization. The couplings between these coordinates are usually much smaller than between the Cartesian coordinates. Furthermore, the curvilinear character of internal coordinates is much better suited for representing motions as valence angle bending and rotation about single bonds. It is easy to select a set $3N - 6$ internals coordinates ($3N - 5$ for linear molecules). Z-matrix is an example of such a set of coordinates. [27] Z-matrices are constructed using valence-type coordinates, *i.e.* by using bond lengths, bond and dihedral angles. The transformation of geometries and derivatives between Z-matrix and Cartesian coordinates is straightforward.[28] Problems occur with cyclic molecules. For acyclic molecules, the set of all bonds, angles, and torsions represents the intrinsic connectivity of the molecule, for a cyclic molecule, the same set of all possible primitive coordinates introduces more than the $3N - 6$ coordinates which are required to define properly the geometry of the molecule. Such a set of coordinates is redundant in some geometric parameters.[28-40] Because only $3N - 6$ of these redundant internal coordinates can be used to transform geometry back to Cartesian coordinates (in Z-matrix coordinates case), other coordinates must be left out from the optimization space during the optimization process. It is rather arbitrary which coordinates should be drop out. Optimizations in Z-matrix coordinates are in general significantly less efficient in comparison to these in internal coordinates.

The set of all primitive coordinates (bonds, valence angles, torsions, additionally augmented by out-of-plane bends and linear bends if necessary) creates a *primitive redundant coordinate system*.[28, 33]

2.2.1. Natural Internal Coordinates

Sometimes it is advantageous to form linear combinations of the primitive redundant internal coordinates to create symmetry-adapted redundant internal coordinates as well as natural or delocalized redundant internal coordinates.[29-32] Additionally, for periodic systems like solids or surfaces, unit cell parameters need to be added.[38-40] This can be done either explicitly or implicitly by defining coordinates that cross the boundaries of the unit cell. Additional coordinates for molecules in nonisotropic media are needed to specify the orientation of the molecule. Also for systems containing more than one fragment, additional coordinates that specify the positions of the fragments relative to each other may be required. The redundant internal coordinates for the reactants and products represent usually a good coordinate system for TS optimization.[33] The transformation of Cartesian coordinates and derivatives to redundant internal coordinates is straightforward. The back transformation of a displacement in redundant internal coordinates to Cartesian coordinates usually is solved using an iterative procedure. [28, 29, 31, 35, 37, 41, 42]

Natural internal coordinates, also referred to as natural valence coordinates, were first introduced to structure optimization by Pulay *et al.*[21, 29, 43] The natural valence coordinates are local, eliminate most redundancies, and conform to local pseudosymmetry. They constitute a complete and nonredundant set of internal displacement coordinates, and therefore assures that in these coordinates Taylor expansion representation of the molecular potential energy (equation (1)) is unique. Character of coordinates originates from vibrational spectroscopy. This involves use of individual bond displacements as stretching coordinates, and linear combinations of bond angles and torsions as deformational coordinates. Choice of linear combinations of bends and torsions is based on local pseudosymmetry using group theory arguments. Construction of the coordinates enables the reduction of both harmonic and anharmonic coupling terms in the potential function in a purely geometrical way. For ring systems, extra attention has to be paid to eliminate redundancies between stretching and bending coordinates. This facilitates the comparison of ring rigidity. The original construction of the deformational coordinates for ring systems involves the whole ring and is therefore nonlocal.

Displacements in the internal coordinates $\boldsymbol{q}$ are related to the Cartesians displacements by the Wilson $\boldsymbol{B}$ matrix,[26]

$$\Delta \boldsymbol{q} = \boldsymbol{B}\Delta \boldsymbol{x} \quad (2)$$

Similarly, the Cartesian forces, $\boldsymbol{g_x}$, are related to the internal forces, $\boldsymbol{g_q}$, by the same matrix

$$\boldsymbol{g_x} = \boldsymbol{B^T g_q} \quad (3)$$

To express $\boldsymbol{g_q}$ from equation (3) the $\boldsymbol{BB^T}$matrix cannot be singular, so that it is possible to construct general inverse matrix, $\boldsymbol{B^{-1}}$,

$$\boldsymbol{B^{-1}} = (\boldsymbol{BB^T})^{-1}\boldsymbol{B} \quad (4)$$

To obtain internal forces,

$$\boldsymbol{g_q} = \boldsymbol{B^{-1}}\ \boldsymbol{g_x} \quad (5)$$

In the case of redundancy (when the rows of the $\boldsymbol{B}$ matrix are linearly dependent), it is impossible to construct general inverse of the transformation matrix.

Nevertheless, natural internal coordinates have some disadvantages. Algorithms for automatic constructions of natural internal coordinates can be relatively complicated. For complex molecular topologies, like multiply fused rings or cages, the assigning algorithm may be unable to correctly generate a set of coordinates. Additionally, for more complex molecular topologies, where more natural internal coordinates are generated, problems of redundancy exist. This can be avoided by removing appropriate redundant coordinates. But this may be arbitrary, like in the case of omitting individual angles and torsions in a Z matrix, and may not give optimal set of natural internal coordinates, especially when a significant number of coordinates that were generated has to be excluded. The redundancy problem can be avoided by carrying out structure optimization in the redundant coordinate space.

2.2.2. Redundant Internal Coordinates

Redundant internal coordinates were also introduced by Pulay *et al.*[28] due to problems in defining physically reasonable redundancy-free natural internal coordinates for bridged polycyclic compounds and cage structures. However, Schlegel[44] was the first one who strongly emphasized the utility of redundant coordinates in geometry optimization by realizing that the strong-coupling terms in the potential function, observed for ring systems when using Z-matrix-type coordinates, can be reduced by representing the potential function (in quadratic

approximation) in redundant valence coordinates, and transforming it to a nonredundant set. The disadvantage of this method is the need of specifying two sets of coordinates and in case of a significant change in geometry the transformation should be repeated. Redundant internal coordinates do not use intermediate coordinate systems. If the coordinates are redundant, the rows of the B matrix (equation (2)) are linearly dependent. By constructing the $\boldsymbol{G}$ matrix (6) and after it diagonalization (7), the redundancy condition can be easily determined.

$$\boldsymbol{G} = \boldsymbol{BuB^T} \tag{6}$$

$$\boldsymbol{G}\,(\boldsymbol{K}\,\boldsymbol{L}) = (\boldsymbol{K}\,\boldsymbol{L}) \begin{pmatrix} \boldsymbol{\Lambda} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \tag{7}$$

The $\boldsymbol{u}$ matrix is usually the unit matrix (it has to be a non-singular matrix). In case $\boldsymbol{u}$ is a diagonal matrix, consisting of triplets of inverse nuclear masses, it is then the same as the spectroscopic spectroscopic $\boldsymbol{G}$ matrix.[26]

Diagonalization provides two matrices; $\boldsymbol{K}$ consists of the first n eigenvectors of $\boldsymbol{G}$, corresponding to nonzero eigenvalues; and $\boldsymbol{L}$ consists of the r redundant eigenvectors of $\boldsymbol{G}$ which correspond to the zero eigenvalues. These matrices are used to define displacements in nonredundant ($\boldsymbol{Q}$), and redundant coordinates ($\boldsymbol{R}$),

$$\Delta\boldsymbol{Q} = \boldsymbol{K^T}\Delta\boldsymbol{q}, \qquad \Delta\boldsymbol{R} = \boldsymbol{L^T}\Delta\boldsymbol{q} \tag{8}$$

$$\Delta\boldsymbol{Q} = \boldsymbol{K^T B}\Delta\boldsymbol{x}, \qquad \Delta\boldsymbol{R} = \boldsymbol{L^T B}\Delta\boldsymbol{x} \tag{9}$$

Both sets of coordinates are kept and used to construct the general inverse matrix

$$\boldsymbol{G^{-1}} = (\boldsymbol{K}\,\boldsymbol{L}) \begin{pmatrix} \boldsymbol{\Lambda^{-1}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{K^T} \\ \boldsymbol{L^T} \end{pmatrix} \tag{10}$$

which is used to transform Cartesian forces to forces in redundant coordinates ($\boldsymbol{g_G}$), in similar form to equation (5),

$$\boldsymbol{g_G} = \boldsymbol{G^{-1} B u}\, \boldsymbol{g_x} \tag{11}$$

By introducing projector $\boldsymbol{P}$

$$\boldsymbol{P} = \boldsymbol{G G^{-1}} = \boldsymbol{G^{-1} G} = \boldsymbol{K K^{-1}} \tag{12}$$

it is possible to in obtain inverse Hessian in redundant coordinates:

$$\boldsymbol{H_G^{-1}} = \boldsymbol{P(PH_xP)^{-1}P} \quad (13)$$

Using gradient and inverse Hessian in redundant coordinates it is possible to update the internal coordinates:

$$\boldsymbol{\Delta q = H_G^{-1} g_G} \quad (14)$$

Displacement in Cartesian coordinates can be generated from the displacement in internal coordinates. Since this transformation between the internal and the Cartesian coordinates is nonlinear, this is usually done iteratively,[43] using the first-order formula

$$\boldsymbol{\Delta x = uB^T G^{-1} \Delta q} \quad (15)$$

This process can be costly due to its iterative character and need for calculating the inverse matrix ($O(N^3)$ scaling).

2.2.3. Delocalized Internal Coordinates

An alternative form of natural internal coordinates present delocalized internal coordinates.[30] These coordinates are fully nonredundant and can be generated for any molecular topology. The scheme for generating a complete set of nonredundant coordinates is considerably simplified. Baker *et al.*[30] argue that, in a mathematical sense, the obtained set of internal coordinate is optimal. The calculation overhead is significantly reduced by carrying the structure optimization in the nonredundant coordinate space. The reduction is particularly pronounced in the transformation of the displacement from the internal coordinate space back to Cartesian coordinates. The authors also introduced constrained optimization within delocalized internal coordinates. Constraints can be imposed by the appropriate Schmidt-orthogonalization procedure. A unit vector corresponding to the primitive internal; bond length, bond angle, and dihedral angle, which should be kept constant, is projected onto the full active subspace and normalized. All other vectors are then Schmidt orthogonalized in turn to this vector. The last vector should be dropped out, due to linearly dependence on the other vectors, leaving $n - 1$ active vectors and one constraint vector. This way opens a possibility to impose constraints on individual bond angles and torsions, as in the case of the Z-matrix approach, and still uses advantages of natural internal coordinates.

The procedure starts by generating an initial set of primitive internal coordinates. Only three types of primitive internal coordinates are considered: all stretches, all planar bends, and

all proper torsions. The coordinates are obtained solely on the atomic connectivity, *i.e.*, which atoms are bound to one another. This can be determined by simple distance criteria. This procedure will typically generate many more primitive internal coordinates than are necessary to describe the $3N-6(5)$ degrees of freedom of a molecule consisting of N atoms. The Wilson $\boldsymbol{B}$ matrix (equation (2)) is then constructed with all n generated primitive coordinates. As a result many rows of the $\boldsymbol{B}$ matrix will be linearly dependent. The redundancies are taken care of by constructing $\boldsymbol{G} = \boldsymbol{B}\boldsymbol{B}^{\boldsymbol{T}}$ matrix, analogous to one in equation (6), and it diagonalization, analogous equation (7),

$$\boldsymbol{G}\,(\boldsymbol{U}\;\boldsymbol{R}) = (\boldsymbol{U}\;\boldsymbol{R})\begin{pmatrix}\boldsymbol{\Lambda} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix} \tag{16}$$

giving $\boldsymbol{U}$, a set of nonredundant eigenvectors, and $\boldsymbol{R}$ corresponding to the redundant set. Eigenvectors in both sets are linear combinations of potentially all the original primitive internals. The $\boldsymbol{U}$ set is the active coordinate set for an unconstrained optimization of a system with no symmetry. If there is any symmetry, then one or more additional coordinates can be eliminated. The active coordinates will then form a subspace of the nonredundant coordinates.

Delocalized internal coordinates are superior over earlier introduced methods for generating internal coordinates. A complicated algorithm determining the molecular topology as well as generating natural internal coordinates is not needed. The generated set of internal coordinates is suitable for any molecular topology. No special consideration involving symmetry considerations are needed, which is often problematic when constructing natural internal coordinates for cages and multiply fused rings. The calculation overhead is reduced by carrying out the optimization in the nonredundant coordinate subspace

To avoid costly iterative back-transformation, but still carry out the optimization in internal coordinates, hybrid delocalized internal Z-matrix approach[31] was proposed. Determination of the optimization step is performed in delocalized internal coordinates, but at the same time the Z-matrix (constructed using individual primitive internal coordinates) is used to convert the new structure into the Cartesian coordinates. Unlike for delocalized internal coordinates, where each coordinate is a linear combination of primitive coordinates, generation of the Cartesian coordinates takes place in a one-step noniterative process, without the need to construct and invert the B-matrix. This approach allows for a highly efficient optimization in delocalized internal coordinates, but at the same time eliminates costly backtransformation (iterative $O(N^3)$ process) which is replaced by with a simple trigonometric Z-matrix to Cartesian conversion (one-step $O(N)$ process)

2.3. Optimization Methods

There is a number of methods for local minimization of nonlinear functions of many variables.[45, 46] The usefulness of a particular method is difficult to assess without considering precise applications. Nevertheless one can define the most important features of an algorithm, like: speed of convergence; numerical stability; scaling with the model size, number of cores, and memory; the overall cost of the optimization. The choice of the best optimization method depends on the nature of the objective function, the number of variables, the availability and cost of evaluating the first and second derivatives.

Local optimization techniques allow for determination of stationary points on PES. Each local minimization performs a number of steps on the PES, usually gradually leading towards lower energy, until no further minimization of the energy is possible. The minimum found this way is the closest one to the starting point – hence the name local minimum. In other words, the starting point lies within the basin of attraction[47, 48] of that local minimum. In general, the basin of attraction of a stationary point, whether it is a minimum or transition state, is a region of a model potential-energy surface for which each search starting from that region converges to one structure. These regions are a generalisation of the catchment areas considered by Mezey,[49, 50] and, among other things, depend upon the choice of the search method and coordinates.

Methods for locating minima can be grouped into three categories, depending on the derivative information used: (a) function value only, (b) function with first derivatives, and (c) function with first and second derivatives. Except force fields methods, where 2nd analytical derivatives are relatively cheap, the most efficient methods, in terms of computational cost, rely on first derivatives of the energy calculated analytically; some use also approximate second derivatives. The first analytical derivatives can be calculated at a cost comparable to the cost of the energy calculations, for most levels of theory that are routinely used for optimization. Derivative-based methods are significantly more efficient than energy-only algorithms. It is always possible to use simplex and pattern search methods [51-54] in case analytic derivatives are not available. Nevertheless, these become less efficient with higher numbers of degrees of freedom.[55, 56]

The determination of equilibrium structures (minima and transition-state structures) is one of the most important areas of applied quantum chemistry. Geometry optimization is usually the first step of each investigation. For larger systems, it is only practical using

analytical gradient techniques. All efficient geometry optimization methods are based on a local quadratic approximation to the surface and involve the Newton-Raphson technique. This formalism utilizes the second derivatives (the Hessian) but since their exact evaluation especially in *ab initio* investigation is usually costly and inefficient, they are most often approximated by updating mechanisms, which involve analytically calculated gradients. Due to rapid increase of computer power, much larger systems became tractable. This, in turn, stimulates development of optimization methods, to better use that power.

The efficiency of structure optimizations can be measured by the number of energy and gradient evaluations needed to achieve convergence. Three factors influence this process: (1) the coordinates used to describe the geometry of the system, (2) the mathematical algorithms used for optimization, and (3) the quality of the quadratic approximation (numerical accuracy and stability of gradients and Hessian).

The significance of the coordinates on the optimization procedure may not be obvious at first sight, since two coordinate systems related by a linear transformation are equivalent in all gradient-optimization methods, given that the gradient vector and the Hessian matrix are properly transformed from one system to another. On this basis it has been claimed that Cartesians are equivalent,[19, 57, 58] or even superior[59] to valence type internal coordinates. This statement[19, 57, 58] overlooks the fact that cubic and higher-order couplings may play an important role. It has been however demonstrated that for medium sized systems, with a reliable initial Hessian matrix and reasonable starting geometries, optimization in Cartesian coordinates can be just as efficient as in Z-matrix coordinates[59] or natural internal coordinates.[60] However, without the initial curvature information (*i.e.* without Hessian matrix) optimization in Cartesian coordinates is extremely inefficient, especially when the system is larger enough.[59, 60] Cartesian coordinates are also less efficient in comparison to internal coordinates when optimization starts with poor geometries, even if the exact Hessian matrix is available at every optimization cycle.[61] When available, a molecular mechanics force constant matrix is usually good and cheap choice as a starting Hessian for *ab initio* calculations.

2.3.1. Direct Methods

Minimization algorithms that do not use derivatives and only require the evaluation of the function are termed direct methods. They are described in a variety of standard textbooks on numerical analysis.[45, 46, 62, 63] Since they do not use derivatives (which are often not available for all levels of theory) these methods have the widest range of applicability. Unfortunately, the downside is that they also have the slowest convergence of all optimization methods.

The simple direct search[53] method is one of these method. By construction, it allows for solving optimization problems with no information about the gradient of the objective function. The algorithm uses a grid of points around the current point, searches for one with the lower value of the objective function than the value at the current point, and then moves the trial solution to this point. In contrast to derivative-based optimization methods, that use information about the gradient or higher derivatives to search for an optimal point, the direct search can be used to solve problems for which the objective function is not differentiable, or even is not continuous.

More sophisticated direct methods exist. For example, the simplex method, also called the Nelder–Mead method where a walk on the PES is defined by the coordinates of all points – the point with the lowest energy is moved to its mirrored image defined by the hyperplane constructed from the other point. The new point replaces the old one in the set and all procedure is repeated until distances and/or energies differences between points are smaller than the threshold. There is also a whole family of the pattern search methods,[64, 65] which use predefined patterns of points around the current point independent of the objective function.

Another example of the direct methods is the sequential univariate search (axial iteration) algorithm.[66] The sequential univariate search method cycles over the coordinates changing one of them at a time. This is performed once (or more times) for all coordinates. We calculate the energy at the initial geometry, then at two displacements along the one of the coordinates. New optimal displacement of the coordinate is found by fitting a parabola to the energy at these three points and displacing the coordinate to the point corresponding to the minimum on that parabola. Algorithm advances to the next coordinate and repeats fitting procedure. The cycle runs through all coordinates until the change in all the coordinates is small enough.

The *modified Fletcher-Powell method* is closely related to the axial iteration method, although it is actually a derivative based method, since it is using numerical derivatives. The method is significantly more efficient than the axial iteration algorithm. This is pronounced

especially for strongly coupled coordinates, where several optimization cycles over all coordinates are required. The energy is calculated at the starting geometry and for positive and negative displacements for all coordinates. A parabola is fitted for each of the coordinates giving a model of the potential energy surface. The predicted change in the coordinates is obtained by finding the minimum on the model surface. Steps are repeated until the change in coordinates is small. This procedure uses the gradients and diagonal elements of the Hessian matrix calculated numerically. The algorithm is not exact for a general quadratic surface, since the offdiagonal terms in the Hessian are neglected. There is, of course, a number of other algorithms that depend only on the function evaluation. Some of them may use derivatives evaluated numerically. These can be relatively efficient but are in general inferior to derivative-based methods. The direct methods are not used for geometry optimization in *ab initio* calculations (unless analytical derivatives are not available); they simply require too many steps.

2.3.2. Gradient Methods

If analytical derivatives are available, derivative-based methods represent a better choice. They can be significantly more efficient with better convergence properties than the function-only algorithms. However, for numerically calculated gradients the overall efficiency is usually not better than for the function-only algorithms. The gradient-type optimization algorithms approximate the potential energy surface by a quadratic expression in terms of the position, energy, and gradient. These methods sometimes use also the approximate Hessian. The initial estimate of the Hessian (or its inverse), can be also updated as the optimization proceeds, and for most methods approaches the true Hessian (or its inverse) if the real surface is quadratic. Additionally, one-dimensional minimization may also be required along each new search direction.

The overall cost, numerical stability, and convergence rate of minimization of a non-quadratic function depend on the accuracy of the line searches, as well as on the initial estimate of the Hessian and its updating scheme (if used). Of course, the nature of the function and the starting coordinates also have influence on the performance of the algorithm. Good initial estimates of the second derivative matrix, for geometry optimization, can be obtained from lower levels of theory or general concepts in chemical bonding. An accurate initial estimate of the Hessian can significantly improve the rate of convergence, but it will not affect the final,

optimized geometry. The final geometry depends only on the gradient (it goes to zero at the stationary point) and not the Hessian.

The simplest gradient method (based on a linear search) is *the steepest-descent algorithm*. In this algorithm the Hessian matrix is the unit matrix (or a constant times the unit matrix) and is not updated. Therefore the search direction is along $p_k = -q_k$. This is the direction in which the function decreases most rapidly. To achieve convergence an accurate linear search is required at each step. The function value decreases rapidly at first, but the final convergence is usually slow. The fixed metric method is closely related to this algorithm – the Hessian is a more general nondiagonal matrix that is not updated.

Another method based on a linear search is the conjugate gradient (CG) algorithm. In the conjugate gradient method, a new search direction is chosen to lower the energy but at the same time remaining at or near the minimum from the previous search direction. The idea arises from the fact that if the Hessian has coupling between the coordinates, the optimal search directions are not orthogonal (like it is in the case of steepest descent method) but are conjugated, *i.e.* $\boldsymbol{x_{new} H x_{old}} = 0$. Two of the most frequently used search directions in conjugate gradient methods are Fletcher–Reeves[67] and Polak-Ribiere. [68]

The conjugate gradient method is one of the older methods, and is suitable for very large systems where storage of the Hessian is not possible; they require less storage than limited memory quasi-Newton methods (only three vectors need to be stored). If the Hessian can be kept in memory, the quasi-Newton methods provide better convergence to the minimum.

2.3.3. Newton Methods

The Newton method is based on a local quadratic approximation of the potential energy surface. Differentiating the equation (1) with respect to the coordinates yields an approximation for the gradient, given by:

$$\boldsymbol{g(x) = g_0 + H_0 \Delta x} \tag{17}$$

By definition at stationary points the gradient must be zero ($\boldsymbol{g(x) = 0}$); equaling the right hand side of the equation (17) to zero and performing transformation we obtain the displacement (within the quadratic approximation) to the minimum:

$$\boldsymbol{\Delta x = -{H_0}^{-1} g_0} \tag{18}$$

Where $\Delta \boldsymbol{x}$ is called the Newton–Raphson step. In Newton's method, the Hessian, $\boldsymbol{H_0}$, is calculated directly at each step. In quasi-Newton methods an approximate Hessian is updated at each step of the optimization.

Variable metric (or quasi-Newton) method falls between the extremes of the steepest descent methods and Newton's method, and represents the most frequently used gradient algorithm. By replacing the calculated Hessian with an approximate one this method increases its usability for extended systems. A starting Hessian is obtained from lower level of theory or taken as a unit matrix (usually scaled), and then is updated using the gradient information gathered during the course of the optimization. The quasi-Newton methods are the most efficient and widely used algorithms, and can be employed effectively for optimizing both minima and transition states.

Real potential energy surfaces are very rarely quadratic; therefore, usually a number of Newton or quasi-Newton steps is required to reach a stationary point. In case of minimization, eigenvalues of the Hessian must be positive (*i.e.*, Hessian must be positive-definite matrix). If n eigenvalues are negative, then the step will be taken toward n^{th}-order stationary point. Therefore, without controlling the size and direction of the step, simple Newton's methods are not robust. If we aim to find a transition state, we must assure that the Hessian has only one negative eigenvalue, with the corresponding eigenvector (the transition vector) roughly parallel to the reaction path. The Hessian updating methods guarantee right properties of the Hessian matrix, depending on optimization type (non-negative definite for minimization and indefinite with one negative eigenvector for TS search).

The quadratic approximation to the potential energy surface holds only for a small region around the current point. This region can be specified by a trust radius, τ. Steps outside this region may lead to unreasonable structures. The optimization is more robust if the step size does not exceed its maximum possible value. An initial estimate of τ is usually updated during the course of the optimization. The update is based on how well the actual change of the energy on potential energy surface fit to energy change on an approximated surface from the previous step. The updating recipe is a ratio of the actual change in energy and predicted change in energy:

$$\rho = \frac{\Delta E}{\boldsymbol{g_0}^{\boldsymbol{T}} \Delta \boldsymbol{x} + \frac{1}{2} \Delta \boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{H_0} \Delta \boldsymbol{x}} \tag{19}$$

Depending on the ρ value, the trust radius can be increased, kept the same, or decreased. Exact values can differ depending on the implementation. If the Newton step is too big, *i.e.* exceeds the trust radius, the simplest approach is to scale it back. A bit more elegant approach is to minimize the energy with a constraint that the step is not bigger than the radius.

2.3.4. Constrained Optimization

On many occasions, it is necessary to apply constraints while optimizing the geometry. These occasions may include scanning potential energy surfaces, coordinate driving, reaction path following, and many more. A nonredundant coordinate systems and simple constraints pose no problem. The constant coordinate can be easily removed from the space of variables being optimized. More general constraints and redundant internal coordinate systems call for more advanced methods like penalty functions, projection methods, or Lagrangian multipliers.

The penalty function method imposes the constraints, $C_i(x) = 0$, by adding an extra term, $\frac{1}{2}\sum \alpha_i C_i(x)^2$, to the Taylor expansion of the PES energy (Equation (1)). Then, the energy is minimized in usual manner. The α_i coefficients must to have large magnitude to ensure that the constraints are satisfied at the minimum. Thus, the optimization may converge much slower in comparison to the corresponding unconstrained optimization.

For optimization in Cartesian space, the preferred method for including constraints is by using Lagrangian multipliers. By adding an extra term, the Lagrangian, $L(\boldsymbol{x})$, is defined,

$$L(\boldsymbol{x}) = E(\boldsymbol{x_0}) + \sum_i \lambda_i C_i(\boldsymbol{x}) \tag{20}$$

Each constrain, $C_i(\boldsymbol{x})$, has a corresponding Lagrangian multiplayer, λ_i. The derivative of the Lagrangian (equation (20)) with respect to the coordinates and the Lagrangian multipliers must be at the critical points zero.

$$\frac{\partial L(\boldsymbol{x})}{\partial \boldsymbol{x}} = E(\boldsymbol{x_0}) + \boldsymbol{g} + \boldsymbol{H}\boldsymbol{x} + \sum_i \lambda_i \frac{\partial C_i(\boldsymbol{x})}{\partial \boldsymbol{x}} \tag{21}$$

The Lagrangian multipliers are optimized along with the geometric variables. Generally, this method converges much faster than the penalty function method and the constraints are satisfied exactly.

Another way of applying linear constraints is the projection method. In this method a projector, $\boldsymbol{P}$, with a set of orthogonal constraint vectors, $\boldsymbol{c_i}$, is defined

$$\boldsymbol{P} = \boldsymbol{I} - \sum_i \boldsymbol{c_i c_i}^{\boldsymbol{T}} / |\boldsymbol{c_i}|^2 \tag{22}$$

The projector removes the constrained directions (where $\alpha > 0$).

$$\boldsymbol{P g_0} + \boldsymbol{P H_0 P} \Delta \boldsymbol{x} + \alpha(\boldsymbol{I} - \boldsymbol{P}) = 0 \tag{23}$$

In case of redundant internal coordinates, the projector has to remove the coordinate redundancies along with the constraint directions. [33]

3. Global Optimization

Global optimization methods are used to find the globally best solution of a problem in question. These methods are of great concern in modern science, technology, and economy. Because of this huge interest, a lot of effort has been dedicated to design, implementation, and testing of global searching algorithms. Many of them have been proposed, some finding application in field of quantum chemistry. The development of methods that efficiently determine the global minima of complex and rugged energy landscapes remains a challenge.

To undergo global optimization, the system has to be described by an objective function. The objective function defines quality of the trial solution. The optimal solution is reached when this function reaches the global minimum. In case of quantum chemistry, the objective function is usually the energy of the molecule; therefore, the optimal solution corresponds to the conformation with the lowest possible energy. This structure is identical with the deepest minimum on the Potential Energy Surface (PES) – the energy landscape of the molecule (see Section 2.1 on Page 7). For most applications of practical interest, physical characteristics of the system lead to an energy landscape with number of local minima, with high barriers between them. The conventional minimization techniques, based on information of the local PES, tend to be trapped within basin of attraction of the given local minimum. It is extremely difficult to find the global minimum using these local optimization methods.

The global structure optimization for big systems is a formidable theoretical challenge. Given that the most stable geometry is likely to be the global minimum, its localization is the object of interest. However, the number of possible global minima increases quickly with the number of atoms in the system, and it was showed by Hoare and McInnes[69, 70] to grows exponentially for Lennard-Jones cluster. In practice, energy landscapes for potentials describing covalently bound materials possess much more rugged nature, which further increases difficulty of the problem. It is therefore suggested, by empirical observations and theoretical arguments, that the number of local minima will generally grow exponentially with number of atoms, *i.e.* $n_{\min}(N) = \exp(aN)$.[71-74] In the same way, the number of transition states

is anticipated to increase as $n_{\mathrm{ts}}(N) = a \exp b(N)$.[71] Both a and b are system-dependent constants. Obviously this scaling precludes exhaustive search in all dimensions, and methods based on a random walk in the configuration space are extremely improbable to find the global minimum. Therefore, a number of strategies for the search has been devised.

3.1. Methods Review

Minimization methods can be classified into two groups, deterministic and stochastic. Great strength of deterministic methods is that there are extremely fast. On the other hand, their weakness lies in their liability to be caught in a local minimum quite easily. On the contrary, stochastic methods have far less probability to be trapped in a local minimum, but there is no guarantee that any stochastic method can converge to the global minimum in a finite number of steps.

Stochastic methods use a rather different strategy for global minimization, and are less likely, in comparison to deterministic derivative-driven approaches, to become trapped in local minima. However, there is no guarantee that these methods will converge to a global minimum in a finite number of steps. Stochastic methods belong to the NP-hard problems. The NP-hard[75] problems are non-deterministic Polynomial-time hard problems. In contrast to a deterministic algorithm, nondeterministic algorithm is an algorithm that can exhibit different types of behaviour on different runs. An algorithm of polynomial time is one which running time is upper bounded by a polynomial expression in the size of its input ($T(n) = O(n^k), k = \mathrm{const}$).

Stochastic methods offer a compromise between reliability and computational cost. For a fixed probability to locate the true minimum, the computational cost scales only as a power law with the number of variables.[76] Stochastic methods perform global minimization through the simulation of a dynamical process for a system (or many systems) on the multidimensional potential energy surface.

There are many approaches that allow the optimizer to escape from local minima. One of such approaches is a popular method of simulated annealing (SA),[77-79] which has enjoyed some success in this area. The SA method is an extension of Metropolis Monte Carlo techniques,[80] which simulates process of slow cooling. The system point is allowed to move around the configuration space starting from an initial, high temperature. Then, the temperature is slowly reduced until it reaches absolute zero. By a slow reduction of the temperature, the trial solutions are less probable to be trapped in local minima. At the end of the cooling process, the

system will be in the global minimum. Unfortunately, this is only true when the cooling process is infinitesimally slow. In real applications, this is unfeasible. Additionally, the presence of high-energy barriers at the saddle points between minima may prevent optimizer to migrate to the adjacent basin of attraction – the system may be controlled kinetically, rather than thermodynamically. Success of the SA method depends strongly on the choice of the cooling schedule. Each geometric cooling schedule is characterized by three parameters: starting temperature, cooling rate, and number of cooling steps. All parameters must be optimized to obtain fine-tuned results. The SA method suffers from the freezing problem for many difficult problems with rugged energy landscapes, since the escape rate from local minima diverges with decreasing temperature. To address this problem a number of variation of the original algorithm[81-86] has been proposed, many of them introduce additional parameters, making them less practical.

The stochastic tunnelling method[87] (STUN) is generalization of the SA algorithm. This approach engages the freezing problem, at the same time reducing the number of problem-dependent parameters to one. The freezing problem in stochastic minimization originates from high barriers between adjacent local minima on the PES. At high temperatures, a particle does not differentiate between the wells and can easily cross the barriers. As the temperature drops during the optimization, the particle will eventually become trapped in one of the wells, not necessarily in the one with the lowest energy minimum. The STUN method allows the particle to tunnel[88, 89] into forbidden regions of the PES, enabling the system to abandon irrelevant minima. The tunnelling is accomplished by applying a nonlinear transformation to the PES. The transformation maps the entire PES from the current lowest minimum to the maximum of the potential onto the interval of [0,1]. The cut-off of the high-energy regions is controlled by the tunnelling parameter. The locations of all minima are preserved. The modification enables the dynamical process to pass through energy barriers of arbitrary height. By continuous updating of the reference energy to new lower laying minima irrelevant features of the PES are successively eliminated.

The J-walking[90] and pivot[91] methods are also related to simulated annealing. The J-walking method is an easy to implement approach that greatly reduces the systematic errors. In standard Monte Carlo simulations, these errors result from quasi-ergodicity, or incomplete sampling of system configuration space due to large potential energy barriers. By coupling the usual random walker with the Boltzmann distribution generated by another random walker at a higher temperature, the J-walking method is able to jump over high barriers. The pivot method

utilizes a series of randomly placed points in optimization space. The points are moved based on their energy. The points with the highest energy are moved so that its energy decreases. This process is repeated iteratively until the system converges. The efficiency of the method depends on the choice of points' distribution and the way they are moved. In the original lowest energy pivot method the pivot point is chosen with a probability based on its energy. In a more efficient version, the nearest neighbour pivot method, the pivot points are chosen to be the nearest neighbour points in the phase space. The main difference between these two methods is the way in which the pivot points are selected, resulting in different ways in which the phase spaces are being searched. There is also another concept, the excitable walkers,[92] in which walkers perform parallel Monte Carlo walks on the potential energy surface, effectively repelling each other in the parameter space.

A different strategy to prevent the system from being trapped in a local minimum is to allow it to behave quantum mechanically, leading to the introduction of tunnelling. This can be accomplished by introducing Gaussian wave packets in imaginary time.[93] The method finds the global energy minimum on a multidimensional PES by finding an approximate solution to the Schrödinger equation in imaginary time. For each particle, its wave function is represented as a single Gaussian wave packet. The whole system is then expressed as a Hartree product of theses single particle wave functions. For each Gaussian wave packet equations of motion are derived. The wave packet is allowed to evolve in time and tunnels through barriers searching for the global minimum on the PES. By setting Planck's constant equal to zero, the classical minimum is found. Another approach imposing quantum mechanical behaviour on the system is the probabilistic quasi-quantal (QQ) method.[94, 95] The method replaces the classical kinetic energy by its quantum mechanical counterpart. The resulting Schrödinger-like eigenvalue equation for a Hamiltonian is then solved. The usual quantum mechanical Hamiltonian is additionally scaled by a parameter, called the "annealing factor". The method is essentially probabilistic in nature. Also a combination of simulated annealing and quantum (or diffusion) Monte Carlo[96] – referred to as ''quantum annealing''[97] – uses delocalization and tunnelling to avoid metastable regions. In contrast to the simulated annealing approach, which exploits classical character of the function that is to be optimized, quantum annealing views the system as quantum-mechanical. Further advantage of Quantum annealing is that it does not require knowledge of the wavefunction. A random walker is allowed to wander through the conformational space. By reducing the system's temperature to zero the resulting quantum ground state energy is obtained. Then, by gradually constraining the wave function (increasing the mass of the walkers), the ground state energy is reduced to its classical limit. The quantum

and simulated annealing avoid different local minima, and therefore, complement each other in general optimization applications. These methods perform well on relatively small atomic clusters, but the main problem is that they become increasingly difficult to implement as the number of dimensions rises. A recent review in an elegant pedagogical article on these and other related methods is available.[98]

Among the deterministic global optimization are the tunnelling method[99] and the packet annealing method.[100] The tunnelling method is composed of a sequence of cycles. Each cycle has two phases, minimization and tunnelling. The minimization phase lowers the current function value until a minimum is obtained. In the tunnelling phase, a new point is found, from which local minimization step will find a new stationary point with no greater value than the energy of the starting minimum. The packet annealing method mix temperature-annealing and spatial-averaging methods. Simulated Annealing contribute to the coarse-graining in objective-function, whereas spatial-averaging methods provide spatial coarse-graining.

There is a number of global optimization techniques by potential energy transformations.[101] By performing a search on deformed or smoothed energy landscape, entrapment in local minima can be avoided. These methods are of heuristic nature, *i.e.* represent experience-based techniques, which involve learning through optimization. As an example, lowering diffusion barriers,[102] stochastic tunnelling,[87] and various generalized ensemble approaches[103-105] can be named. In the ideal case the original energy landscape is transformed into a funnel landscape that leads toward the global minimum. These methods have been very successful; nevertheless, most of them require a considerable amount of fine-tuning or additional *a priori* information on the system. Some surface deformations may displace or merge minima, causing problems with connecting back to the original landscape. Moreover, there is usual no guarantee of finding the global minimum.

The tabu search method is a discrete problems approach, which tries to transcend local optimality by penalizing certain moves. In this approach, the search is guided away from areas that have already been explored, trying to force the better coverage of all-important regions of the solution space. The tabu optimizer records visited parts of the optimization space keeping them unavailable for the system. After a predefined period, moves lose their tabu status, and become once again accessible.

The Local Elevation method[106] enhances searching of conformational space in molecular dynamics simulations. The approach can be used for Molecular Dynamics (MD) and Monte Carlo (MC) simulations. Together with the conformational flooding method,[107] the

Local Elevation method introduced memory-dependence into molecular simulations. The memory-dependent term modifies the potential energy surface and prevents the simulation to revisit already sampled configurations. This simple manipulation of the potential energy leads to increased probability of visiting new parts of the configuration space. The Local Elevation method can be seen as a continuous variant of the Tabu search method.[108-110]

Metadynamics[111] is a powerful method for exploring the properties of the multidimensional free energy surfaces (FESs). The method improves sampling by means of coarse-grained non-Markovian dynamics (*i.e.* a stochastic process with memory) in the space defined by a few collective coordinates. The key aspect of the method is the presence of a history-dependent potential term. This term fills the minima in the FES, as the optimization proceed, allowing the optimization to efficiently escape from the local minimum and to further explore of the FES. The FES is represented as a function of a finite number of relevant collective coordinates.

The energy landscape paving (ELP) method[112] combines ideas from energy surface deformation and tabu search. The method avoids some of their pitfalls, having a very general applicability. The approach performs a low-temperature Monte Carlo (MC) simulation, except that the energy expression is modified to steer the search away from regions that have been already explored. The modifications may lead to a slow convergence, since the optimizer does not distinguish between important and unimportant regions of the landscape.

The Basin-hopping method[113,114] is a global optimisation approach that involves a transformation of the potential energy surface. The deformation of the hypersurface preserves the global minimum as well as the energies of all the local minima. The "basin-hopping" technique transforms the potential energy surface into a collection of interpenetrating staircases representing basins of attraction for all minima. The method exploits the features of energy landscape for efficient relaxation to the global minimum. Any point in the configuration space is associated with the local minimum obtained by a geometry optimization started from that point. This manipulation effectively removes transition state regions from the optimization problem. Global Optimization on Funneling Landscapes[115] is a variant of basin-hopping method.

The Umbrella Sampling[116] method represents another way to improve sampling of a system. An energy barrier that separates two regions of the configuration space may cause a poor sampling in Metropolis Monte Carlo simulation. Umbrella Sampling prevents this from

happening by replacing standard Boltzmann weighting for Monte Carlo sampling by a potential chosen to cancel the influence of the energy barrier.

In addition, there is a knowledge-based approach, which neither uses trial solutions nor involves modification of the PES. In geometry optimization problem, it is mainly used for prediction of the tertiary structures of proteins.[117] The knowledge-based methods identify analogies between a given case, being a subject of the investigation, and a database, having all-important information on similar cases. For protein folding, the amino acid (AA) sequence of the molecule is compared with AA chains of other proteins in the database, to recognise the secondary structures, motifs, domains or ligand interactions. Based on these similarities the three-dimensional structures can be predicted. Outside chemistry, it is also very often used in economics, machine learning, and others fields.

The method of genetic algorithm[118-123] (GA) is a stochastic minimization global technique which requires no quantum formulation, does not modify the PES, is undeterred by barriers between minima, and does not require any a priori knowledge about the system. The GA method is inspired by concept of natural evolution. Similarly to the pivot methods, a population of candidate solutions (trial solutions) is kept. Members of the population compete with each other, through selection, breeding, and mutation operations, for survival. The fittest individuals are allowed to pass their genetic characteristics (geometry) on to later generations. This happens through many generations. Ultimately, at the end of the optimization process, the fittest candidate represents the best solution to the problem in question (the global minimum). In case of geometry optimization, trial solution is a geometry of a system, and global minimum is the solution with the lowest energy among the population. Each individual geometry is assigned a fitness based on its potential energy. The lower the potential energy of a structure, the fitter it is. The fitter the structure, the more often it is allowed to mate. Usually two parent solutions are mated to produce one offspring, which contains some genetic material from both parents. Of course, from conceptual point of view there is neither any limit on the number of parent structures contributing to the one child, nor how many children are being constituted. The size of the population is usually kept constant, so its average fitness changes by allowing fitter individuals to replace older one, with the lowest fitness score.

A very important modification of the standard genetic algorithm for optimization introduces local optimization step for each trial solution. Local relaxation of the structures additionally improves the gene pool of the population. All structures that were high in energy but, at the same time, were lying in the basin of attraction of low lying minimum can now

contribute to the final solution (by being included in the population – which would not be possible without relaxation). This modified GA algorithm is called the hybrid genetic algorithm. The hybrid GA algorithm is based on Lamarck's concepts of evolution, since the members of the population pass on characteristics that thay acquired during their lifetime, *i.e.* they pass on parts of the relaxed geometry.

The main advantage of the GA approach is that it is not a greedy algorithm, *i.e.* it does not solve the local optimization problem at expense of the global solution. The genetic operations of breeding and mutation, which are stochastic in nature, very often yield children structures drastically different from their parents. These procedures enable structures to escape from local minima, and, therefore, enable algorithm to search the configuration space extensively, with the final goal of achieving the fittest possible individual. Some recent articles explore usage of GA in chemistry. [124-128]

3.2. Genetic Algorithm

The genetic algorithm (GA) is a stochastic search technique, based on the principles of natural selection, which can be applied to find the best possible solution, the so-called global minimum. Each GA run starts with generation of random structures, *i.e.* trial solutions, for the initial population. In hybrid GA algorithm, the structures are then optimized using standard local optimization methods like quasi-Newton or Trust Radius Optimization (TRO), and afterwards undergo a fitness evaluation. Standard GA algorithms do not include a local optimization part. The fitness parameter, usually based on the energy of the structure, measures the quality of the trial solution. The population evolves, through crossing over and mutation, for a certain number of generations; with the priority to crossover and mutate based on the fitness parameter, so that structures with better fitness undergo recombination more often. The GA runs stop when either the convergence criteria, *e.g.* the difference in the average energy between two generations, is smaller than a threshold, or the number of generations is exceeded.

The Genetic Algorithm is a search technique inspired by concept of Darwinian natural evolution, although in case of combining it with local optimization it corresponds to Lamarckian evolution, since parents pass on a part of characteristic that their acquired, *i.e.* parts of the relaxed structure, to their children. This coupling of the local minimization with GA was proven to be efficient for global optimization,[129] and leads to transformation of the potential energy surface (PES) into a stepped surface – with each step corresponding to a basin of

attraction of a local minimum. This simplification of the PES reduces the configuration space that the GA has to search.

Every GA algorithm starts by creating an initial population – a randomly generated set of structure, so called trial solutions (Figure 2). Each structure from this set is then optimized to the nearest local minimum; in the non-hybrid GA the local optimization step is omitted. Structures then undergo a fitness evaluation to measure the quality of the trial solution. The fitness (f_i) of each structure is evaluated as a function of the structure's dynamically scaled relative energy (ε_i),

$$f_i = \frac{exp(-\alpha\varepsilon_i)}{\sum_j exp(-\alpha\varepsilon_j)} \tag{24}$$

where α is a constant, and

$$\varepsilon_i = \frac{E_i - E_{min}}{E_{max} - E_{min}} \tag{25}$$

is the dynamically scaled relative energy of a given structure, E_i is the energy of the structure, and E_{max} and E_{min} are the highest and lowest energies of structures in the population, respectively. Only a certain number of structures from the initial population is kept. These structures are allowed to evolve, by mating and/or mutation, to create a new generation of structures. The crossover generates child structures by taking two parents from the population, cutting them through an arbitrary plane, and recombining them. The method is similar to the one used by Chuang *et al.*,[130] but had to be adopted to work with the Rigid Body approximation. Parent structures are chosen based on their fitness evaluation using a roulette wheel selection.[131, 132] The child structure is accepted if no atoms overlap, no atoms are unbound, and the new structure is not similar to any of previously obtained structures. After generating the declared number of children, all structures are optimized to local minima, and then incorporated into the existing population to create a new population. Structures in the new population undergo fitness evaluation, and only a predefined number of structures with the lowest energies is kept in the population. The whole cycle is repeated until the convergence criteria are met or the maximum number of generation is exceeded, then GA finishes and the last set of structures represents the result of the search – the so-called final population. A very important step of GA is similarity recognition. By comparing the structures no species with redundant information has been keep within the population. This increases the variety of the genes in the pool, and gives better quality of recombined structures.

In the traditional Genetic Algorithm (TGA) the coordinates of the whole system are coded into strings of binary numbers. The mating process is carried on through a one-point crossover between two individuals. Mating is accompanied with an occasional mutation of a few bits into their complement. The problem with binary coding of the coordinates becomes pronounced with the increased number of atoms. For large systems, encoding and decrypting can become the bottleneck of the computation. For that reason, several GA approaches for locating global energy minima use real-valued coding of the coordinates. First attempts used internal coordinates, like atom distances or bonds angles. Finally, space-fixed (SF) coordinates were proposed, as an equivalent solution. Advantage of using directly Cartesian coordinates is that the size of the problem increases linearly with the number of atoms. For internal coordinates, the scaling is quadratic.

Holland's schema theorem explains the power of genetic algorithms. The theorem states that in each subsequent generation subsets of solution with above-average fitness increase exponentially.[133] The Building Block Hypothesis attempt to explains the functioning of TGA based on the Holland's schema theorem. The hypothesis states that the genetic algorithm performs heuristic adaptation by identifying and recombining "building blocks", *i.e.* parts of the trial solutions. Instead of trying to construct a high-performance solution from every possible combination, the final solution is built from the best partial solutions of past samplings, through their gradual improvement. The theorem holds only if the coded bits of information are substantially related to the physical nature of the problem. This is certainly true for internal coordinates, but does not hold for SF coordinates, since one Euclidean distance can be mapped into an infinite set of Cartesian coordinates. Nevertheless, SF coordinates cope better with larger systems than binary coded. Moreover, the most successful applications of the GA method use derivative information to relax structures to a local minimum on the surface. These methods belong to the hybrid genetic algorithm (HGA).

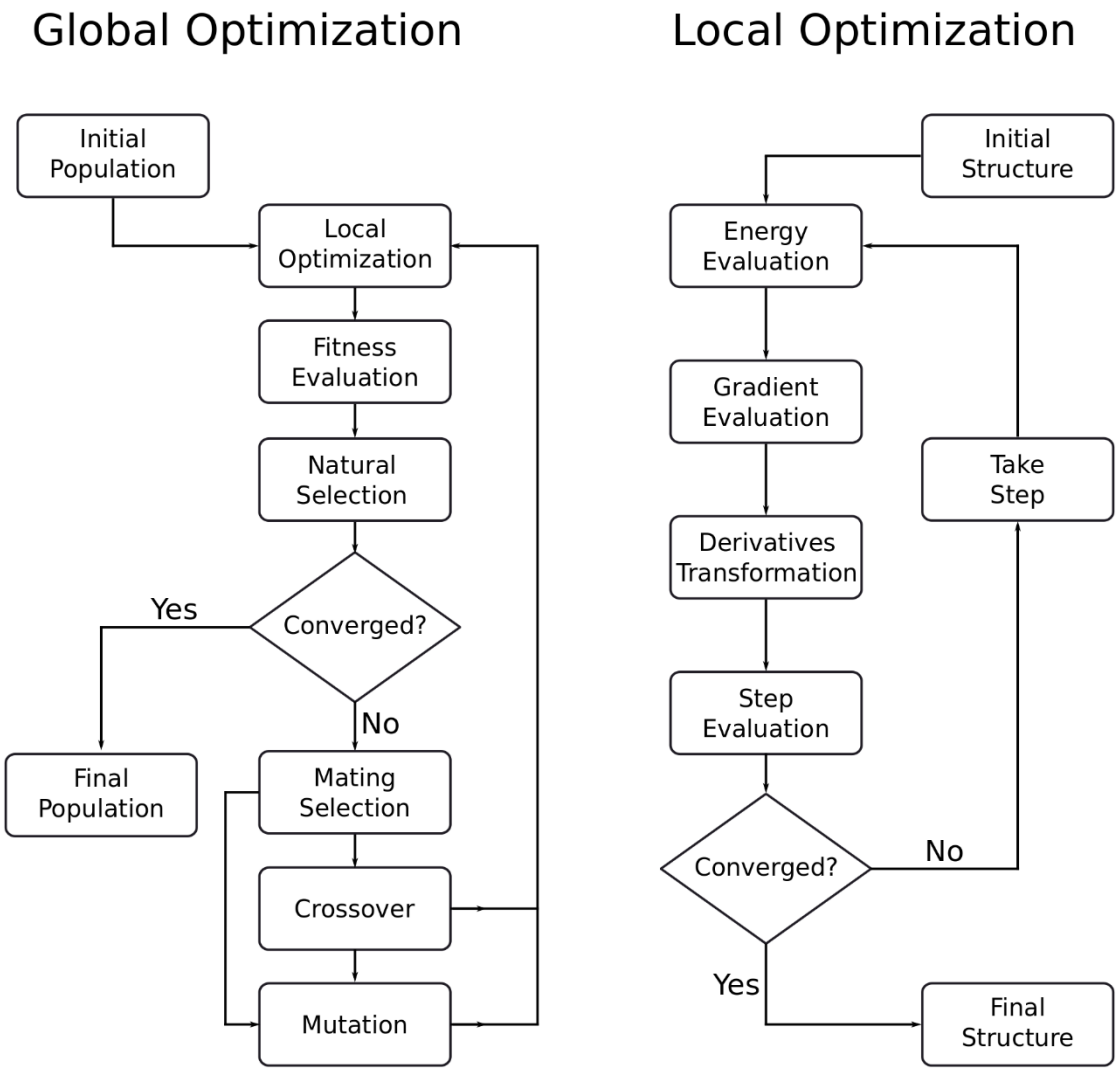

Figure 2. Schemes of global and local optimizations of the genetic algorithm in our implementation of hybrid GA. Local optimization part, right side, incorporates derivatives transformation due to Rigid Body approximation.

The genetic algorithm scheme was firstly applied to the problem of cluster optimization. In the early 1990s, Hartke, as first, proposed the use of GA optimization method for cluster geometries. In his pioneer work[134] he used TGA to localize the global minimum for the Si_4 cluster. The cluster geometry was binary encoded in a set of internal coordinates, and the genetic operators were acting in a bitwise fashion on the binary strings In later work,[135] he extended this treatment to larger cluster using growth strategy, *i.e.* seeding Si_n calculations with the Si_{n-1} structure. Niesse and Mayne[136] were able to minimize silicon cluster up to ten atoms without restrictions using SF real coding and gradient-driven minimization for each structure. The

results improved previously reported minima. Xiao and Williams[137] used TGA to investigate molecular clusters of benzene, naphthalene, and anthracene. For each molecule, geometry was binary encoded as six parameters: three translations of the centre of mass and three rotations around the centre of mass. They obtained convergence for dimer cluster of each investigated molecule, and for the trimmers and tetramers of benzene. An important step in the evolution of GA algorithm was made by Zeiri.[138] He used the real-valued Cartesian coordinates without further encoding. This way, the cluster could be represented in terms of continuous variables; and removed the costly requirement for encoding and decoding geometry into binary genes. The scheme was further improved by Deaven and Ho[129] by introducing a gradient driven local minimisation of the cluster energy for each newly generated cluster. The introduction of local minimisation transforms the cluster PES into a stepped surface, with each step corresponding to a basin of attraction of a local minimum. This transformation reduces the space that the GA has to search, and therefore simplifies the PES and greatly facilitates the global optimization. Using local minimisation step, change character of the GA algorithm from Darwinian to Lamarckian evolution, since individual members pass on, to their offspring, a proportion of the characteristics that they have acquired during their "life time". The acquired characteristics relate to the geometries after local minimisation. The hybrid GA algorithm has much improved efficiency in comparison to the traditional GA method. In the same article, Deaven and Ho made another significant improvement to the cluster optimisation GA method, they introduced the 3-dimensional cut and splice crossover operator, giving it more physical meaning. The operation was adapted into most subsequent works. The results on carbon clusters bound by a tight binding potential and Lennard-Jones clusters with 2–100 atoms from subsequent article,[139] show that the authors were able to localize many low energy minima that had not been found previously. Since the first application of the GA algorithm to geometry optimization there has been a fast and rapid development of this method. Further information on the current state of the advancement in the field can be found in recent reviews.[140, 141]

In comparison to clusters, usage of genetic algorithms to obtain structural models for periodic systems like surfaces and crystals, is new concepts. Bush *et. al.,* [142] used genetic algorithms together with energy minimisation to predict the crystal structures of complex inorganic solids. In their implementation of the GA algorithm, the binary string stored the coordinates of all the atoms in the cell. Two-point crossover scheme helped to retain genetic diversity, and conjugate gradient minimisations of several randomly chosen individuals served to improve the quality of 'gene pool' available for breeding. Top ten individuals of the GA step were output into the second stage of the calculation in which their lattice energy was minimised. The efficacy of the

method was demonstrated by yielding accurate solutions for complex crystal structures. In the following work, Woodley *et. al.*[143] used genetic algorithm to generate plausible crystal structures with the unit cell dimensions and constituent elements as an input. The quality of the structures were assessed using a cost function based on the refined bond valence model. The lattice energy was minimised using a local. The effectiveness of the approach was showed by generating number of crystals, starting from configurations without knowledge of the atomic arrangement in the unit cell. Bazterra *et. al.*[144] presented a new computational scheme to model crystal structures of organic compounds based on a modified genetic algorithm. The model had no assumptions on the crystallographic group nor to the number of molecules in the unit cell. Real-valued Cartesian coordinates and Euler angles between molecules were used as variables that were optimized in the genetic search. The molecules themselves were treated as rigid bodies. As an example, global optimization of crystal structures of the benzene, naphthalene, and anthracene molecules was performed. The results showed good agreement with the experimental data. Chuang *et. al.*[130], applied the hybrid genetic algorithm to find the reconstruction of the semiconductor surface. The algorithm used Cartesian representation of atomic positions and real-space representations for crossover and mutation operators, and the trail structures were optimized locally. Additionally, authors maintained the structural diversity of the population using a similarity recognition scheme. In this scheme, the surface energies of the structures were compared, and if the energy deference fell within a given threshold, the structures were analyzed for topological similarity. Oganov *et. al.,*[145] developed an efficient and reliable methodology for crystal structure prediction, by merging *ab initio* calculations with a specifically devised evolutionary algorithm. The method allows to predict the most stable crystal structure and a number of low-energy metastable structures without requiring any experimental input. Test cases included ionic, covalent, metallic, and molecular structures with up to 40 atoms in the unit cell. Authors reported the method to be highly successful. Abraham *et. al.*[146] used genetic algorithm to determine the global minimum energy configurations of crystal structures. The algorithm requires no prior knowledge about the ionic configuration within the unit cell, nor information about size, shape, or symmetry of the unit cell. The algorithm uses a real-space representation of the coordinates, and performs periodic cut as the crossover operation. Briggs et. al.,[147] presented genetic algorithm for determining the atomic configuration of crystallographic steps. To find the atomic structure of the steps oriented along given direction in the plane of the surface, the algorithm uses information on a stable surface orientation with a known reconstruction. The efficacy of the method was exemplify by finding structures for several types of monatomic steps on Si(114) surface. Zhu *et. al.,*[148] presented an

evolutionary algorithm for the prediction of surface reconstructions. The authors used a representation in which the system has three parts: vacuum, surface, and substrate. Vacuum and substrate regions are predefined. Only the surface region is subject to the GA optimization. The number of surface atoms varies up to a predefined maximum number. To allow surface reconstruction, the cell size is also variable. The algorithm can be used to automatically explore stable and low-energy metastable configurations of surfaces.

Genetic algorithms for cluster and surfaces are conceptually the same. Both keep the population of trial solutions. At each cycle, new trial solutions are obtain by genetic operations like crossover, mutations. Creation of new structures proceeds until given convergence criterion is meet. The only difference is within the implementation of genetic operations. Of course, for cluster optimization additional criterion concerning 3D rotations of clusters has to be considered. For periodic calculations, coordinate system is fixed, and usually related to dimensions of the unit cell.

4. Rigid Body Genetic Algorithm

4.1. Introduction

The use of Rigid Body coordinates allows reducing the computational effort of finding local minima. Number of factors accomplishes the reduction: decreasing the number of degrees of freedom which have to be optimized; reduction of couplings between atoms coordinates, and separating stiff and flexible coordinates. All these elements provide a perfect mean to treat systems with distinguishable subsystems – like surface and adsorbate molecules – which can be then treated separately, as unchangeable entities during the course of optimization.

In the rigid body approximation the whole system is divided into disjoint sets of atoms or molecules which form rigid bodies. Determining the position of N_{at} atoms requires a set of $3N_{\mathrm{at}} - 6$ coordinates in a nonlinear molecule, while the same system with N_{rb} predefined rigid bodies requires only $6N_{rb} - 6$ independent coordinates; and, therefore, the advantage of introducing rigid-body coordinates is pronounced for systems with a large number of atoms and a relatively small number of rigid bodies, *i.e.* when $N_{\mathrm{rb}} \ll N_{\mathrm{at}}$. In this case the reduction of the search space leads to a rapid and fast localization of a stationary point, a local minimum or saddle point, which can be cumbersome or even unfeasible in the unconstrained case. This, of course, can greatly decrease the computational cost of investigating a molecular system.

The RB method not only reduces the optimization space by freezing some degrees of freedom and, hence, reducing number of possible minima, but also restricts the optimizer from searching unwanted regions of the PES. These regions could correspond to very unlikely molecular conformations, like highly oxidised states, or cases when part of the molecule is "submerged" under the surface. Within the RB approach we strictly require that no alternation to the composition of the rigid body happens. Even more, rigid bodies cannot undergo any conformational changes. What is left as a subject of optimization is relative placement of the rigid bodies with respect to each other. We treat all atoms explicitly, yet we forbid them to move freely. This approach seems to be perfect to study physical phenomena like molecules

adsorption, aggregation, diffusion, melting, freezing, vaporization, condensation, sublimation, etc. Application of the RB method can range from homogenous/heterogeneous catalysis, to enzyme catalysis, photocatalysis, phase-transfer catalysis, and more.

The rigid body approximation also allows for separation of intra- and intermolecular degrees of freedom.[42] In this method, the intramolecular interactions are neglected, while only intermolecular interactions are taken into account. This segregation of degrees of freedom leads to a division between weak and stiff modes. Again, the necessary assumption is that investigated system does not undergo a substantial change in molecules' connectivity, like breaking or formation of the bond. This way, weakly bounded systems (*i.e.* systems with a distinct division between strong intermolecular forces, and weak intramolecular interactions) can be treated in a simplified fashion. The approach allows modeling of adsorption processes, nanoparticles interaction, liquids, molecular clusters and complexes, crystallization – *i.e.* every phenomenon that do not involve chemical reaction.

Throughout the optimization process each rigid body is kept frozen, *i.e.* all the internal degrees of freedom are fixed and only relative position of the rigid bodies, translation and rotation, are being subject to the optimization problem. The construction of rigid bodies precludes applications of this method to investigate chemical reactions in which an atom, or a group of atoms, is transferred from one molecule to another. Rigid bodies are defined using proper internal coordinates (*i.e.* rigid-body coordinates). Each rigid body can be fully described by six independent coordinates, the most general choice, following Chasles' theorem,[149] are three translations of the center of mass, and three rotations with respect to the center of mass of the rigid body.

There are three major ways to introduce rigid body scheme. The first one is to evaluate the derivatives in a rigid-body framework, and use them with standard optimization techniques. This approach is easy to implement for empirical and semi-empirical methods. It is, therefore, usually used with force fields and molecular dynamics calculations. The second option is to perform constrained optimization in Cartesian or internal coordinates. In the former case, this requires definition of a proper Lagrangian, in the latter, first a projection form Cartesian to internal coordinates has to be performed, and then all frozen coordinates have to be projected out. The third option is to introduce proper internal coordinates, ones that will intrinsically address issues of freezing parts of the system. This approach lacks overhead calculations due to constraining the system. It also seems to be much better suited for *ab initio* calculations,

since it can easily utilize Cartesian derivatives used by most of the quantum mechanics codes, and avoids need of redefining derivatives in the rigid-body framework.

Most of the standard quantum chemistry codes evaluate energy and its derivatives in the Cartesian coordinate system. To allow the rigid body optimization, a proper transformation of first- and second-order derivatives must be defined. To avoid singularities during derivatives transformation form Cartesian coordinates to rigid-body coordinates, Rodrigues' rotational formula – an efficient algorithm for rotating each rigid body in space – was used. The advantage of this approach is that no backward transformation is required; the calculated step, in rigid-body coordinates, corresponds directly to transition and rotation of a rigid body in the Cartesian coordinates. Of course, by performing a projection onto rigid-body subspace, we change the dimensionality of the problem in question, but this should not influence the topography of the PES. All stationary points should preserve the topology of the unconstrained PES, although some stationary points will be unavailable.

Each rigid body has assigned its own Cartesian coordinate system, attached and rotating with the molecule. The position of the Cartesian systems is defined with respect to the laboratory-fixed frame. The choice of translational degrees of freedom is straightforward, but care is required in case of the rotational coordinates. The Euler angles suffer from the problem of information loss during transformation of derivatives from Cartesian coordinates, due to singularities at the origin of the coordinate system. Similar problems occur when polar coordinates are used. Unit quaternions provide a singularity-free description of pure rotation, but the quadratic unit-norm constraints make it less attractive for structure optimization proposes. An alternative offers an angle-axis representation based on the Euler's rotation theorem, which states that each arbitrary rotation of a rigid body around its center can be presented as rotation about one axis.

The main advantage of using rigid-body coordinates is that there is no back-transform step, like in standard internal coordinates. Therefore, there is no need to calculate inverse of the $\boldsymbol{B}$ matrix. This can be cumbersome for large number of particle, since matrix inversion scales N^3 with matrix size. Displacements obtained in rigid-body coordinates correspond directly to translations and rotations of the molecules (rigid bodies) in Cartesian coordinates. This is a one-step procedure, involving substitution of the translations and rotation into the Rodrigues' formula. In contrast, step transformation in standard internal coordinates is an iterative scheme. Additionally, imposed constraints are intrinsic to the design of internal coordinates. Therefore,

no additional procedures to impose constraints are needed; like projecting frozen distances and angles out of optimization space.

The idea of introducing constrained optimization, so as to obtain rigid bodies, is not new to the optimization problems; and it is widely used to answer a variety of questions in quantum chemistry.[42, 150, 151] Nevertheless no use of it was reported so far for *ab initio* slab calculations with periodic boundary condition. In our implementation we introduce the Rigid Body approximation at the local optimization step, by implementing proper transformation coordinates and gradients from the Cartesian to the rigid-body coordinates, in the existing code. At the same time GA steps – initialization, crossover, and mutation – had to be modified to work within the same framework. In generation of the initial population predefined number of rigid bodies is taken and randomly distributed, using translation and rotation, on top of the surface. Obtained structures are checked if any rigid body is not unreasonably close or far from the surface or other rigid bodies; they are also compared with previously obtained structures, so as to not recalculate them.

4.2. Methodology

To describe a system in the rigid body approximation the relationship between the Cartesian coordinates and the rigid-body coordinates, which are specific type of internal coordinates, must be found. This relationship provides means to express the gradient (first derivatives of energy) and the Hessian (second derivatives of energy) in the rigid-body coordinates. Derivatives are then used to predict a step on the PES in the rigid-body frame.

In this work, we use the angle-axis representation for rotational coordinates of a rigid body (following Chakrabarti *et.al.*[150]). The rotation is described by the Rodrigues' rotation formula. The formula allows for direct transformation of the step into Cartesian coordinates. The step in the Rigid Body (RB) coordinates corresponds to infinitesimal translation and rotation in Cartesian coordinates, and it is obtained in a one-step process by substituting the RB step into an equation describing rotations and translations of all rigid bodies in Cartesian coordinates. Therefore, we avoid costly back-transformation typical for other internal coordinates. In addition, since the design of the Rigid Body coordinates by definition imposes constraints of a desired character on the system, we avoid extra work in relation to introduction of constraints. For the Cartesian coordinates this is usually done using the Lagrange multiplier method, while for internal coordinates it is done by proper projection constructed to impose

expected constraints. Of course, in both cases the introduction of constraints leads to an increase in computational cost.

Let us consider a system of N molecules represented as rigid bodies. Following Schröder[151] we define two coordinate systems; a complex-fixed Cartesian coordinate system, with origin in the center of mass of the complex; and space-fixed system, which axes coincide which axes of the complex-fixed Cartesian coordinate system. In addition each rigid body has defined Cartesian coordinate system with the origin in the center fo mass of rigid body.

The first rigid body, *i.e.* surface, is treated as frozen. Its position is not subject to the optimization process. Position of each other rigid body is fully describe by 6 coordinates: 3 translations along the axes of space-fixed system, $\boldsymbol{r_i} = [r_{ix},\ r_{iy},\ r_{iz}]$, and 3 rotations around the axes of rigid body's coordinate system, $\boldsymbol{t_i} = [t_{ix},\ t_{iy},\ t_{iz}]$. In total, $N-1$ rigid bodies are defined using $6(N-1)$ independent coordinates. Displacement in the rigid-body coordinates contains values of rotations and translations for each rigid body,

$$\boldsymbol{s_q} = (\boldsymbol{r_1}, \boldsymbol{t_2}, \ldots, \boldsymbol{r_{N-1}}, \boldsymbol{t_{N-1}})^T. \tag{26}$$

Displacements in the Cartesian coordinates ($\Delta\boldsymbol{x}$) can be expressed as a sum of rotations and translations of all N rigid bodies,

$$\Delta\boldsymbol{x} = \boldsymbol{x}' - \boldsymbol{x} = \sum_{j}^{N-1} (\boldsymbol{r}_j + \boldsymbol{t}_j) = \sum_{j}^{N-1} (\boldsymbol{R}_j \boldsymbol{x}_j + \boldsymbol{t}_j), \tag{27}$$

where $\boldsymbol{R}_j$ is the rotation matrix acting on the Cartesian coordinates of j^{th} rigid body – $\boldsymbol{x}_j$, the $\boldsymbol{R}_j$ matrix is given by Rodrigues' rotation formula

$$\boldsymbol{R}_j = \mathbf{I} + [\boldsymbol{p}]_x \sin\theta + [\boldsymbol{p}]_x^{\,2}(1 - \cos\theta) \tag{28}$$

Where $[\boldsymbol{p}]_x$ is a cross product matrix

$$[\boldsymbol{p}]_x = \begin{bmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{bmatrix} \tag{29}$$

$$\theta = \sqrt{{p_1}^2 + {p_2}^2 + {p_3}^2} \tag{30}$$

of a vector $\boldsymbol{p} = [p_1, p_2, p_3]$, indicating direction of an axis of rotation of j^{th} rigid body and θ is an angle of rotation about $\boldsymbol{p}$.

More general, to find the transformation between derivatives, we consider a set of n Cartesian coordinates $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^T$, and a set of r rigid body coordinates $\boldsymbol{q} = (q_1, q_2, \dots, q_m)^T$. Small changes in rigid-body coordinates transform to change in Cartesian coordinate as

$$dx_i = \sum_j \frac{\partial x_i}{\partial q_j} dq_j \,. \tag{31}$$

In matrix formulation, change in the Cartesian coordinates ($\delta\boldsymbol{x}$) is given by change in rigid-body coordinates ($\delta\boldsymbol{q}$)

$$\delta\boldsymbol{x} = \boldsymbol{B}\delta\boldsymbol{q}. \tag{32}$$

The $\boldsymbol{B}$ matrix is similar to Wilson's matrix. Its elements are given by

$$(\boldsymbol{B})_{ij} = \frac{\partial x_i}{\partial q_j}. \tag{33}$$

Using the definition of gradients in the rigid-body coordinates and the Cartesian coordinates, the transformation of the coordinates (7), and the $\boldsymbol{B}$ matrix (9), the relation between them can be given by

$$\boldsymbol{g_q} = \boldsymbol{B^T g_x} \tag{34}$$

Similarly, the relation between Hessians in the rigid-body coordinates and the Cartesian coordinates can be given by

$$\boldsymbol{H_q} = \boldsymbol{B^T H_x\, B + K} \tag{35}$$

Where the $\boldsymbol{K}$ matrix is the correction due to the non-zero Cartesian gradient

$$\boldsymbol{K} = \sum_{k=1}^{n} (\boldsymbol{B'})_{kij}\, [\mathbf{g_x}]_k \tag{36}$$

and the $\boldsymbol{B'}$ matrix is given by

$$(\boldsymbol{B}')_{kij} = \frac{\partial^2 x_k}{\partial q_i \partial q_j} \tag{37}$$

Equations (34) and (35) allow for direct transformation of gradients and hessians, using only $\boldsymbol{B}$ and $\boldsymbol{B}'$ matrix; both can be derived from equation of displacement in Cartesian coordinates (27).

There is no additional effort needed to perform transformation from the rigid body coordinates to the Cartesian coordinates. The step prediction on the PES in the rigid-body coordinates

$$\Delta \boldsymbol{q} = \boldsymbol{H}_{\boldsymbol{q}}^{-\mathbf{1}} \boldsymbol{g}_{\boldsymbol{q}} \tag{38}$$

corresponds to translations and rotations of the rigid bodies. Rotational and translational coordinates can be substituted into equation (27), to yield a step directly in the Cartesian coordinates. Neither iterative method nor back transformation of coordinates is required (it scales as $O(n^3)$, and can be the bottleneck for larger systems).

4.3. Implementation

To enable the usage of the rigid body approximation as a local optimization scheme with a genetic global optimization scheme, changes to the existing codes had to be done.

In our implementation the global optimization part is performed by the DODO program,[131] which uses an interface to the local optimization program (QMPOT[152]). The VASP program[3] evaluates single point energies and gradients, which are then used by the QMPOT program to optimize the structure. The optimization is done by moving the structure on the potential energy surface (PES) towards the local minimum. The global optimization program keeps a population of structures during each global search, and also controls how many of these structures are being optimized at the same time.

Child structures are obtained by cutting parent structures by an arbitrary plane and recombining them. Whether rigid body is associated with one or the other part of the structure depends on which side of the cutting plain lays its centre of mass. In the recombined structure rigid bodies have the same orientation as in the parents' structures. Mutation is performed by choosing random number of rigid bodies, rotating them around their centres of mass, and translating them along $\boldsymbol{c}$-vector direction away from the surface.

This implementation is suited only for surface models with one type of rigid bodies. Further improvement of the algorithm to include optimizations for different types of rigid bodies and for cluster systems is possible and straightforward.

4.4. Rigid Body Genetic Algorithm Performance

The newly developed Rigid Body Genetic Algorithm (RBGA) method, was tested by running 3 independent GA optimization runs for a system consisting of four methane molecules at the MgO(001) surface. The global minimum for this system is well known to be the ROT structure. It is a monolayer structure, with methane molecules adsorbed at every other magnesium site; each molecule is adsorbed in on-top position at the site maintaining the dipod configuration; neighboring molecules are rotated with respect to each other, by 90 degrees around surface normal going through the carbon atom.

The system was a subject of previous investigations, both experimental and theoretical. The global minimum structure is resolved and recent results are very consistent about its structure. Therefore, due to the possibility of relating the results to experimental and other theoretical studies, it is a perfect model to benchmark the performance of the RBGA method. The main concern was whether the RBGA algorithm will be able to localize this structure. However, so far all studies involved less strict methods based on chemical intuition. This is the first report of using global optimization methods to investigate methane monolayer structure at the MgO(001) surface.

The surface is a 2A × 2A slab model - A is two times the Mg-O distances. This gives eight magnesium and eight oxygen binding sites at the top of the model (more detailed description of the surface model can be find in Section 5.2 on page 56). Four methane molecules at that surface correspond to monolayer coverage (one methane molecule per 2 magnesium sites). Each GA run has begun with 100 structures in the initial population.

Figure 3 shows progress of all three independent GA investigations. The y-axis represents the energy difference with respect to the energy of the final structure – supposed global minimum – obtained during the optimization. The x-axis shows number of the structures calculated over the optimization. The solid line corresponds to structures with the lowest energies in each GA populations. The dotted lines correspond to the average energies of 20 and 50 structures. All of the values are used to monitor convergence of the genetic algorithm. The calculation can be considered as converged when the average energies change by less than a

preselected threshold. The size of the threshold depends on the system. For the methane/MgO system, small differences in energies can translate to huge structural rearrangements. That calls for a very strict threshold. Small changes in the average energies also imply that the global minimum is stable, *i.e.* it has not changed over the last few cycles.

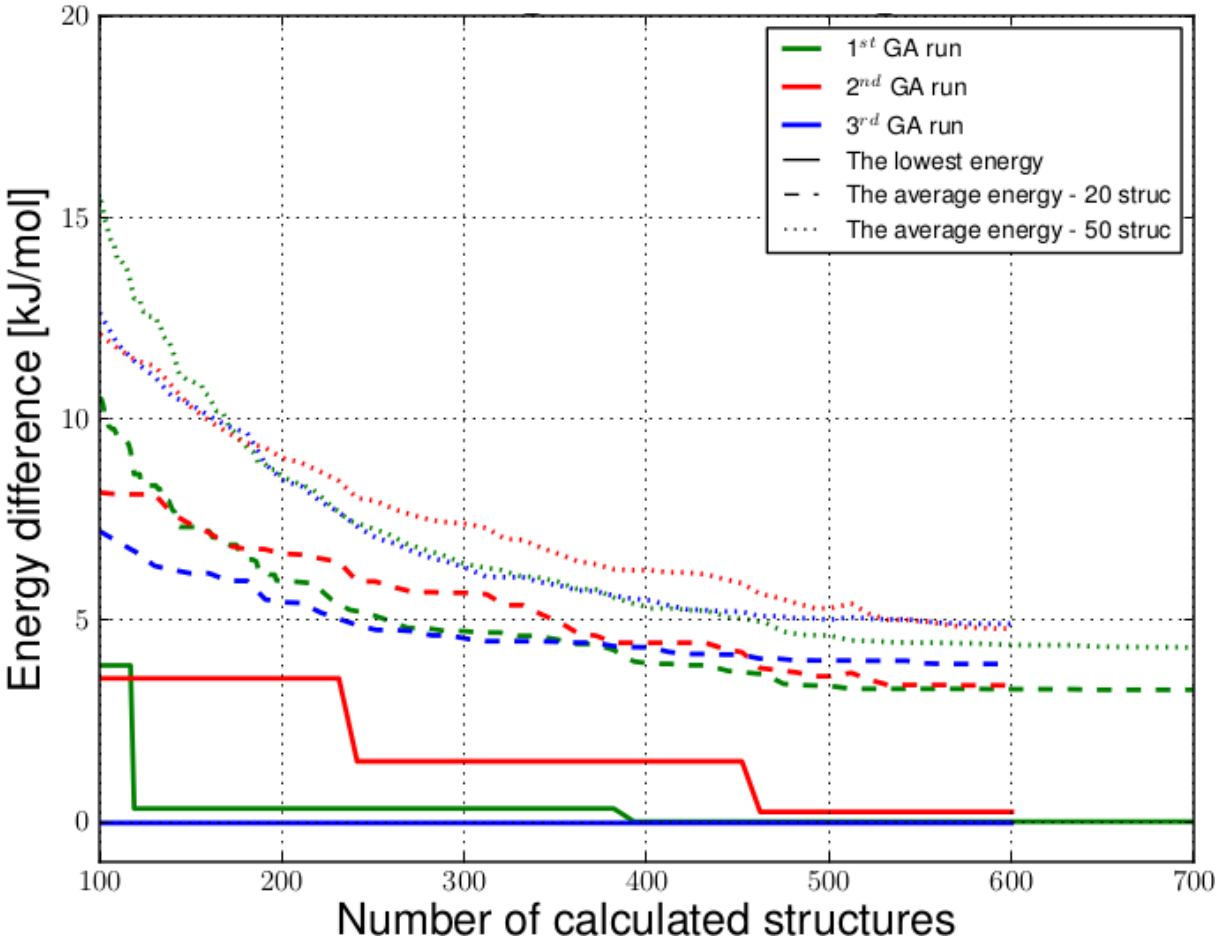

Figure 3. Performance of Genetic Algorithms for each run: first run with 700 calculated structures, second and third run with 600 calculated structures. The lowest energy structure for second GA run is 0.24 kJ/mol higher in energy with respect to the two other. Energy difference is relative with respect to the final global minimum. The small difference in the final energy of the global minimum for second GA run will disappear at the full relaxation step of the optimization process.

The first and third RBGA runs localized the same global minimum. The first GA converged after optimizing 400 structures (green solid line). The second GA run did not converge to the same structure as the other two optimization runs. However, the difference between the final structures is only 0.24 kJ/mol in total energies, which translate to 0.06 kJ/mol

per molecule. This difference is negligible, since at the full optimization step all structures undergo further optimization, and in the end all structures converged to the same final global minimum structure. Also a visual check confirmed that the second GA final structure was indeed a slightly distorted ROT configuration. The reason why this has happened may be due to the fact that the local optimization ran out of cycles and, therefore, the structure did not fully converge. In the third run, the global minimum was already found after optimizing structures in the initial population (solid blue line).

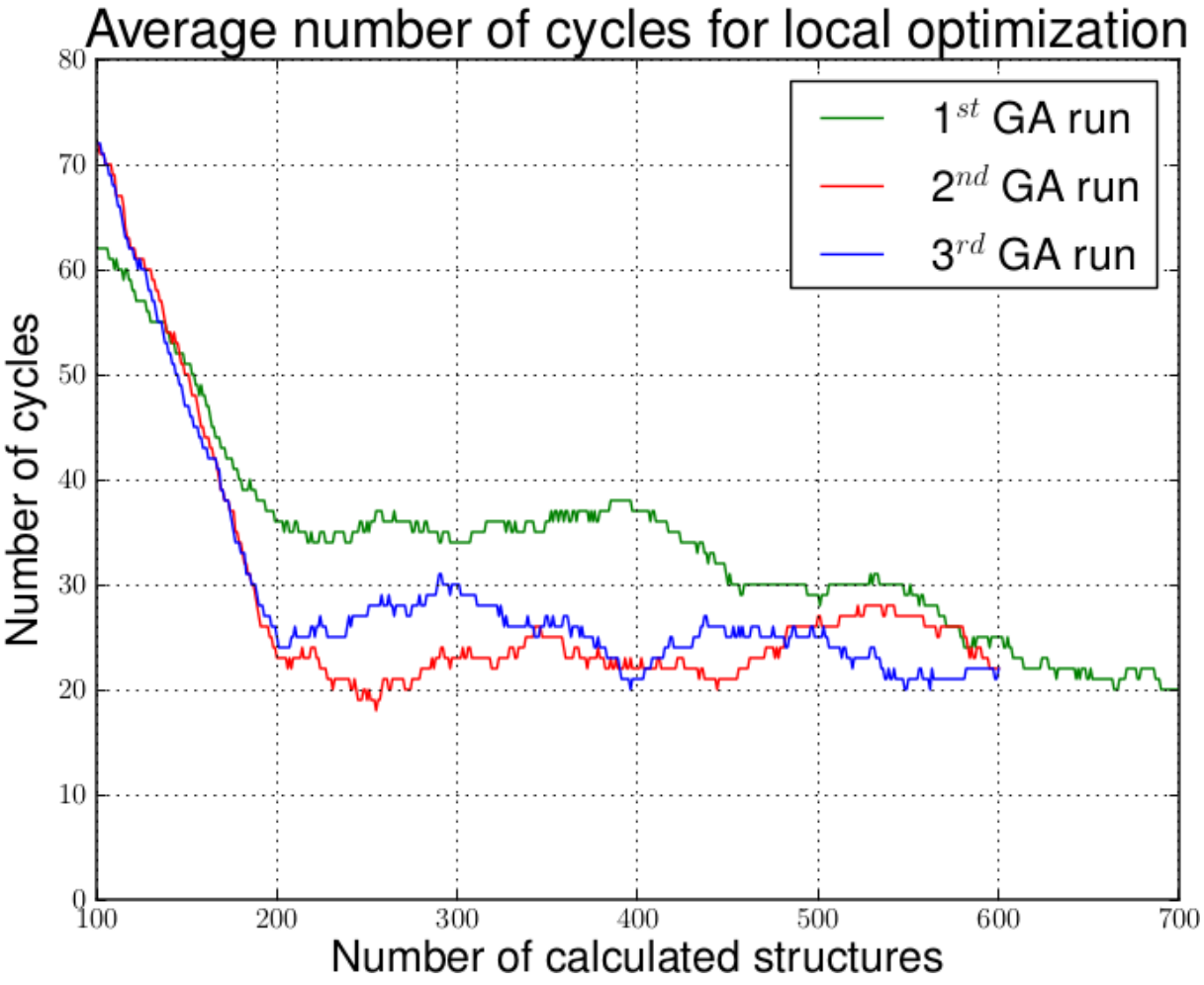

Figure 4. Average number of local optimization cycles needed to reach a local minimum, calculated as a simple moving average over last 100 optimizations.

Figure 4 shows average number of local optimization cycles for last 100 calculations. This value is relatively high for the initial populations, above 60 cycles, as the randomly generated structures are expected to be strongly disordered. Number of cycles drops significantly within second hundred structures, since the child structures inherit parts of the

ordered character from parents. The goal of the performance analysis was to confirm that all three RBGA optimizations were able to localize the global minimum for the monolayer coverage of the methane at the MgO(001). That goal was reached. The obtained geometries, along with interaction energies, thermodynamic frequencies, and frequency analysis are discussed in the Section 5.7.2 on page 72.

5. Case Study: Methane, Ethane, Propane on MgO(001)

After implementation and validation, the newly developed RBGA method was used to investigate interaction of hydrocarbons with a magnesium oxide surface.

5.1. Motivation

The interaction of individual molecules with crystal surfaces and with the internal surfaces of microporous materials such as zeolites or metal–organic frameworks is an active research area of catalysis. Interaction of hydrocarbon molecules with oxide surfaces is particularly interesting in the context of oxidative coupling of methane (OCM).[153] As a support or an active catalyst MgO crystal plays an important role in the OCM process. Although, the OCM reaction requires irregularities[154] and defects at the oxide surface, investigation of hydrocarbons interaction with the clean MgO surface allows for instructive insight on the adsorption, since adsorption is the first step of catalytic hydrocarbon conversion and synthesis reactions. MgO can be considered a prototype of square lattice metal oxide surface, which is of interest in the framework of heterogeneous catalysis. Similarly, methane can be considered the most basic example of a hydrocarbon. Therefore, the combination of these two model systems has been the subject of a numerous previous studies. The similarity in lattice parameters of methane and magnesium oxide (4.18 Å in the former[155] and 4.21 Å in the latter[156]) drives methane to adopt a square lattice commensurate with the surface structure.

The structure of thin methane films was a subject of a number of diffraction experiments. The first investigation of the adsorption of methane on the MgO(001) surface was performed by Coulomb and his co-workers.[157] In the neutron diffraction experiment they characterize the structure of deuterated methane film adsorbed on a highly homogenous MgO(100) surface. The main result of this study suggested that both the bilayer and the monolayer coverage of CD_4 formed a square $c(2 \times 2)$ structure that commensurates with the surface lattice. The best matching with experimental data was obtained for a lattice constructed

with methane molecules in dipod configurations. Nevertheless, the diffraction data was not sufficiently precise to differentiate exactly between the possible orientations of the methane molecules. Helium atom scattering experiments and force field calculations for the MgO surface with adsorbed a CH_4 layer exclude the face-down configuration (tripod) of methane as unlikely.[158] Later experiments showed that CD_4 grows as a commensurate square lattice to 3 layers at 50 K[159] and to 4 layers at77 K. [160] Subsequent helium atom scattering[161] shows that the real-space lattice determined from magnetic-sector mass spectrometer (MSMS) for the CH_4/MgO(001) monolayer is a commensurate $c(2 \times 2)$ square lattice. The lattice was estimated to 4.21 Å (at 33–36 K), and adsorption energy to 136 meV (13.1 kJ/mol). Larese *et al*,[162] used inelastic neutron scattering (INS) experiments to study rotational diffusion of an adsorbed methane monolayer. Matching the observed transition energies with molecular field potential function they arrived at the conclusion that methane molecules prefer the dipod configuration on the MgO(100) surface. Further studies by Larese and others[159, 163, 164] confirmed those results. Quasielastic neutron scattering, employed by Gay *et al.*[165] to study rotational diffusion of an adsorbed methane monolayer, showed that the molecules are rotationally ordered below 20 K, rotationally disordered around the twofold axis perpendicular to the surface in the dipod configuration between 20 K and 40 K, and are in a free-rotor state above 40 K. The experiment also excludes the tripod configuration. The most recent calculations of monolayer and bilayer lattices of methane as a spherical molecule on the MgO(001)structure, reported by Bruch *et al.,*[166] show that stability of c(2 × 2) commensurate square monolayer (and bilayer lattices) is reproduced with a surface energy corrugation that gives a large gap in phonon density of states of the commensurate methane monolayer, that is present in the incoherent inelastic neutron scattering. Tait *et al.*,[167] obtained adsorption energies for seven small n-alkane molecules on the MgO(100) structure from high-quality temperature-programmed desorption data. For methane, the adsorption energies are 11.1 kJ/mol and 12.6 kJ/mol for zero-limit coverage and monolayer coverage, respectively.

Also a variety of theoretical methods have been utilized to investigate the orientation of a single methane molecule, as well as the whole methane monolayer. Deprick and Julg[168] performed Hartree-Fock electronic structure calculation using three model systems: a single methane molecule on a surface constructed of point charges, a $CH_4Mg_4O_4$ cluster, and four methane molecules above a point charge surface. The results showed that out of three possible orientations an isolated molecule should have the dipod (010) configuration, with the monopod configuration being less stable, and the tripod configuration even more unstable. The monolayer could have the (010) orientation as well as the (011) configuration (ROT configuration with all

methane molecules rotated 45 degrees in the same direction around 2 fold axis perpendicular to the surface). In contrast, the results from the semi-empirical potential model used by Girard and Girardet to calculate the interaction energy of a methane molecule adsorbed on a MgO substrate with square symmetry, show that the tripod configuration is slightly favoured over the dipod configuration (by 13 meV – *i.e.* 1.25 kJ/mol) for an isolated molecule.[169] Alavi[170] has studied a monolayer molecular crystal of methane adsorbed on the (100) surface of MgO using molecular dynamics. Methane was modelled as a five interactions site rigid tetrahedral. The interaction between sites on different molecules was described by Lennard-Jones and electrostatic terms. The predicted structures configuration of molecules was over the magnesium sites with a tripod orientation. Semiempirical potentials used for classical determination of equilibrium structure of a single methane molecule[193, 195] and monolayer [196] showed that the tripod orientation for the molecules was significantly more stable than the dipod one. Ferrari *et al.,*[171] applied calculations at the density functional level of theory employing the LCGTO-DF (Linear Combination of Gaussian-Type Orbitals Density Functional) package combined with the gradient corrected Becke-Perdew (BP) functional, to predict methane interaction with the MgO clusters model. According to the results, methane bound weakly to Mg^{2+} and O^{2-} corner sites, but did not adsorb on the 100 planar surface. The adsorption energies for the corner sites lie between 1-3 kcal/mol (4-13 kJ/mol). For interactions with the MgO (100) planar surface four on-top and bridge sites were investigated. The potential energy curves for the Mg^{2+} site at the MgO (100) planar surface were totally repulsive for all configurations. The potential energy curves for adsorption at the O^{2-} site of the MgO(100) surface showed that the tripod structure interaction is purely repulsive, and while in cases of monopod and dipod-bridge positions a local minimum is present, both configurations become unbound after the Basis Set Superposition Error (BSSE) corrections. Nonetheless, the investigation considered dipod configuration only in bridge position. Todnem *et al.*[172] employed embedded cluster models to investigate adsorption of methane at the MgO(001) surface employing Modified Coupled-Pair Functional (MCPF) method with extrapolation of the correlation energy to the full CI complete basis set limit (PCI-X). A single molecule adsorption was investigated at on-top magnesium and oxygen sites using MgO_5^{8-} and Mg_5O^{8+} surface model clusters, respectively. The results predicted that the dipod configuration is preferred over the tripod configuration; but only by 0.4 kJ/mol. Combining adsorption energies for a single molecule with interadsorbate interaction energies obtained as four times half the pair-interaction energy of two methane molecules (interaction energy of one methane molecule with four nearest neighbours), the authors arrived at adsorption energies at monolayer coverage. The reported value of the adsorption energy, 8.5 kJ/mol, substantially underestimates

the experimental one, probably due to finite cluster effects. Drummond *et al.*[173] used the generalized-gradient approximation exchange-correlation functional of Perdew and Wang PW91, with the plane-wave basis set, and periodic boundary conditions to avoid disadvantages of employing point charges, of an arbitrary magnitude, to embed a finite, often small-size system. Both surface-methane interactions, and methane-methane interactions were investigated. The results confirmed the dipod configuration at the magnesium site to be 0.61 kJ/mol lower then tripod at oxygen site; used models contained one methane molecule per unit cell which corresponded to monolayer converges ($\theta = 1$). In case of larger unit cell, containing four methane molecules but corresponding to monolayer coverage, result predicts the ROT structure (where neighbouring methane molecules are rotated 90 degrees with respect to each other) to be the most stable, with the PAR structure (methane molecules with the same relative orientation) 4.98 kJ/mol higher in energy. The authors do not provide the interaction energy, referring to the general failure of standard DFT in properly describing the dispersion effects (underestimation of binding). Further studies included investigation of the second and third layer.[174, 175] Trevethan *et al.*[176] utilized DFT level of theory with the B3LYP hybrid functional in an embedded cluster approach implemented in the Gaussian Used for Embedded Systems Studies (GUESS) methodology. The authors point that in the case of the methane there is virtually no binding to the MgO surface at all. Total binding energy is 20 meV (1.93 kJ/mol) for monopod and 50 meV (4.82 kJ/mol) for dipod. Tripod configuration was not investigated. The authors points out that the computational method they employed (B3LYP hybrid-DFT) underestimates the dispersion interactions. The dispersion interaction is expected to contribute dominantly to the binding of hydrocarbon molecules to the surface. Therefore, they suggested that more sophisticated and expensive quantum chemistry methods are required to evaluate more accurate binding energies. Pisani *et a,.*[177] compared the physisorption of methane molecules at the (100) surface of MgO using three techniques: PBE, B3LYP, and Periodic Local MP2 (periodic LMP2) method. In contrast to the very weak binding obtained for the DFT level of theory (B3LYP, PBE), the periodic LMP2 calculations gave a pronounced adsorption to the surface. The dipod geometry had a minimum at 8.1 kJ/mol, which was nearly 3 kJ/mol lower in energy then the tripod configuration. By taking into account the attractive interaction energy per molecule between the methane molecules (calculated by Todnem *et al.*[172]) which also contributes to the adsorption energy, authors estimated the periodic LMP2 adsorption energy to about 11 kJ/mol, which gives even better agreement with the experiment. Tosoni *et al.*[178] examined the methane adsorption by a hybrid approach, *i.e.* by combining MP2 calculations with the extrapolation to the complete basis set limit with DFT+D calculations. The hybrid

energies obtained by the authors are sum of the low-level energy for the periodic structure and the high-level correction for the methane–cluster and methane–methane pair interaction. The final estimate to the adsorption energy was 13.3 kJ/mol. In order to compare this result to the measured Arrhenius desorption barriers, authors used thermal enthalpy contributions and a substantial zero-point energy calculated from DFT+D vibrational frequencies to obtain experimental desorption energies.

Table 1 contains summary of the interaction energy between methane and the MgO(001) surface obtained from experiments and previous theoretical investigations. Additionally, where it was available the Mg-C distances are reported.

Table 1. Summary of all interaction energies found by previous experimental and theoretical investigations.

Method		Interaction energy [kJ/mol]	Distance Mg-C [pm]
QMS/MSMS[161]	Experiment	13.12	-
TPD (monolayer)[167]	Experiment	12.63	-
TPD (zero coverage limit)[167]	Experiment	11.10	-
LCGTO-DF/BP[171]	Cluster	repulsive	-
PCI/ANO[172]	Cluster	8.5	290
GUESS/B3LYP[176]	Cluster	4.82	280
LMP2[177]	Periodic	11	300
PBE+D (monolayer)[178]	Periodic	13.98	309
MP2/CBS:PBE+D + ΔCCSD(T) + model corr.[178]	Cluster	13.31	329
CCSD(T):MP2//MP2:PBE+D (monolayer)[179]	Cluster	-15.7±0.7	310
CCSD(T):MP2//MP2:PBE+D (single molecule)[179]	Cluster	-13.0±0.6	-
CRYSTAL 09 (QZVP)[180]	Periodic	14.52	307

5.2. MgO(001) Surface Model

At monolayer coverage, one CH_4 molecule adsorbs above every second magnesium site. Therefore to model monolayer coverage, the smallest possible model of the surface has to contain two magnesium sites in the top layer. This gives an A×A surface cell, where A is two times the Mg-O distance. To describe two CH_4 molecules that are not related by symmetry a $A\sqrt{2}\times A\sqrt{2}$ surface cell consisting of 4 Mg^{2+}–O^{2-} units, is required. In our investigation we use a larger model surface – 2A × 2A supercell that contains in total 16 adsorption sites, 8 magnesium sites (green spheres) and 8 oxygen sites (red spheres) in the top layer (Figure 5). This is the smallest possible model allowing to describe 4 CH_4 molecules in the ROT structure, with methane molecules adsorbed at magnesium sites, in a dipod configuration (two hydrogen atoms pointing down to the surface in directions of oxygen atoms). Moreover, neighbouring methane molecules are rotated by 90° with respect to each other. This structure is referred as the lowest energy configuration.

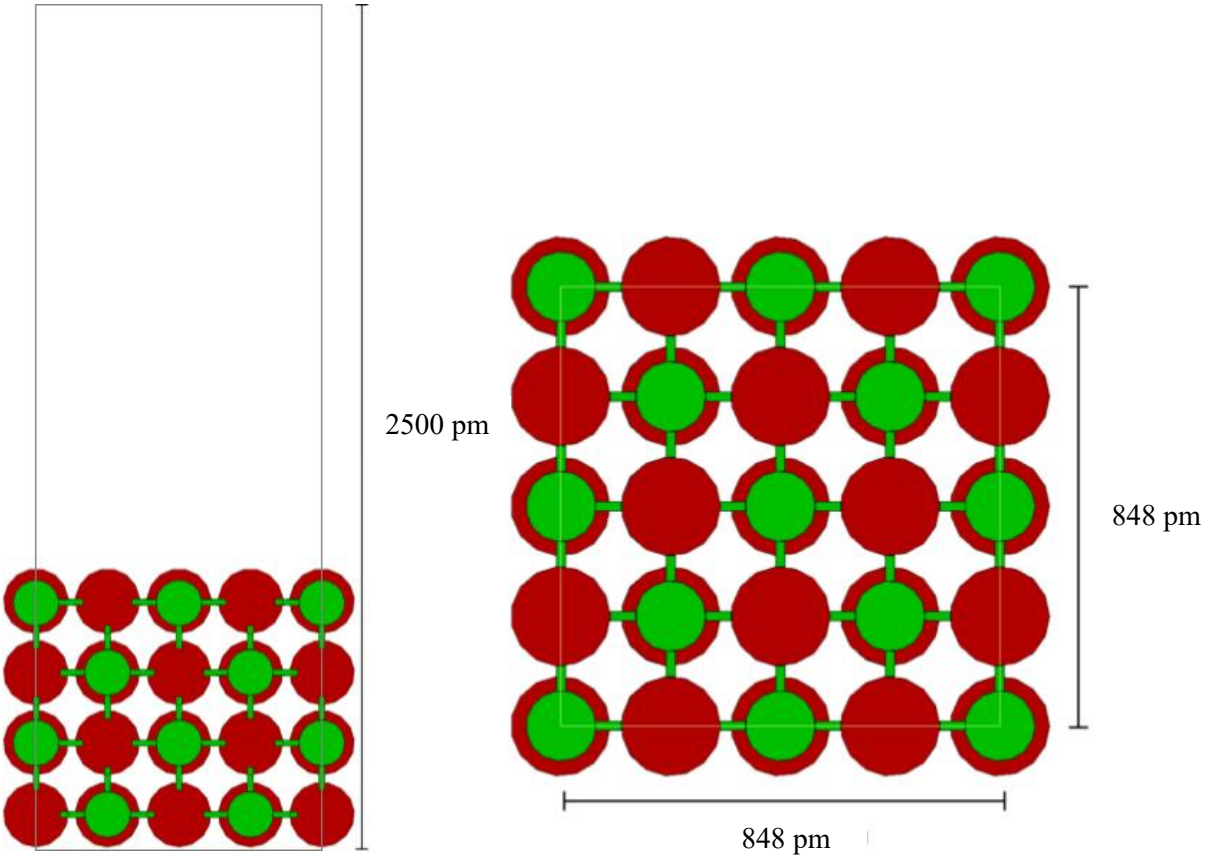

Figure 5. Model of the MgO(001) slab. Red dots represent O^{2-} ions ($\mathbf{124-128\ pm}$); green dots represent Mg^{2+} ions ($\mathbf{86\ pm}$). Supercell has $\mathbf{848\times848\times2500\ pm^3}$ size. Vacuum height, after subtracting surface height, is $\mathbf{186.5\ pm}$ which is enough for a proper description of the monolayer; no interaction takes place between upper and lower edges of the surface.

The MgO(001) surface model has high symmetry – plane symmetry group $p4m$ – with two rotational centres of order four, and reflections in four different directions (diagonals, horizontal, and vertical). The high symmetry of the surface was taken into account during the similarity recognition step. To stay consistent within the method the lattice parameter for bulk MgO (without adsorbed methane) was optimized using the PBE+D method yielding a value of 4.24 pm.

Additionally, for the adsorption of a single molecule we used a smaller unit cell - the A × A super cell, in order to compare results with previous calcualtions done by Drummond *et al.*[173] This cell corresponds to monolayer coverage of methane. As already mentioned, the disadvantage of a small unit cell is that the ROT structure could not be investigated, due to translational symmetry.

5.3. Methods

Global structure optimizations is performed using a genetic algorithm of our implementation (DODO).[131] Local structure optimizations are performed using the QMPOT[152] program. Both programs were accustomed to work within the rigid-body framework. The Vienna Ab initio Simulation Package (VASP),[181, 182] with the Perdew, Burke, and Ernzerhof (PBE) exchange correlation functional,[183, 184] and electron-ion interactions described by the projector augmented wave method (PAW),[185, 186] is used to evaluate single point energies and gradients. Only the valence electrons are explicitly considered. A semiempirical C_6-term is added to include long-range dispersion contributions as suggested by Grimme (PBE+D),[187] and implemented by Kerber for periodic systems. Following previous study C_6 parameters for Ne atom were used for Mg^{2+} ions.[178]

During the GA step, all calculations are performed with $1 \times 1 \times 1$ Monkhorst-Pack k-point mesh, which, for the final local structure optimization, was increased to a $2 \times 2 \times 1$ mesh.

Local optimizations, during global optimization, are performed using soft potentials with an energy cutoff of 300 eV and a global threshold for energy change was set to 10^{-4} eV/Å, so as to assure good accuracy – time ratio. This is important due to vast number of local optimizations during each GA run. The final structure optimizations are obtained with standard potentials and a higher energy cutoff of 600 eV. The stationary point is localized if the forces acting on ions are smaller than 5×10^{-3} eV/Å.

After the final optimization step, many structures still had imaginary frequencies. Therefore, further structure optimization had been performed using Normal Mode optimization code.[188] The maximum gradient component threshold for Normal Mode calculations was set to 1×10^{-4} eV/Å in gradient. Nevertheless, in some cases it was impossible to remove all imaginary frequencies using this method.

The Normal Mode optimization[189] introduce the harmonic approximation in order to define transformation between coordinates. Starting from the molecular harmonic vibrational Hamiltonian, and introducing the normal mode coordinates, the linear relation between the Cartesian displacements and the normal mode coordinates can be found. The search for the energy minimum is based on quadratic approximation. The normal mode gradient and the second derivatives of the energy at an optimization point, defines the step towards the minimum energy. The step is described by the RFO method (the modified Newton–Raphson formula). Normal Mode optimization achieves better accuracy and therefore allows for tighter convergence criterions.

Table 2. Number of imaginary wavenumbers with respect to step size used to calculate finite differences (POTIM parameter). Reference structure: global minimum for six methane molecules the MgO(001) surface (Figure 16).

Step Size [nm]	Number of imaginary wavenumbers	Imaginary wavenumbers [cm^{-1}]
0.0005	3	19, 32, 34
0.0010	1	14
0.0015	0	-
0.0020	1	4
0.0025	2	9, 16
0.0030	2	9, 35
0.0035	2	18, 54
0.0040	4	8, 21, 43, 70

To calculate the Hessian matrix, finite differences are used. Each ion is displaced in the direction of each Cartesian coordinate. The obtained forces are used to determine the Hessian matrix. In this study, the IBRION parameter was set to five, *i.e.* all atoms are displaced in all

three Cartesian directions, without any symmetry considerations. The POTIM parameter, which determines the step size used to calculate forces, was set to 0.0025 nm. As showed by later investigation, reduction of the step size to 0.0015 nm allows for elimination of imaginary wavenumbers (Table 2). Although, the test was performed only on one structure and further investigation would be needed.

5.4. Optimization Strategy

Each GA investigation consists of 4 subsequent steps. The first step is pre-optimization. At this stage each structure (methane, ethane, propane and surface models) are optimized separately. The second step is the genetic algorithm optimization. During that stage a vast number of local localizations have to be performed, with each one of them as an independent job. The global optimization program can easily control the number of local optimizations, which are done at the same time. Therefore, the global search can be easily parallelized over a number of independent local optimization jobs. To ensure fast convergence; and, consequently, good performance of the explorations of potential energy surface (PES), each structure optimization is performed using low convergence criteria. At the third stage, refinement, around 100 structures, found by genetic algorithm, with the lowest energies are recalculated using tighter convergence criteria, in order to obtain better relaxed structures. The fourth and final step is full relaxation, where atoms are allowed to fully relax (except the two bottommost layer of the MgO, which simulate bulk characteristic). Around 20 of the most energetically favourable structures, from the refinement step, undergo a full relaxation with the same, tight convergence criteria. In case of methane, additional optimization with Normal Mode code was performed to eliminate imaginary wavenumbers.

The algorithm starts with a population of 100 structures. Initial structures are obtained by random distribution of 4 rigid bodies on the MgO(001) surface model and optimized to the local minimum within the basin of attraction. One generation evolves to another through crossover operations. Two structures, called parents, are selected from the population to create one child. The selection probability is proportional to the value of the fitness function of the given structure. A constant number of children are allowed to mutate, through random rotations of 1 or 2 rigid bodies with respect to their centres of mass.

5.5. Thermodynamic Properties - Microscopic–Macroscopic Connection

It is possible under certain circumstances to observe single molecules in the experiment, nevertheless, the vast majority of experimental methods measure macroscopic quantities of matter made of large numbers of molecules. The behavior of such ensembles of molecules is governed by the laws of thermodynamics. Chemical reactions and chemical properties can be defined in terms of fundamental variables of thermodynamics, such as enthalpy, entropy, free energy.

Transition from the microscopic regime to the macroscopic is made by recognition that the Born–Oppenheimer potential energy surface, although defined by electronic energy function of nuclei coordinates, is fundamentally a classical construct. The motion of the nuclei on the PES surface is also accounted for in a quantum mechanical way, by molecular vibrations, even at zero temperature, since the lowest vibrational energy level for any bound vibration is non-zero. The energy of the lowest vibrational level, in the harmonic oscillator approximation, can be determined as $1/2\,\hbar\omega$ (where $\hbar$ is Planck's constant divided by 2π, and ω is the vibrational frequency). Therefore, the sum of energies of all molecular vibrations defines the zero-point vibrational energy (ZPVE). The internal energy at 0 K for a molecule can then be defined as

$$U_{T=0} = E_{elel} + \sum_{i}^{modes} \frac{1}{2}\hbar\omega_i \tag{39}$$

The harmonic oscillator approximation introduces implicit errors in calculations of ZPVE. Nevertheless, the weak modes, which are *least* harmonic in character, have very small vibrational frequencies associated with them, and therefore do not contribute much to the ZPVE (since it is linear in the frequencies). However, small frequencies have a significant influence on change in entropy.

To describe collections of molecules using the statistical mechanics, certain macroscopic variables must be held constant by external influence. The enumeration of these conditions defines an 'ensemble'. One commonly used type of it is the 'canonical ensemble', sometimes referred to as the NVT ensemble, where the constants are the total number of particles N, the volume V, and the temperature T. For the canonical ensemble, it is written as

$$Q(N,V,T) = \sum_{i} e^{-E_i(N,V)/k_B T} \tag{40}$$

where i runs over all possible energies of the system. By assuming that the ensemble is an ideal gas of N molecules (*i.e.* they do not interact with one another), the partition function may be rewritten as molecular partition function q.

$$Q(N,V,T) = \frac{[q(V,T)]^N}{N!} \tag{41}$$

Therefore, reduce the problem of finding the ensemble partition function $Q(N,V,T)$ to finding the molecular partition function $q(V,T)$. By assuming that the molecular energy can be expressed as a sum of electronic, translational, rotational, and vibrational energies, a further simplification of the problem can be made.

$$q(V,T) = q_{elec}(T) \cdot q_{trans}(V,T) \cdot q_{rot}(T) \cdot q_{vib}(T) \tag{42}$$

and individual components – electronic, translational, rotational, and vibrational – of thermodynamic functions can be treated separately, within the respective approximations.

The partition function characterizes statistical properties of a macroscopic system in thermodynamic equilibrium; and allows thermodynamic functions of the system (U – the total energy, G – Gibbs free energy, S – entropy) to be expressed as its function or its derivatives. Within the canonical ensemble, and by thermodynamic definitions,

$$U = k_B T^2 \left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} \tag{43}$$

$$H = U + pV \tag{44}$$

$$S = k_B \ln Q + k_B T \left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} \tag{45}$$

$$G = H - TS \tag{46}$$

The transformation of the potential energy surface, determined by the electronic structure calculation, to produce thermodynamic data is straightforward. It requires an optimized structure with its associated vibrational frequencies. Since experimental data are typically measured as heat (enthalpy) of formation (ΔH_f) or free energy of formation (ΔG_f), it is usually worthwhile to compute the frequencies and then the thermodynamic variables.

In practice, what we are interested in is to find the standard enthalpy change of reaction, or any other thermodynamic function,

$$\Delta H = H_{\text{product}} - H_{\text{substrat}} \quad (47)$$

In this work, change in any thermodynamic function (X) is expressed as

$$\Delta X = \frac{X(M_N \cdot S) - X(S) - N \cdot X(M)}{N} \quad (48)$$

Where $X(M_N \cdot S)$ – is the value of a thermodynamic function for the surface with adsorbate layer consisting of N molecules per unit cell, $X(S)$ – is the value of a thermodynamic function for a clean surface, $X(M)$ – is the value of a thermodynamic function for the molecule.

Unfortunately, it was impossible to eliminate all imaginary frequencies for some of the structures. This is crucial, since these frequencies do not contribute to calculations of thermodynamic functions. Additionally, fact that they are present testifies that in some cases we are missing few small frequencies; and this, in turn, has an influence on entropy calculation, as changes in value of small frequencies have a huge influence on change in entropy.

5.6. Interaction Energy

In this work the formation energy, ΔE_{form}, is defined as the energy of surface with adsorbate layer consisting of N molecules per unit cell, $E(M_N \cdot S)$, minus the energy of a clean surface ($E(\text{S})$), and minus the energy of the adsorbed molecules ($N \cdot E(M)$)

$$\Delta E_{form} = E(M_N \cdot S) - E(S) - N \cdot E(M) \quad (49)$$

Using definition of the formation energy, we can define the interaction energy per molecule, ΔE, as:

$$\Delta E = \frac{\Delta E_{\text{form}}}{N} \quad (50)$$

Which can allows for an easy comparison between models with different number of molecules. The interaction energy can be further decomposed into the interaction energy between the surface and the whole adsorbate layer, ΔE^*, and the lateral interaction, ΔE_L,:

$$\Delta E = \Delta E^* + \Delta E_L \tag{51}$$

$$\Delta \mathrm{E}^* = \frac{\mathrm{E}(\mathrm{M_N} \cdot \mathrm{S}) - \mathrm{E}(\mathrm{S}) - \mathrm{E}(\mathrm{M_N}//\mathrm{M_N} \cdot \mathrm{S})}{\mathrm{N}} \tag{52}$$

$$\Delta E_L = \frac{E(M_N//M_N \cdot S) - N \cdot E(M)}{N} \tag{53}$$

The notation $M_N//M_N \cdot S$ implies the calculation of a layer structure (M_N) with the same geometry as if in the presence of the surface ($M_N \cdot S$).

The lateral interaction describes energy change due to interaction of the molecules in the layer, which can be both, attractive and repulsive. Similarly, the surface-layer interaction describes interaction of the whole layer with the surface.

5.7. Methane at the MgO(001) Surface

The first step of the investigation was to examine the adsorption of a single molecule at the MgO(001) surface. To do this, two models of the surface were used, one corresponding to one-fourth coverage, second to monolayer coverage. Obtained PBE results, with and without dispersion corrections, were related to previous theoretical studies by Drummond *et al.*[173]

The next step was to model monolayer coverage with four methane molecules and the $2A \times 2A$ unit cell from the single molecule step. All three independent GA runs predicted the ROT structure to be the global minimum. This result is in an agreement with what earlier studies predicted to be the most stable monolayer structure.[173]

Subsequently, higher CH_4 coverage of the MgO(001) surface was analysed: with five and six methane molecules. The global minimum structure for five molecules contains the ROT structure as the first adsorption layer.

Motivated by the results of five molecules on top of the MgO surface, we performed additional GA investigation with additional one, two, and four methane molecules on top of the ROT monolayer structure (neighbouring molecules rotated by 90° with respect to each other). The resulted structures contained 4+1, 4+2, amd 4+4 methane molecules. This way, we were able to observe the formation of the second layer of methane.

For ethane on the MgO surface, in the GA runs, each ethane molecule was modelled as two separate rigid bodies, *i.e.* as two methyl groups. The presented results contain structures with one, two, and four molecules in the 2A × 2A MgO supercell.

Finally, propanes, in the staggered conformation, were added to the MgO surface; with four different loadings spanning from one to four molecules in the 2A × 2A MgO supercell. Results for three and four propane molecules are only preliminary, and will need further investigation in the future.

5.7.1. Single Molecule Coverage

Interaction of a single methane with the MgO(001) surface was subject of a number previous investigations.[172, 173, 178]

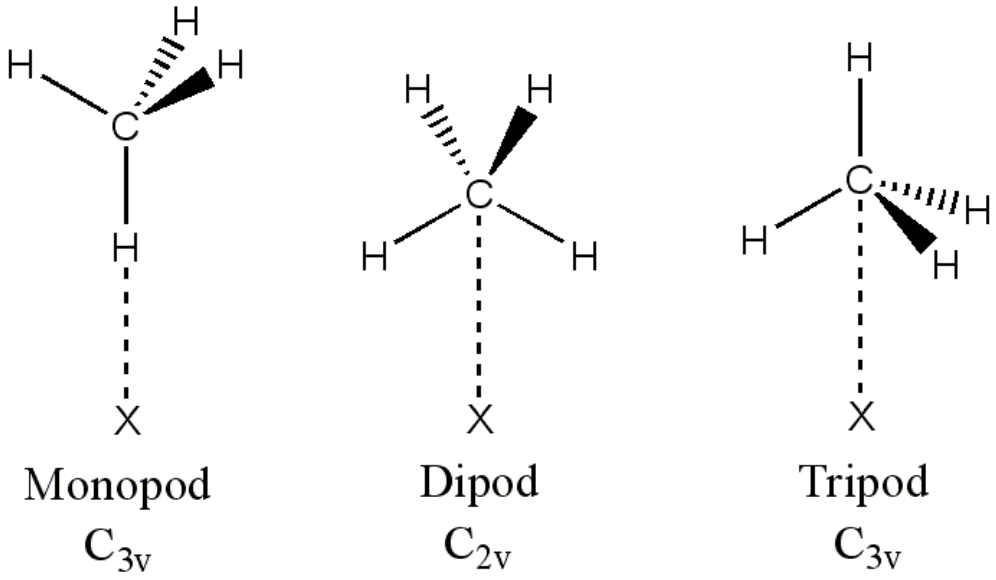

Figure 6. Possible orientations of the methane molecule at the MgO(001).

Figure 6 shows the three possible on-top orientations of the methane molecules. Bridge and hollow sites were found as disfavoured with respect to on-top binding as presented in the previous theoretical investigations.[173] Also during this investigation bridge and hollow sites were not found in any of the obtained structures. In most of the results, the dipod configuration, on-top magnesium sites with two hydrogen atoms pointing into directions of adjacent oxygen sites, was predominant. Some of the molecules were significantly tilted, with third hydrogen atom pointing in the direction of yet another oxygen site. Also monopod configurations were present, especially for the structures with the second layer of methane. All molecules were adsorbed in on-top positions above magnesium sites. Other configurations were not observed.

The single molecule adsorption was investigated using two models of the surface: the smaller one – $424 \times 424 \times 2500\ pm^3$ unit cell – corresponds to monolayer coverage ($\theta = 1$); the larger one – $848 \times 848 \times 2500\ pm^3$ – corresponds to one quarter of the monolayer coverage ($\theta = 0.25$). Molecules were allowed to optimize freely, i.e. no local symmetry was imposed on them. Structure names indicate local symmetries of the initial configurations.

Table 3. Relative energies in kJ/mol (and distances C-X in pm) for six arrangements of methane above the MgO(001) surface. Smaller supercell: 4 layers slab, $\mathbf{4 \times 4 \times 1}$ k-point grid. Larger supercell: 4 layers slab, $\mathbf{2 \times 2 \times 1}$ k-point grid. Abbreviations mono, di, and tri stand for monopod, dipod, and tripod, respectively. Mg and O stand for on-top binding above magnesium and oxygen sites, respectively.

Supercell $[pm^3]$	Functional	di Mg	mono O	di O	tri Mg	tri O	mono Mg
$424 \times 424 \times 2500$ $(\theta = 1)$	PW91 [1]	0.00	0.61	0.63	1.03	2.40	2.95
	Distances C-X	(325)	(350)	(380)	(325)	(355)	(369)
	PBE+D[2]	0.00	12.47	15.87	2.56	15.14	6.19
	Distances C-X	(307)	(357)	(360)	(349)	(351)	(374)
	PBE+D[3]//PBE+D[2]	0.00	21.34	25.16	9.27	24.07	13.39
	PBE//PBE+D[2]	1.86	9.47	12.98	0.00	12.10	3.30
$848 \times 848 \times 2500$ $(\theta = 0.25)$	PBE+D[2]	4.15	0.00	9.82	5.85	1.73	4.50
	Distances C-X	(302)	(359)	(366)	(307)	(359)	(478)
	PBE+D[3]//PBE+D[2]	8.50	0.00	8.34	9.27	1.49	0.68
	PBE//PBE+D[2]	0.74	0.00	16.49	4.57	1.81	10.46
CH_4/Mg_9O_9	MP2[4]	0.00	---	---	1.48	---	4.26
	Distances C-X	(325)	---	---	(350)	----	(334)

[1] Drummond *et al.*[173] (6 layers slab, 8 × 8 × 1 k-point grid)

[2] Tosoni's dispersion[178]

[3] Standard dispersion

[4] BSSE corrected MP2/aug'-cc-pVQZ energy, only C-Mg distance optimized (ESI Tosoni *et al.*[178])

Results for the smaller unit cell (Table 3) confirm that the dipod configuration on top of magnesium site is favourable arrangement, both for PW91 and PBE+D results. The energy differences between the structures are more pronounced for PBE+D results. The PW91 approach predicts the monopod configuration on oxygen site, to be the next higher in energy minimum, but only by 0.61 kJ/mol, whereas, the PBE+D functional predicts the tripod configuration at magnesium site to be the second one in energy, exceeding the minimum by additional energy of 2.56 kJ/mol. PW91 and PBE are generalized gradient approximation (GGA) functionals. Although both functionals are independent from each other, they should be numerically equivalent for most purposes.[190] Therefore, the difference in energies and energetic ordering of the structures can be attributed to dispersion correction used for PBE results.

PBE+D results for the larger unit cell points to the the monopod configuration as the preferred one over the dipod configuration by 4.15 kJ/mol. Tripod configuration over oxygen site lies between them, 1.73 kJ/mol above the monopod. PBE results without dispersion, still give the monopod configuration as the energy most preferable situation. Nevertheless, this time the dipod Mg configuration is only 0.74 kJ/mol higher in energy. In general, removing dispersion correction, or altering to standard one, does not change the global minimum, only relative order of other structures.

During optimization for the larger unit cell, the tripod configuration at the magnesium site yielded the dipod one, also at magnesium site but slightly tilted. The energy difference between these two structures is small (1.70 kJ/mol). Similarly, the dipod configuration at oxygen site yielded the monopod configuration at the same site. The dipod C_{2v}-O configuration (which is in fact monopod configuration) is 7 pm higher above the oxygen site than the monopod C_{3v}(1)-O structure. Additionally, in contrast to the monopod C_{3v}(1)-O structure which do not have any symmetry, the dipod C_{2v}-O structure possess C_s symmetry with respect to the surface. The energy difference between these structures is large by 9.82 kJ/mol.

Last row of Table 3 shows results by Tosoni *et al.*[178] obtained with MP2 on the Mg_9O_9 cluster model with one methane molecule (only at magnesium site). The energy separation between dipod and tripod for MP2 is similar with PBE+D results (1.70 kJ/mol), but the MP2 method predicted monopod to be significantly higher in energy than PBE+D, 4.26 and 0.45 kJ/mol, respectively. This difference might be due to the fact that, when fully optimized, methane molecule is tilted. In case of the MP2 results, one C-H bond, pointing towards the surface, is parallel to the surface normal.

Table 4. PBE+D interaction energies in kJ/mol. During optimization dipod C_{2v}-O strucutre yield monopod $C_{3v}(1)$-O structure (for larger unit cell). Tripod $C_{v3}(3)$-Mg structure optimized towards tilted dipod C_{2v}-Mg structure (for larger unit cell).

Supercell $[pm^3]$		dipod C_{2v}-Mg	monopod $C_{3v}(1)$-O	dipod C_{2v}-O	tripod $C_{3v}(3)$-Mg	tripod $C_{3v}(3)$-O	monopod $C_{3v}(1)$-Mg
$424 \times 424 \times 2500$ $(\theta = 1)$	interaction energy	-14.03	-1.56	1.84	-11.46	1.12	-7.84
	interaction surface – layer	-7.41	5.28	8.39	-4.63	7.95	-1.29
	lateral interaction	-6.62	-6.84	-6.54	-6.83	-6.83	-6.55
	Im freqeuncies	1	2	4*	3	2	4*
$848 \times 848 \times 2500$ $(\theta = 0.25)$	interaction energy	-12.61	-16.76	-6.94	-10.91	-15.03	-12.26
	interaction surface – layer	-12.82	-16.78	-6.96	-11.07	-15.03	-12.26
	lateral interaction	0.21	0.02	0.02	0.16	0.01	0.00
	Im freqeuncies	0	2	2	1	2	5

*not fully optimized

Lateral interactions for the $424 \times 424 \times 2500$ unit cell are very similar for all considered configurations. In case of the big unit cell the lateral interactions are slightly positive due to the deformation of the methane when in contact with the surface.

Even though structures were located as stationary points all of them, except one, are no local minima (Table 5). Only the dipod-Mg configuration, for the $848 \times 848 \times 2500$ unit cell, is a minimum. Many of the obtained structures are higher stationary points – with two and more imaginary frequencies. Increasing the number of single point used for Hessian evaluation to 4 points (Table 6) did not result in a significant decrease in the number of imaginary frequencies. What is worth mentioning is that now the dipod-Mg configuration of the smaller unit cell is the only minimum.

Table 5. Values of imaginary wavenumbers in $\boldsymbol{cm^{-1}}$ obtained with two points central numerical differentiation for single methane molecule at the MgO(001) surface. Two top-most layers of MgO were unfrozen.

Supercell $[pm^3]$	Imaginary wavenumber					
	dipod C_{2v}-Mg	monopod C_{3v}(1)-O	dipod C_{2v}-O	tripod C_{3v}(3)-Mg	tripod C_{3v}(3)-O	monopod C_{3v}(1)-Mg
$424 \times 424 \times 2500$ $(\theta = 1)$	46	28 86	30 71 82 148	37 69 94	28 62	20 47 146 162
$848 \times 848 \times 2500$ $(\theta = 0.25)$	-	23 51	20 45	6	7 48	19 26 37 52 66

Table 6. Values of imaginary wavenumbers in $\boldsymbol{cm^{-1}}$ obtained with four points central numerical differentiation for single methane molecule at the MgO(001) surface. Two top-most layers of MgO were unfrozen.

Supercell $[pm^3]$	Imaginary wavenumber					
	dipod C_{2v}-Mg	monopod C_{3v}(1)-O	dipod C_{2v}-O	tripod C_{3v}(3)-Mg	tripod C_{3v}(3)-O	monopod C_{3v}(1)-Mg
$424 \times 424 \times 2500$ $(\theta = 1)$	-	23 42	42 63 82 132	37 77 87	32 76	40 150 160
$848 \times 848 \times 2500$ $(\theta = 0.25)$	44	18	63	53	51 57	53 54 77

Imaginary wavenumber for the dipod C_{2v}-Mg structure, 46 cm^{-1} (Table 5) and 44 cm^{-1} (Table 6), are rocking motions of a molecule through the vertical plane. Two imaginary wavenumbers for the monopod C_{3v}(1)-O, in range of 18-28 cm^{-1} and 42-86 cm^{-1}, correspond to rocking and twisting motions, respectively.

Based on electronic energy alone dipod-Mg is the most stable configuration for the smaller supercell (Table 7), with stabilization energy of 14.03 kJ/mol, *i.e.* 2.57 kJ/mol higher than the tripod-Mg. However, after including corrections from ZPVE and thermal contributions this dissimilarity diminishes to 0.16 kJ/mol in the total energy and enthalpy.

Table 7. Thermodynamic adsorption data in kJ/mol. Calculated for $\boldsymbol{T = 47\ K}$ and $\boldsymbol{p = 0.1 MPa}$ for the $\boldsymbol{424 \times 424 \times 2500\ pm^3}$ supercell.

supercell	dipod C_{2v}-Mg	monopod C_{3v}(1)-O	dipod C_{2v}-O	tripod C_{3v}(3)-Mg	tripod C_{3v}(3)-O	monopod C_{3v}(1)-Mg
ΔE_{el}	-14.03	-1.56	1.84	-11.46	1.12	-7.84
ΔE_{ZPV}	3.06	2.18	0.43	0.47	1.11	0.44
ΔU	-11.70	-0.16	1.49	-11.54	1.79	-8.24
ΔG	-5.05	6.49	8.12	-5.30	7.67	-1.48
ΔH	-12.09	-0.55	1.10	-11.93	1.40	-8.63
$-T\Delta S$	7.03	7.03	7.01	6.62	6.26	7.13
$\Delta U - \Delta E_{\mathrm{el}} - \Delta E_{\mathrm{ZPV}}$*	-0.73	-0.78	-0.78	-0.55	-0.44	-0.84
$\Delta U - \Delta E_{\mathrm{el}}$	2.33	1.40	-0.35	-0.08	0.67	-0.40
$\Delta H - \Delta E_{\mathrm{el}}$	1.94	1.01	-0.74	-0.47	0.28	-0.79
$\Delta G - \Delta E_{\mathrm{el}}$	8.98	8.05	6.28	6.16	6.55	6.36

*thermal contribution

For the big unit cell the difference between dipod-Mg, the one that is thought to be the most favourable configuration of a single molecule, and monopod-O is even more pronounced (Table 8); from 4.15 kJ/mol at the electronic energy level to 5.53 kJ/mol for the total energy.

Table 8. Thermodynamic adsorption data in kJ/mol. Calculated for $\boldsymbol{T = 47\ K}$and $\boldsymbol{p = 0.1 MPa}$ for the $\boldsymbol{848 \times 848 \times 2500\ pm^3}$ supercell.

supercell	dipod C_{2v}-Mg	monopod C_{3v}(1)-O	dipod C_{2v}-O	tripod C_{3v}(3)-Mg	tripod C_{3v}(3)-O	monopod C_{3v}(1)-Mg

ΔE_{el}	-12.61	-16.76	-6.94	-10.91	-15.03	-12.26
ΔE_{ZPV}	3.70	2.24	1.22	3.32	1.18	0.22
ΔU	-9.75	-15.28	-6.44	-8.06	-14.47	-12.93
ΔG	-2.93	-8.64	0.09	-2.04	-7.98	-6.22
ΔH	-10.14	-15.67	-6.83	-8.45	-14.85	-13.32
$-T\Delta S$	7.21	7.03	6.92	6.41	6.87	7.10
$\Delta U - \Delta E_{\mathrm{el}} - \Delta E_{\mathrm{ZPV}}$	-0.84	-0.76	-0.72	-0.47	-0.62	-0.89
$\Delta U - \Delta E_{\mathrm{el}}$	2.86	1.48	0.50	2.85	0.56	-0.67
$\Delta H - \Delta E_{\mathrm{el}}$	2.47	1.09	0.11	2.46	0.18	-1.06
$\Delta G - \Delta E_{\mathrm{el}}$	9.68	8.12	7.03	8.87	7.05	6.04

5.7.2. Monolayer Coverage

Monolayer coverage of the methane molecules at the MgO(001) surface was intensively investigated during validating performance of the RBGA implementation. As the result three independent GA runs were performed; one up to 700 calculated structures, two up to 600 calculated structures. All runs predict the ROT structure as the global minimum for this system. The ROT structure consists of methane molecules adsorbed in the dipod configuration at magnesium sites. Every second magnesium site is occupied (monolayer coverage). Additionally, each methane molecule is rotated 90 degrees with respect to the surface normal passing through the carbon atom.

Although some of the structures possessed small imaginary frequencies (mostly within range of 5 to 20 cm^{-1}), the structures were not optimized using Normal Mode optimization. In this cases, the NMopt optimization yielded structures with more imaginary frequencies than the starting one.

5.7.2.1. Structures

The final refinement of populations, each consisting of 60 structures, resulted in 20 distinguishable final structures. Eight with the lowest relative energies are presented in Figure 7 –Figure 14. All the structures contained methane molecules in the dipod configuration. Some of the slightly distorted – titled – due to the repulsive interaction between hydrogen atoms. There were structures with higher energy (not showed here) that contained methane molecules in monopod configuration, but these were rather negligible cases.

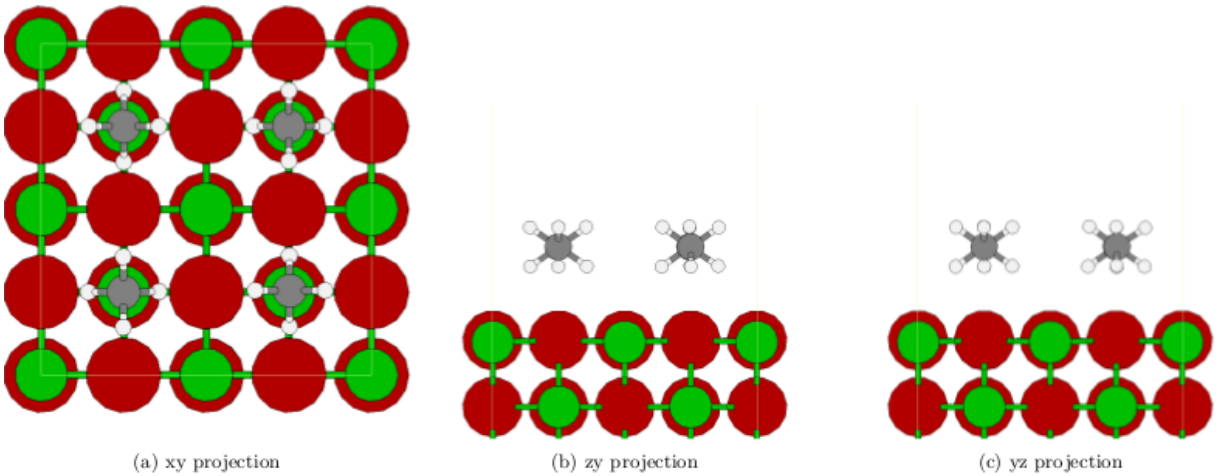

Figure 7. Structure number 1 (ROT) - global minimum

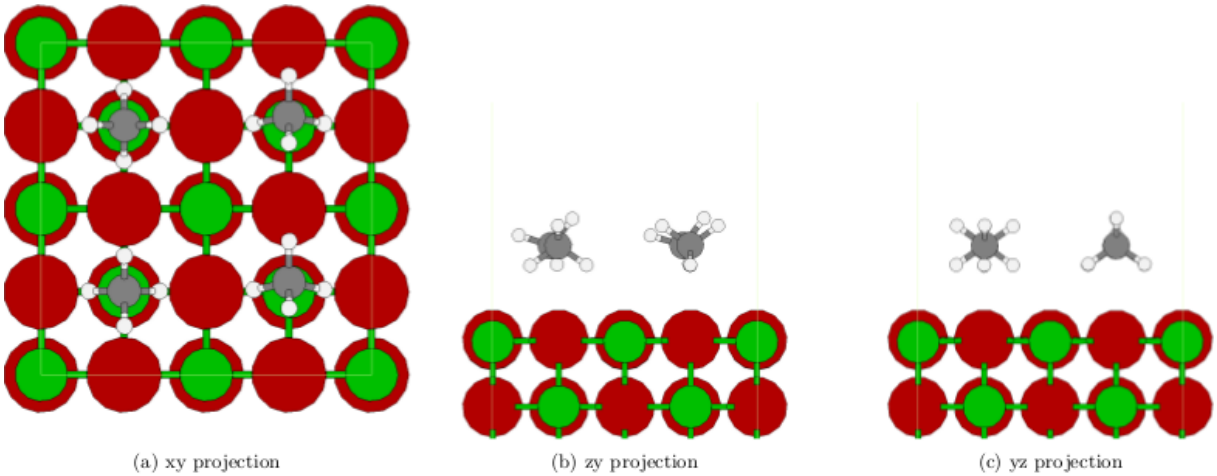

Figure 8. Structure number 2 (1.51 kJ/mol)

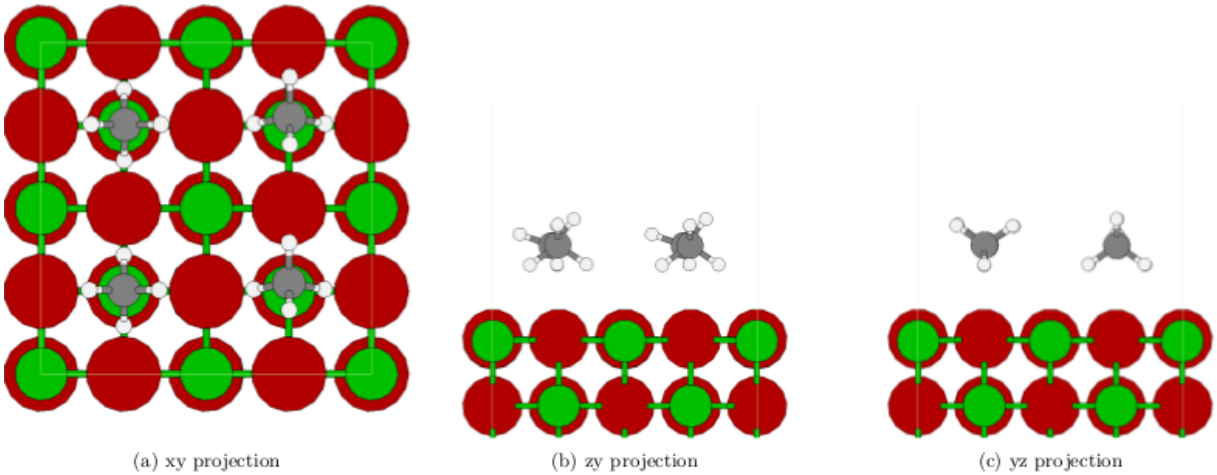

Figure 9. Structure number 3 (1.57 kJ/mol)

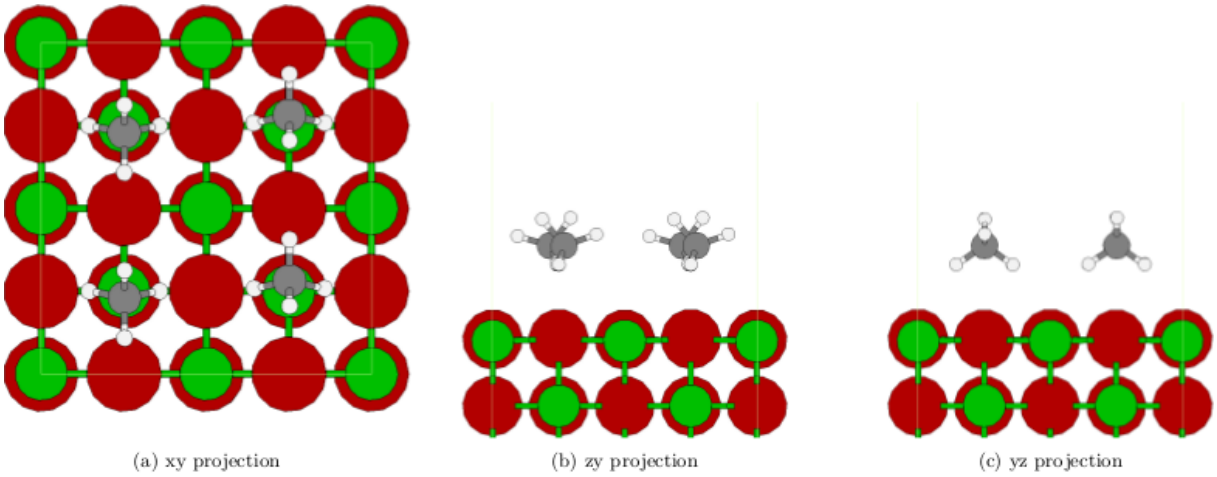

Figure 10. Structure number 4 (PAR) (3.10 kJ/mol)

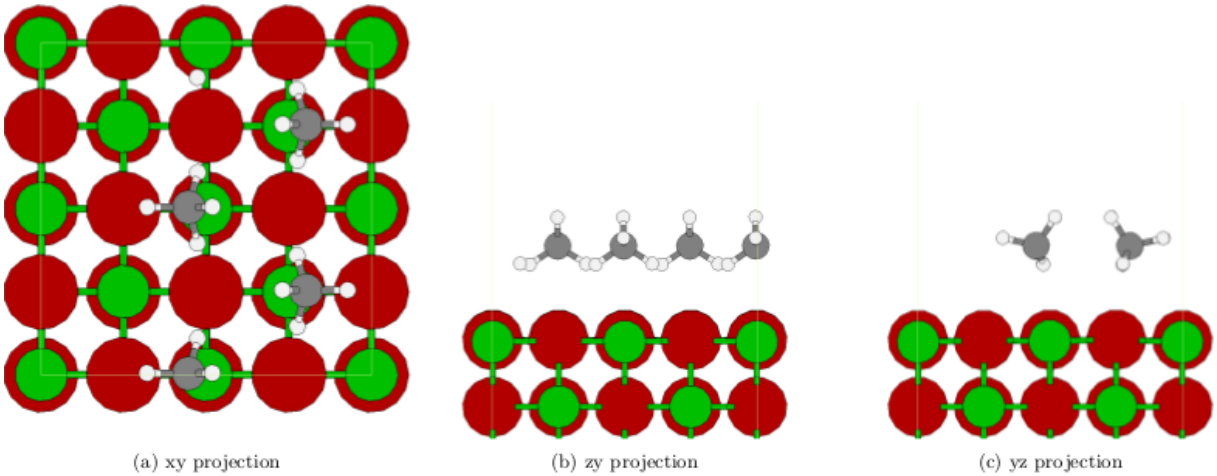

Figure 11. Structure number 5 (WIRE PAR) (4.57 kJ/mol)

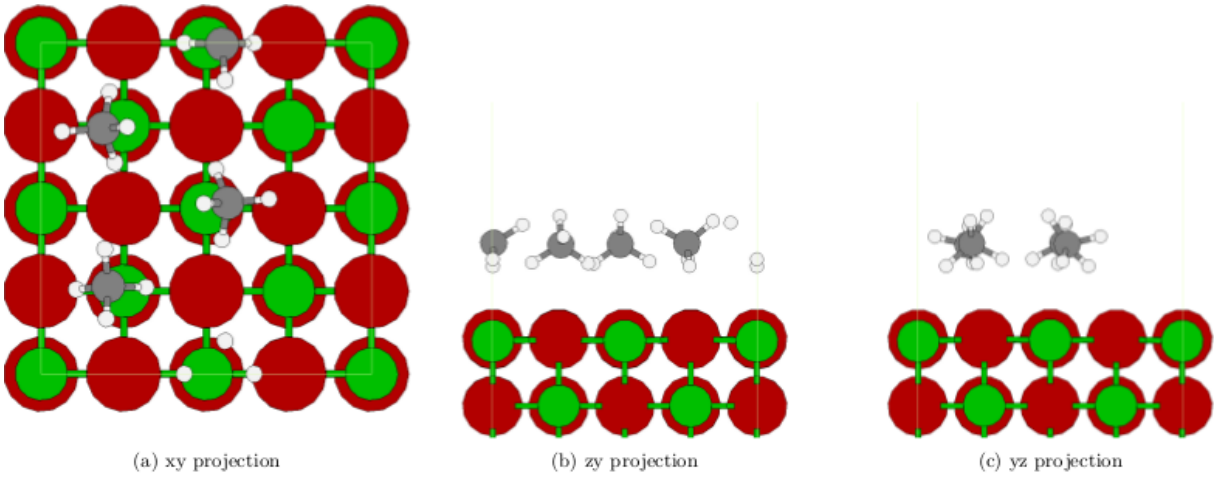

Figure 12. Structure number 6 (5.12 kJ/mol)

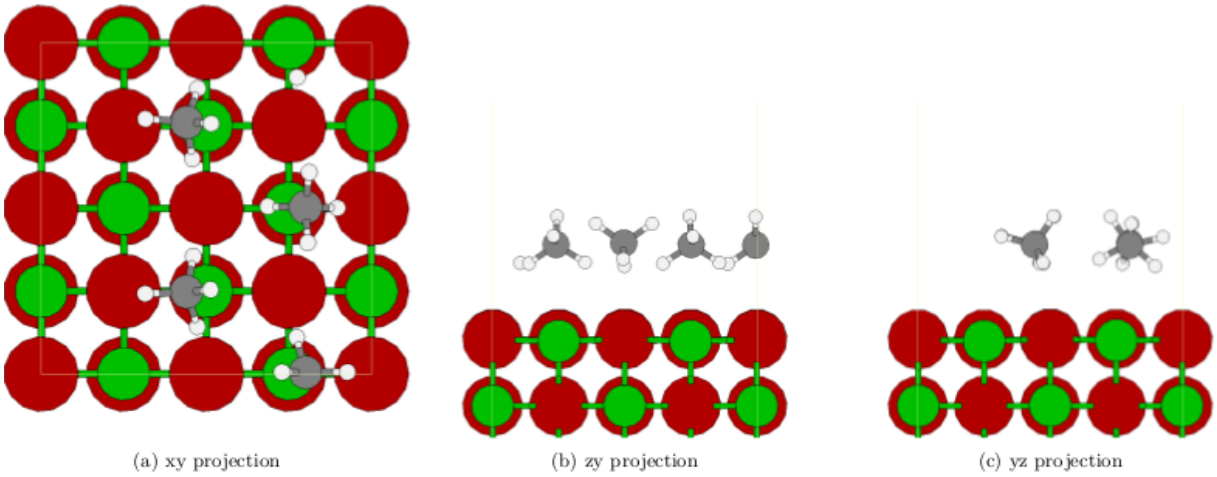

Figure 13. Structure number 7 (5.37 kJ/mol)

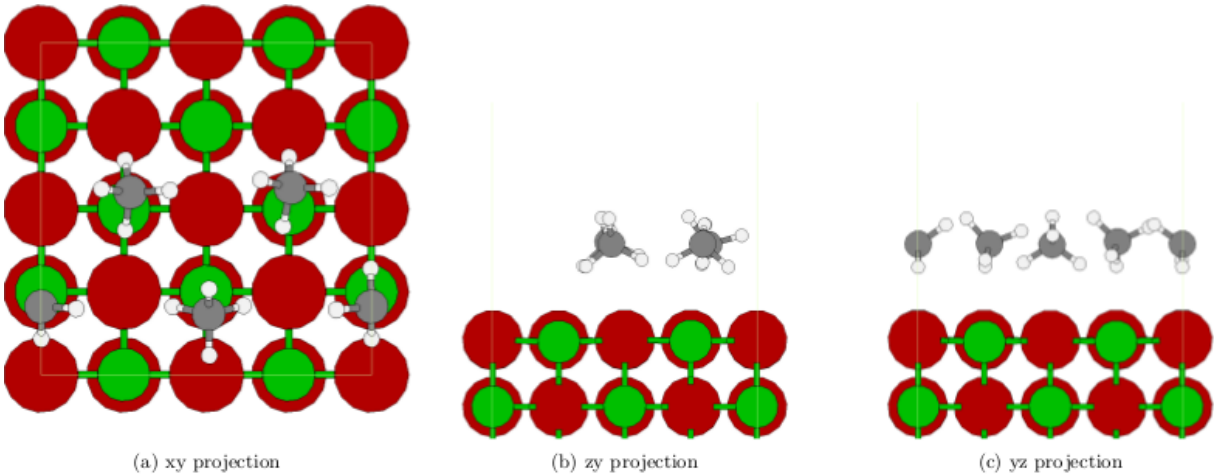

Figure 14. Structure number 8 (6.00 kJ/mol)

The first four structures are variations of possible dipod configurations of the methane molecules adsorbed at every other magnesium site, maximizing the distance between carbon atoms. Further four have methane adsorbed at adjacent magnesium sites, forming rows on the surface. Other structures, not shown, occasionally had methane molecules adsorbed in monopod configuration at oxygen sites. The methane molecules adsorbed at the ROT structure (Figure 7) are rotated with respect to each other 90 degrees around the normal axis to the surface going through the carbon atom. This way distances between neighboring are minimized, therefore minimizing repulsive interaction. Although methane molecules at the PAR structure (Figure 10) are not rotated with respect to each other, they are tilted so as to minimize repulsive interaction between hydrogen atoms of different methane molecules, by increasing the distance between them. The WIRE-type structures (Figure 11 –Figure 14) have methane molecules adsorbed in the dipod configuration; nevertheless all of the configurations are distorted due to repulsive interaction between hydrogen atoms.

5.7.2.2. Energies

As we go along the structures numbers in the Table 9, the interaction energy steadily decreases; However by relatively small number. Along with declining interaction energy both surface-layer interaction and lateral interaction decline as well, but even at smaller rate. The change in total relative energies for structures 5-8 is less pronounced than in case of previous four structures. This is expected since molecules are very close to each other, so there is not much rotational freedom to try to decrease the repulsion between the molecules. Similarly, a

small difference in the interaction energy for structures 2 and 3 is not surprising. They differ by rotating 90 degree around C_{2v} axis of one of the methane molecules.

Table 9. Interaction energies, total relative energies and imaginary wavenumbers for methane monolayer structures.

Structure	Imaginary wavenumbers [cm^{-1}]	Total relative energies [kJ/mol]	Energy per molecule [kJ/mol]		
			Interaction energy	Interaction energy surface - layer	Lateral interaction
1 (ROT)	-	0.00	-14.79	-9.97	-4.82
2	-	1.51	-14.41	-9.74	-4.67
3	-	1.57	-14.40	-9.77	-4.63
4 (PAR)	-	3.10	-14.01	-9.49	-4.53
5 (WIRE-PAR)	-	4.57	-13.65	-9.19	-4.46
6	8	5.12	-13.51	-9.04	-4.47
7	12	5.37	-13.45	-9.07	-4.38
8	20	6.00	-13.29	-8.86	-4.43

The structures were not optimized using Normal Mode optimization, since on average it produced structures with more imaginary frequencies than starting one. Therefore, structure 1 (the ROT structure) has a smaller C-Mg distances then 4xCH_4 structure (Table 14 – also the ROT structure), which was further optimized using NMopt (the resulting structure has lower energy of 0.11 kJ/mol).

The structure number 4 (PAR) corresponds to a single molecule in the smaller unit cell (Section 5.7.1 on page 65). The interaction energy of the smaller model (14.02 kJ/mol, Table 4) and the larger one (14.01 kJ/mol, Table 9) is almost the same.

The C-Mg distances increase with increasing energy (Table 10). A significant jump in Mg-C distances occurs for structures number 6-8, with the WIRE-PAR structure being somewhere in between.

Table 10. Mg – C distances in pm, and wavenumbers of imaginary frequencies, for methane monolayer structures.

Structure	Mg – C distances [pm]				Average	wavenumber imaginary frequencies
1 (ROT)	303	303	303	304	303	-
2	305	306	306	307	306	-
3	305	305	305	305	305	-
4 (PAR)	304	304	305	305	305	-
5 (WIRE-PAR)	307	308	308	308	308	-
6	311	313	313	314	313	8
7	308	311	311	314	311	12
8	312	314	314	323	315	20

Table 11. Thermodynamic adsorption data kJ/mol, at $\mathbf{T = 47\ K}$ and $\mathbf{p = 0.1 MPa}$ obtained for methane monolayer structures, $\mathbf{848 \times 848 \times 2500\ pm^3}$ supercell.

#Structure	1	2	3	4	5	6	7	8
ΔE_{el}	-14.79	-14.41	-14.40	-14.01	-13.65	-13.51	-13.45	-13.29
ΔE_{ZPV}	3.86	3.59	3.61	3.55	4.01	3.83	3.76	3.55
ΔU	-11.58	-11.45	-11.39	-11.10	-10.37	-10.42	-10.40	-10.43
ΔG	-5.08	-4.99	-5.08	-4.63	-3.73	-3.77	-3.82	-3.95
ΔH	-11.97	-11.84	-11.78	-11.49	-10.76	-10.81	-10.79	-10.82
$-T\Delta S$	6.89	6.85	6.70	6.86	7.04	7.05	6.97	6.87
$\Delta U - \Delta E_{\mathrm{el}} - \Delta E_{\mathrm{ZPV}}$	-0.65	-0.63	-0.60	-0.64	-0.73	-0.74	-0.71	-0.69
$\Delta U - \Delta E_{\mathrm{el}}$	3.21	2.96	3.01	2.91	3.28	3.09	3.05	2.86
$\Delta H - \Delta E_{\mathrm{el}}$	2.82	2.57	2.62	2.52	2.89	2.70	2.66	2.47
$\Delta G - \Delta E_{\mathrm{el}}$	9.71	9.43	9.32	9.39	9.92	9.74	9.62	9.34

5.7.3. Higher Coverage

Another step of the investigation was to increase the loading of the methane molecules at the MgO(001) surface by adding one or two methane molecule (five or six in total). Since the result for five methane molecules at the MgO(001) surface resembled the ROT structure with an additional methane molecule on top of the first layer, we decided to follow that lead and investigated the formation of a second layer by putting additional one, two, and four methane molecules on top of the ROT structure. This procedure reduced the optimization problem since only one, two, or four molecules were allowed to optimize on top of the ROT layer of methane, that, in turn, lead to fast convergence of the calculations (lower degrees of freedom). This is similar to Drommund‘s approach for his investigation of second, and also third layer, of methane at the MgO(001) surface.[173-175] Results for two methane molecules on top of the ROT structure were compared with results for six methane molecules at the MgO(001) surface (the same amount of methane molecules in total). The structure of the first layer differed in both cases. In case of six methane molecules at the MgO(001) surface it was determined to be the WIRE structure, even though this structure had lower interaction energy per molecule then the global minimum obtained for two methane molecules at the ROT structure. This result puts in question GA optimization ability to locate the global minimum in a robust way.

5.7.3.1. Methane at Bare Surface

The first step on investigating higher coverage of methane molecules was performed by introducing additional rigid bodies, and allowing all of them to freely optimize during GA calculations. Of course, this increased number of possible minima, and, therefore, computational load needed to search through the optimization space.

Firstly, coverage of five methane molecules, which corresponds to $\theta = 1.25$ coverage, was examined. The results (Figure 15) lead to conclusion that the fifth molecule starts to form second layer on top of the ROT structure (Figure 7); this, in turn, motivated the investigation of the second layer of the methane by performing geometry optimization with the monolayer ROT structure as the “surface” (described in Chapter 5.7.3.2 at page 86).

The structure with six methane molecules (which corresponds to $\theta = 1.5$ coverage) was investigated next. Contrary to expectation, the global minimum for this structure was not the ROT structure with two methane molecules forming the second layer on top of the first one. The first layer of the global structure for this system (Figure 16) is in the WIRE-type

configuration, similar to the structure number 6 form the monolayer results (Figure 12). Further investigation of the second layer at the ROT structure revealed that, in fact, was not the global minimum for this loading.

Interestingly, all methane molecules forming the second layer, for both structures, were in the monopod configuration. This turned out to be true also for the complete second layer, obtained later in the course of this study. In contrast, Drummond *et al.*[174] predicted that the second layer will also consist of methane molecules in the dipod configuration.

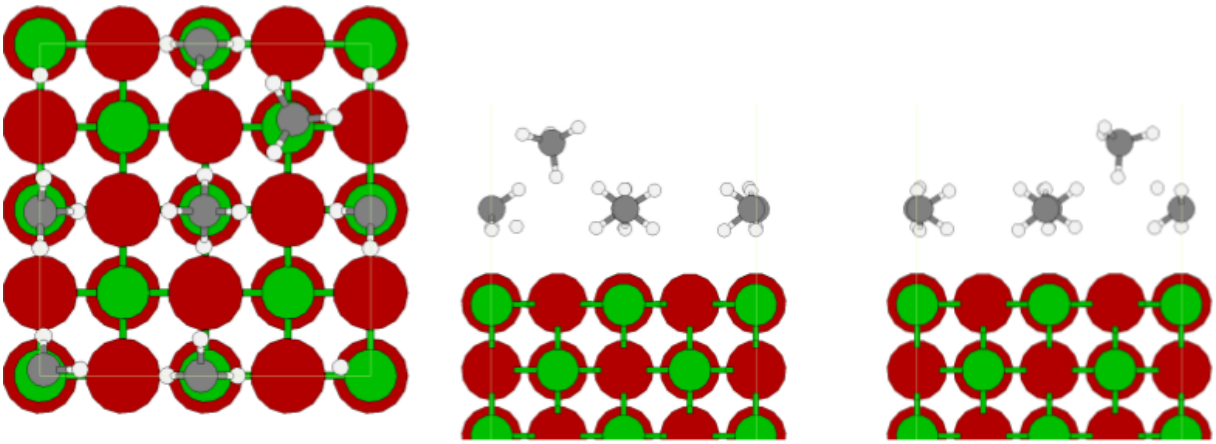

Figure 15. Global minimum for structure five methane molecules the MgO(001) surface. All molecules were allowed to optimize.

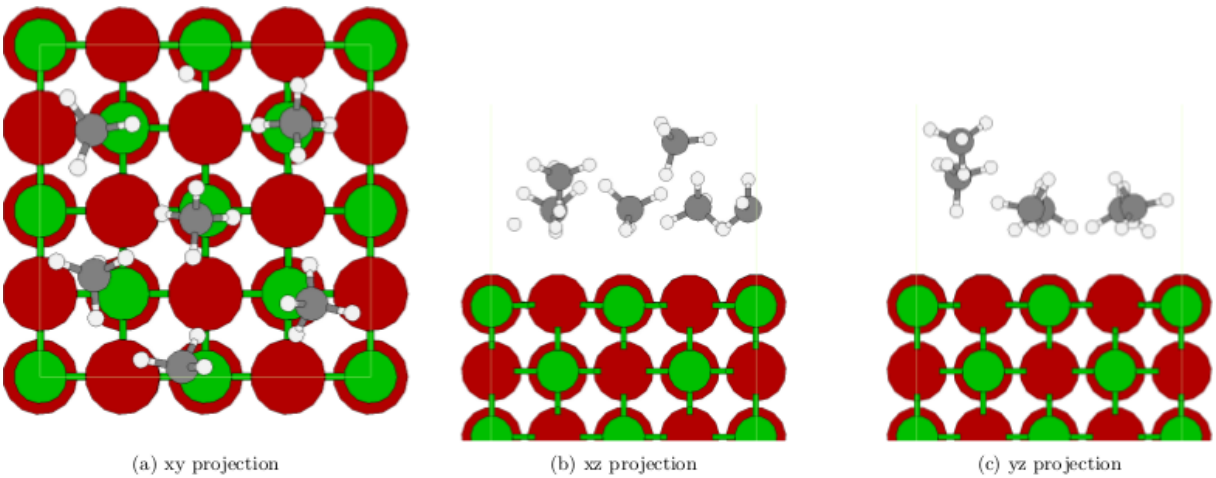

Figure 16. Global minimum for structure six methane molecules the MgO(001) surface. All molecules were allowed to optimize.

The structure with five molecules (Figure 15) is identical to the ROT structure with an additional methane molecule in the monopod configuration (one hydrogen pointing in the

direction of magnesium site) forming a second layer. The structure with six molecules (Figure 16) is different. Both "additional" methane molecules point with hydrogen atoms towards magnesium sites. One methane molecule is significantly lower than the second one.

The higher coverage systems were noticeably more difficult to re-optimize after the GA step. During the GA mating process, in some cases more than four methane molecules were placed in the first methane layer. This was possible due to fulfillment of the criterion that atoms are not too close to each other. Unfortunately, these structures, when optimized, yielded stationary points with more than one imaginary frequency (*i.e.* higher stationary points at the PES).

During re-optimization step these additional molecules (5^{th} and 6^{th}) present in the first layer, were pushed from it to form the second layer, on top of the first one. Nevertheless, the ejection of additional methane molecules from the first layer increase significantly to the computational load. Additionally, despite the fact that both structures were subject to the NMopt optimization, it was impossible to eliminate all imaginary frequencies (Table 12).

Most of the previous literature claims the dipod structure to be the preferable configuration for a single molecule, therefore, we additionally included energies for dipod configuration in Table 12, although results for the monopod configuration on-top of oxygen site yield lower values (Table 4).

Both structures, with five and six methane molecules, have two imaginary wavenumbers that correspond to translations (Table 12). Interaction energy per molecule increases with the number of molecules. This trend continues up to four molecules, when it starts to drop. Lateral interaction increases steadily with the number of molecules. This is expected, since lateral interaction is proportional to the coverage. Simultaneously, by increasing the number of molecules the surface-layer interaction contributes less to the interaction energy. This is also awaited, as the second monolayer starts to grow, the molecules within this layer interact less with the surface. They are shielded by the first layer, which has the strongest interaction with the surface.

Table 12. Interaction energies for different loading of methane molecules on the MgO(001) surface, all energies in kJ/mol per molecule, values of imaginary wavenumbers in $\boldsymbol{cm^{-1}}$. All unit cells are identical ($\mathbf{848 \times 848 \times 2500\ pm^3}$).

# molecules	1*	1**	2	3	4	5	6
Coverage	0.25	0.25	0.50	0.75	1.0	1.25	1.50
interaction energy	-12.61	-16.76	-12.82	-13.58	-14.82	-13.96	-13.03
interaction surface – layer	-12.82	-16.78	-10.68	-10.35	-9.97	-8.15	-6.45
lateral interaction	0.21	0.02	-2.15	-3.22	-4.85	-5.81	-6.58
Imaginary wavenumbers $[\boldsymbol{cm^{-1}}]$	0	23; 51	0	0	0	34	9; 16

*dipod configuration (at Mg^{2+} site)

** monopod configuration (at O^{2-} site)

Single molecule in the monopod configuration has the strongest interaction energy – 16.76 kJ/mol – that is almost 2 kJ/mol higher than interaction energy of a molecule in the ROT configuration (14.82 kJ/mol). This result suggests that for low coverage ($\theta = 0.25$) all molecules should assume monopod configuration. For two methane molecules the drop in the interaction energy is significant, almost 4 kJ/mol. For this coverage, both molecules are adsorbed at the adjacent magnesium sites in dipod configuration, and rotated by 90 degrees with respect to each other (like in the ROT configuration).

The interaction energy increases with the number molecules, up to the monolayer ROT structure that has the strongest interaction energy. For coverages higher than monolayer coverage, the positive gain in attractive interaction between the molecules is counterbalanced by the decrease in the surface-layer interaction. In result, the total interaction energies for structures with higher coverage are lower than the interaction energy for the monolayer.

Comparing structures for five and six molecules loadings of methane molecules, some questions arise. Switching WIRE structure, the first layer for six-methane structure, to one with lower energy (*i.e.* the structure WIRE PAR from Figure 11), may produce structures with

stronger interaction energy per molecule. Additionally, further study revealed (Section 5.7.3.2, on page 86) that using the monolayer ROT structure as the first layer, and introducing additional methane molecules on top of it, yields structures with lower energies (stronger interaction energy per molecule). The fact that the GA optimization scheme was not able to localize this structure as the global minimum put in question its robustness, *i.e.* whether the GA can properly localize the global minimum.

Table 13. Interaction energies obtained using embedded method (QMPOT) and CRYSTAL 09.

Method	Result [kJ/mol]	Distance Mg-C [pm]
CCSD(T):MP2//MP2:PBE+D (monolayer)[179]	-15.7±0.7	310
CCSD(T):MP2//MP2:PBE+D (single molecule)[179]	-13.0±0.6	-
CRYSTAL 09 (QZVP)[180]	14.52	307

Interaction energy for the ROT structure (Table 12) is very close to the result acquired for periodic calculations with orbital basis set with CRYSTAL 09 (Table 13). The model in this study corresponded to the ROT structure, although, it consisted of three MgO layer and two methane layers on the opposite surfaces of the slab. The difference (0.30 kJ/mol) in results between VASP and CRYSTAL 09 can be associated with dissimilarity in the models. The CCSD(T):MP2//MP2:PBE+D calculations – embedded CCSD(T):MP2 calculations at the structure optimized at the MP2:PBE+D level – gives higher desorption energies by almost 1 kJ/mol with respect to TPD and periodic PBE+D,. This energy is also near to the estimated CCSD(T):MP2//MP2:PBE+D values for a single molecule in the dipod configuration (Table 13). However, results obtained in this study show that a single molecule in the monopod configuration has stronger interaction energy.

Both structures, with five and six methane molecules, have two imaginary wavenumbers that correspond to translations (Table 12). Interaction energy per molecule increases with the number of molecules. This trend continues up to four molecules, when it starts to drop. Lateral interaction increases steadily with the number of molecules. This is expected, since lateral interaction is proportional to the coverage. Simultaneously, by increasing the number of molecules the surface-layer interaction contributes less to the interaction energy. This is also

awaited, as the second monolayer starts to grow, the molecules within this layer interact less with the surface. They are shielded by the first layer, which has the strongest interaction with the surface.

Table 14. Distances between carbon atoms and magnesium sites for different loading of methane molecules on the MgO(001) surface, all distances in pm. Additionally for methane molecules in the second layer (monopod configurations) distances between hydrogen pointing towards the surface and magnesium sites are given. For structures with five and six methane molecules ($5xCH_4$ and $6xCH_4$, respectively) hydrogen – magnesium distances (H-Mg) for the second layer molecules are reported (all in the monopod configuration, distance taken to the hydrogen from the C-H bond pointing towards the magnesium site).

Structure	Distances [pm]							
	C-Mg						H-Mg	
$1xCH_4$*	302							
$1xCH_4$**	359							
$2xCH_4$	302	302						
$3xCH_4$	307	308	308					
$4xCH_4$	308	308	308	308				
$5xCH_4$	307	308	308	308	541		432	
$6xCH_4$	313	315	318	316	494	545	389	448

*dipod configuration (at Mg^{2+} site)

** monopod configuration (at O^{2-} site)

The monopod configuration at the oxygen site was only obtained for $\theta = 0.25$ coverage. The additional GA investigation (not included in this work) for two ($\theta = 0.5$) and three ($\theta = 0.75$) methane molecules showed as global minima structures with the dipod configuration only. That may suggest that for coverage higher than $\theta = 0.25$ attractive interaction between molecules may cause them to form islands on the surface, forbidding the monopod configuration from being observed.

Table 15. Number of imaginary wavenumbers and their values [cm^{-1}] for different loading of methane molecules on the MgO(001) surface, calculated with two and four point central differences.

Number of Single Points	# molecules	Methane					
		1*	2	3	4	5	6
2	Number of Imaginary wavenumbers	0	0	0	0	1	2
	Imaginary wavenumber	-	-	-	-	34	9; 16
4	Number of Imaginary wavenumbers	1	1	1	0	3	2
	Imaginary wavenumber	44	32	27	-	3;15; 21	10; 25

*dipod configuration (at Mg^{2+} site), larger unit cell

Table 16. Thermodynamic functions in kJ/mol, for different loading of methane molecules on the MgO(001) surface, $\boldsymbol{T = 47\mathrm{K}}$, $\boldsymbol{p = 0.1\mathrm{MPa}}$, $\boldsymbol{848 \times 848 \times 2500\ \mathrm{pm}^3}$ supercell.

# molecules	1*	1**	2	3	4	5	6
ΔE_{el}	-12.61	-16.76	-12.82	-13.58	-14.82	-13.96	-13.03
ΔE_{ZPV}	3.70	2.24	3.88	3.40	3.56	3.36	3.34
ΔU	-9.75	-15.28	-9.55	-10.73	-11.85	-11.19	-10.35
ΔG	-2.93	-8.64	-3.26	-4.62	-5.47	-4.82	-3.83
ΔH	-10.14	-15.67	-9.94	-11.12	-12.24	-11.58	-10.74
$-T\Delta S$	7.21	7.03	6.69	6.50	6.77	6.76	6.91
$\Delta U - \Delta E_{\mathrm{el}} - \Delta E_{\mathrm{ZPV}}$	-0.84	-0.76	-0.61	-0.55	-0.59	-0.59	-0.66
$\Delta U - \Delta E_{\mathrm{el}}$	2.86	1.48	3.27	2.85	2.96	2.77	2.67
$\Delta H - \Delta E_{\mathrm{el}}$	2.47	1.09	2.88	2.46	2.57	2.38	2.28
$\Delta G - \Delta E_{\mathrm{el}}$	9.68	8.12	9.57	8.96	9.35	9.14	9.19

*dipod configuration (at Mg^{2+} site)

** monopod configuration (at O^{2-} site)

The C-Mg distances for the first layer of the $6xCH_4$ structure are significantly longer. The difference is due to the change in the structure of the first layer in presents of the second layer molecules. In comparisons to previous structures, the layer is significantly higher above the surface, and corresponds to the structure number six for monolayer coverage (Figure 12 on page 74). The C-Mg distances of these two structures are relatively close to each other (Table 10 on page 77).

Imaginary wavenumber for structure with five methane molecules (Table 15), 34 cm^{-1}, corresponds to a twisting motion. Two imaginary wavenumbers for structure with six methane molecules, 9 cm^{-1} and 16 cm^{-1}, are rocking and twisting motions of two methane molecules belonging to the second layer.

5.7.3.2. Methane at the ROT Structure

Motivated by results obtained for five methane molecules at MgO(001), we decided to investigate formation of the second methane layer at top of the ROT structure, assuming this would lead to global minima for five, six, and eight methane molecules structures. In this case the ROT structure was treated as a "surface", above which one, two, and four methane molecules were randomly initialize, it the initialization phase of the GA, and then optimized during the optimization phase. Because there are only one, two, and four rigid bodies, the optimization space was further reduced. This lead to decrease in computational cost and, in turn, allowed the investigation of the full second layer. Additionally, since rigid bodies are not allowed to penetrate the surface, this prevented the GA optimizer from including more molecules into the first layer via the crossover method. This problem caused huge difficulties at the full optimization step, for structures with five and six ethane molecules. Of course, at the full relaxation step, the first layer was also optimized.

By increasing the number of molecules, formation of the second layer is observed. All additional methane molecules assume positions above magnesium sites that were not occupied by the first layer molecules. All second layer methanemolecules are in the monopod configuration.

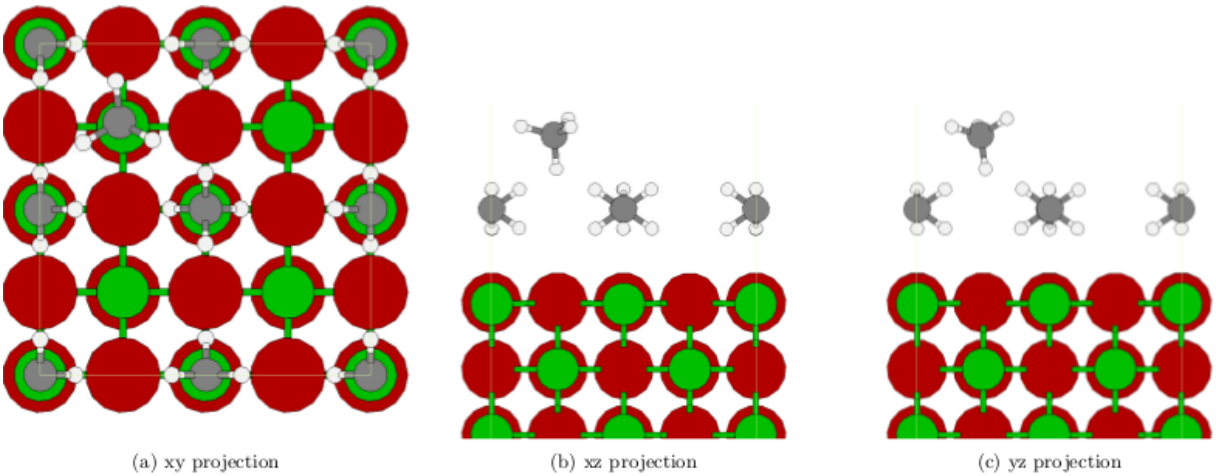

Figure 17. Global minimum for one methane molecule at the ROT structure (coverage $\boldsymbol{\theta = 1.25}$). Throughout the GA step only one molecule in the second layer was allowed to freely optimize.

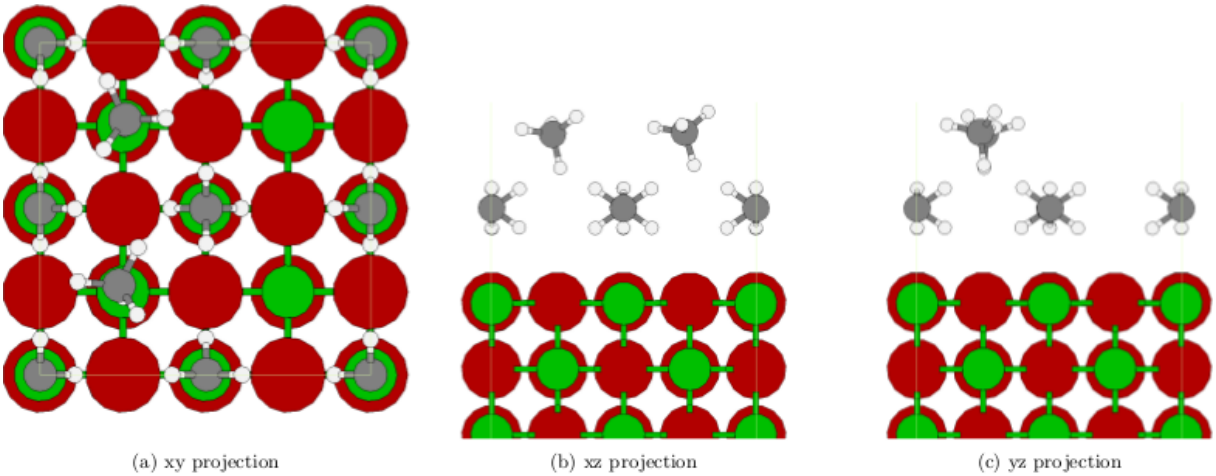

Figure 18. Global minimum for two methane molecules at the ROT structure (coverage $\boldsymbol{\theta = 1.50}$). Throughout the GA step only two molecules in the second layer were allowed to freely optimize.

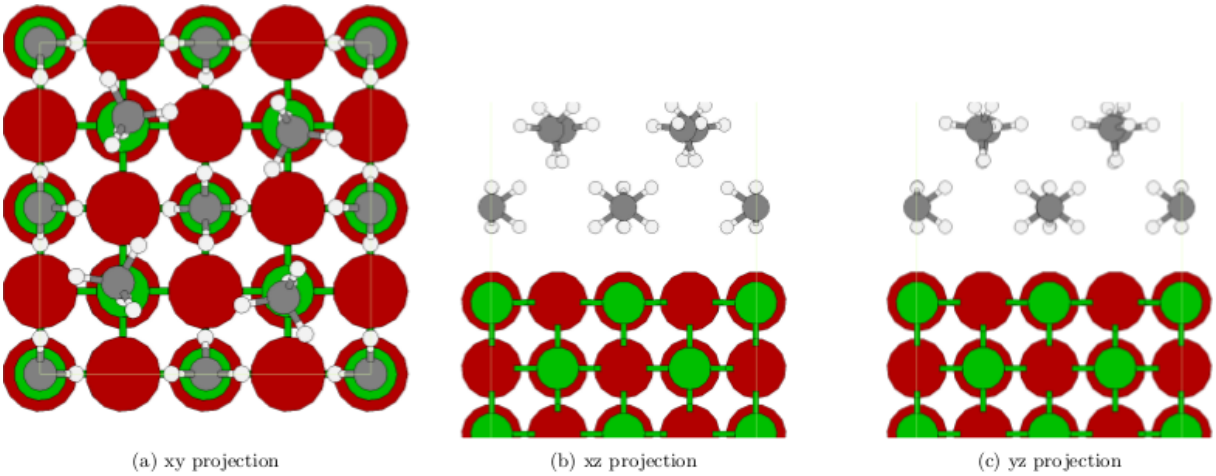

Figure 19. Global minimum for four methane molecules at the ROT structure (coverage $\boldsymbol{\theta = 2.00}$). Throughout the GA step only four molecules in the second layer ware allowed to freely optimize.

The second layer formation was also investigated by Drummond *et al.*[174]. By altering possible arrangements of the molecules in the first and second layer, they determined that the bilayer structure contain the ROT configuration as the first layer, and "rotated" configuration as the second layer. The "rotated" configuration corresponds to the ROT structure, but with all molecules rotated in the same direction by 45 degrees around normal to the surface. Molecules of the second layer adsorbed at the vacant surface magnesium sites. The global minimum structure obtained by the GA investigation, was also found in this study, but it lies 18.78 meV (18.12 kJ/mol) higher in energy than the Staggered-Rotated minimum. The Staggered-Rotated

minimum corresponds structure with first layer in the Staggered configuration (*i.e.* the ROT configuration) and second layer in the Rotated configuration (*i.e.* the ROT configuration with molecules rotated by 45 degrees with respect to the normal axis). The layers are translated with respect to each other, so that all magnesium sites are occupied.

Table 17. Comparison of interaction energies for different loading of methane molecules on the MgO(001) surface, all energies in kJ/mol. Labels "1 at ROT", "2 at ROT", and "4 at ROT" correspond to one, two, and four methane molecules at the ROT structure, respectively.

	Methane					
Structure	$4xCH_4$	$5xCH_4$	$6xCH_4$	1 at ROT	2 at ROT	4 at ROT
# molecules	4	5	6	5	6	8
interaction energy	-14.82	-13.96	-13.03	-13.93	-13.89	-14.37
interaction surface – layer	-9.97	-8.15	-6.45	-8.17	-6.97	-5.43
lateral interaction	-4.85	-5.81	-6.58	-5.76	-6.92	-8.94
Imaginary frequencies	0	1	2	1	3	4

The energy difference between $5xCH_4$ and "1 at ROT" structures (Table 17) is negligible, the structures are essentially the same, difference comes from carbon distances to the surface (compare with Table 19). Similarly, the imaginary frequencies for both structures are similar (Table 18).

By comparing the fully optimized layer structure of six methane molecules ($6xCH_4$) with only second layer optimized structure (2 at ROT), the latter has lower interaction energy. This shows that our hybrid RBGA algorithm was not able to localize global minimum in the case of six fully optimized molecules. The problem is caused by inability of the local optimization step to localize the minimum properly. The crossing-over routine generates structures with five and six methane molecules in the first layer. These structures correspond to higher order stationary points that are difficult to optimize to local minima.

Interaction energies for the monolayer ($4xCH_4$) and bilayer (4 at ROT) are very similar. The difference is only 0.45 kJ/mol per molecule. However the second layer has no interaction with the surface. The surface – layer interaction for the bilayer structure is almost a half of interaction for the monolayer structure.

Table 18. Number of imaginary wavenumbers and their values [cm^{-1}] for different loading of methane molecules on the MgO(001) surface, calculated with two and four point central differences.

Number of Single Points	# molecules	Methane					
		$4xCH_4$	$5xCH_4$	$6xCH_4$	1 at ROT	2 at ROT	4 at ROT
2	Number of Imaginary wavenumbers	0	1	2	1	3	4
	Imaginary wavenumber [cm^{-1}]	-	34	9 16	43	7 15 39	5 18 23 36
4	Number of Imaginary wavenumbers	0	3	2	3	4	5
	Imaginary wavenumber [cm^{-1}]	-	3 15 21	10 25	21 30 41	12 28 36 41	7 11 20 32 34

*dipod configuration, larger unit cell

Number of imaginary wavenumbers rises with the number of molecules on the surface, both for bare MgO calculations and calculations at the ROT structure. The Normal Mode optimization was not able to eliminate all imaginary modes. The high number of imaginary wavenumbers is most probably caused by strongly anharmonic character of the PES. This increases with the number of single point used to evaluate frequencies. All imaginary

wavenumbers, for each of the structure, correspond to rocking and twisting motions of the molecules in the second methane layer.

Distance between the magnesium site and carbon atom of the second layer molecule increases with increasing number of molecules (Table 19). In addition, distances from the surface of the first layer increase slightly for the structure with full double layer coverage (from 304 to 306 pm) In case of the $6xCH_4$ structure, the carbon-magnesium spacing is significantly different. This, as it was previously discussed, can be related to the substantial change of the first layer structure.

Table 19. Distances between magnesium and carbon of methyl group adsorbed at magnesium side, and oxygen and hydrogen of the second methyl group (the one pointing to the surface). The presented hydrogen – magnesium distances (H-Mg) reported only for the second layer molecules (all monopod configurations, the distance between the hydrogen from the C-H bond pointing towards the magnesium site).

Structure	Layer	Distances [pm]							
		C-Mg				H-Mg			
4xCH4	1st	308	308	308	308				
5xCH4	1st	307	308	308	308	432			
	2nd	541							
6xCH4	1st	313	315	318	316	389	448		
	2nd	494	545						
1 at ROT	1st	304	304	304	304				
	2nd	543				434			
2 at ROT	1st	304	304	305	304				
	2nd	550	543			445	436		
4 at ROT	1st	306	306	306	306				
	2nd	558	557	560	561	455	454	459	459

As already mentioned, the $5xCH_4$ and "1 on ROT" structures are essentially the same. However, after including ZPVE and thermal energies the difference is slightly bigger (Table 20).

The interaction energy of the "2 at ROT" structure is higher than for the 6xCH_4 structure. The "2 at ROT" global minimum in not present in the population for the 6xCH_4 structure. GA was not able to localize this structure. The problem may arise from the fact that more than four methane molecules are being pushed into the first layer during the crossover step of the GA. This generates higher order stationary points. The local optimizer was unable to full relax these strucutres. Consequently, the obtained minima were nowhere close to the stationary points. This may explain why GA failed to localize the global minimum with six methane molecules. On this account, the failure calls for more robust local optimizer. Alternative solution would be a better crossover mechanism that prevents the algorithm from placing too many molecules in the first layer.

Table 20. Thermodynamic functions in kJ/mol per molecule, for different loading of methane molecules on the MgO(001) surface, $\boldsymbol{T = 47\mathrm{K}}$, $\boldsymbol{p = 0.1\mathrm{MPa}}$, $\boldsymbol{848 \times 848 \times 2500\ \mathrm{pm}^3}$ supercell.

# molecules	4xCH_4	5xCH_4	6xCH_4	1 at ROT	2 at ROT	4 at ROT
ΔE_{el}	-14.82	-13.96	-13.03	-13.93	-13.89	-14.37
ΔE_{ZPV}	3.56	3.36	3.34	3.48	3.26	3.02
ΔU	-11.85	-11.19	-10.35	-11.03	-11.25	-11.94
ΔG	-5.47	-4.82	-3.83	-4.72	-4.85	-5.57
ΔH	-12.24	-11.58	-10.74	-11.42	-11.64	-12.33
$-T\Delta S$	6.77	6.76	6.91	6.69	6.79	6.76
$\Delta U - \Delta E_{\mathrm{el}} - \Delta E_{\mathrm{ZPV}}$	-0.59	-0.59	-0.66	-0.57	-0.62	-0.59
$\Delta U - \Delta E_{\mathrm{el}}$	2.96	2.77	2.67	2.90	2.64	2.43
$\Delta H - \Delta E_{\mathrm{el}}$	2.57	2.38	2.28	2.51	2.25	2.04
$\Delta G - \Delta E_{\mathrm{el}}$	9.35	9.14	9.19	9.21	9.04	8.80

The electronic interaction energy for the 4xCH_4 and "4 at ROT" structures are very similar, with 4xCH_4 0.45 kJ/mol lower in energy. However, after incorporating ZPVE and thermal contributions, the "4 at ROT" structure occurred to be lower by 0.09 kJ/mol in the total internal energy and enthalpy.

5.8. Ethane at the MgO(001) Surface

Ethane molecules, at the GA step, were modelled as two rigid bodies, two methyl groups, each optimizing separately. The goal was to gain additional freedom during optimization, through the possibility for the ethane molecules to switch between staggered and eclipsed conformations when in contact with the surface. Unfortunately, it also introduced some problems. During initialization and at the crossing over steps of the GA optimization, a lot of structures with unreasonable geometries were generated. Vast number of such cases signals necessity for introducing better way to model whole molecules with internal degrees of freedom, *i.e.* a way to notify system that two or more rigid bodies constitute a bigger molecule. In addition, while optimizing the relative position of them, the whole integrity should be preserved.

Study included three different coverage of the ethane molecules, starting from one molecule ($\theta = 0.25$ coverage), up to four molecules (monolayer coverage ($\theta = 1.0$)), without three molecules. The obtained structures (Figure 20 –Figure 22) correspond to three different coverages of ethane. For each system, ethane molecules assume the same configuration at the surface. One methyl group adsorbs at the magnesium site in the dipod configuration, the other methyl group adsorbs at the oxygen site in the monopod configuration. Overall, the ethane molecule preserves its staggered conformation. One molecule (Figure 20) adsorbs in the C_s symmetry with respect to the surface. For more molecules this symmetry is not present. Two ethane molecules adsorb at adjacent magnesium sites to maximize lateral interaction (Figure 21). For the monolayer structure, the ethane molecules create a pine-tree pattern at the surface (Figure 22).

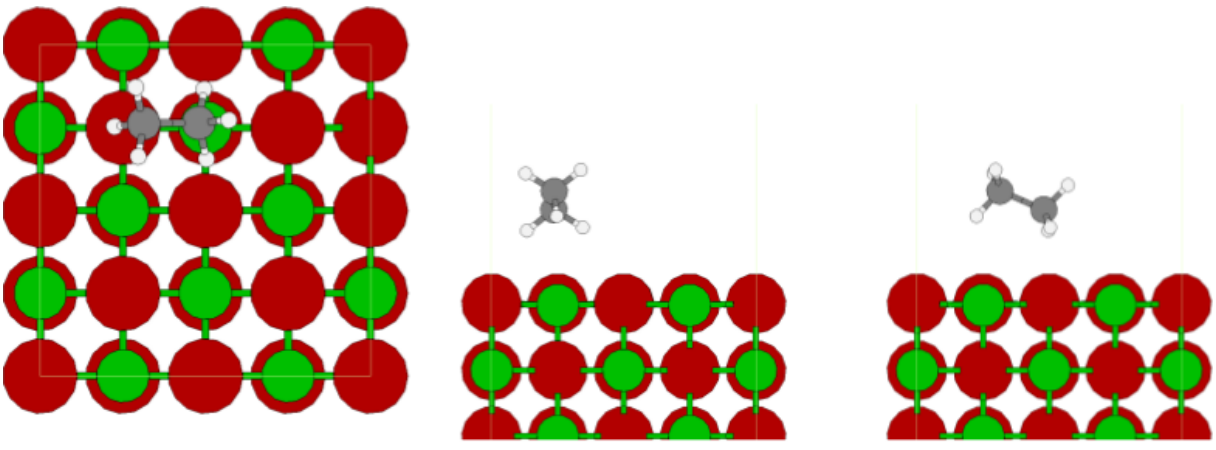

Figure 20. Global minimum for one ethane molecule at the MgO(001) surface (coverage $\boldsymbol{\theta = 0.25}$).

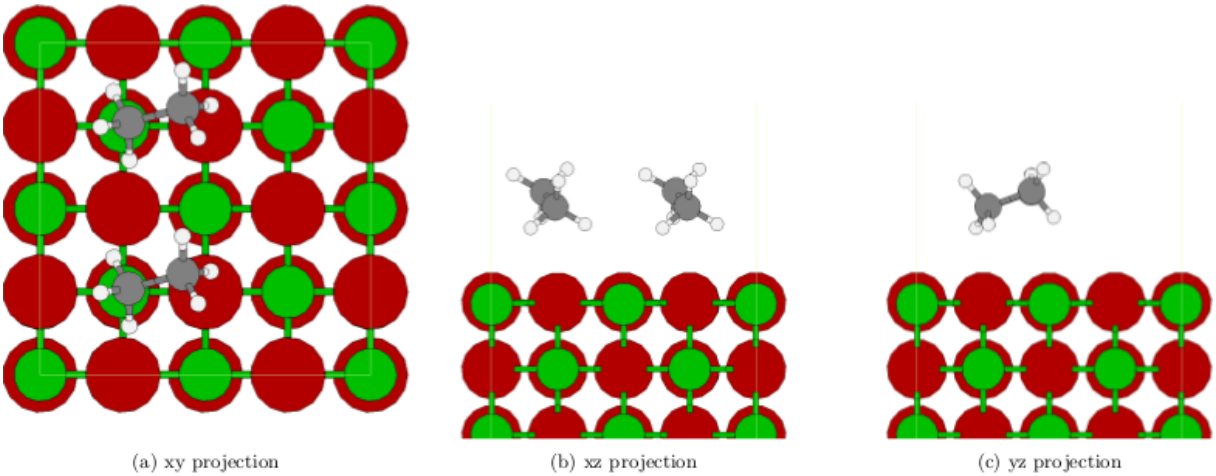

Figure 21. Global minimum for two ethane molecules at the MgO(001) surface (coverage $\boldsymbol{\theta = 0.50}$).

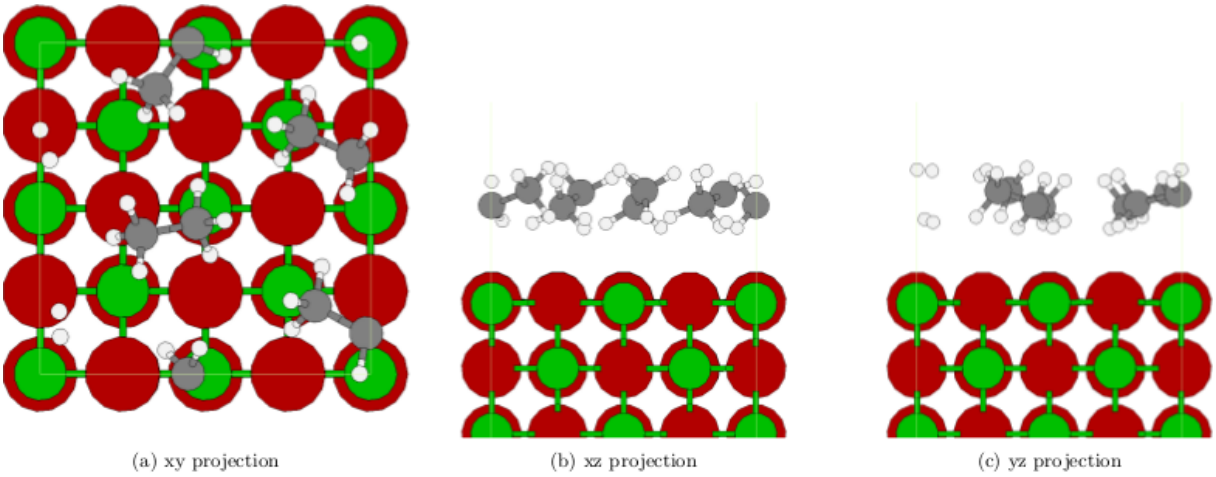

Figure 22. Global minimum for four ethane molecules at the MgO(001) surface (coverage $\boldsymbol{\theta = 1.00}$).

The interaction energies rise as the number of ethane molecules in the system increase (Table 21). At the same time surface - layer interaction decreases. This is due to enlargement in distances between the surface and molecules (Table 22). This effect is counterbalanced with increase in lateral interactions.

The interaction energy for the ethane monolayer coverage is by ~9 kJ/mol higher with comparison to the interaction energy of methane monolayer. The increase in the energy is mainly due to stronger lateral interaction, ~7 kJ/mol (4.85 kJ/mol for the methane monolayer, Table 12). Increase in the surface – layer interaction is less significant, ~2 kJ/mol (9.97 kJ/mol for the methane monolayer, Table 12).

The CCSD(T):MP2//MP2:PBE+D adsorption energy by Boese *et al.,*[179] for an isolated molecule equals to 18.9±0.8 kJ/mol. It is almost by 4 kJ/mol higher than the interaction energy obtained with the plain periodic PBE+D calculations. The same difference in the energy for these methods was observed for an isolated methane molecule.

Table 21. Interaction energies for different loading of ethane molecules at the MgO(001) surface, all energies in kJ/mol.

		Ethane		
Energies [kJ/mol]	# molecules	1	2	4
interaction energy		-15.03	-18.95	-23.83
surface – layer interaction		-15.27	-13.93	-12.30
lateral interaction		0.24	-5.03	-11.53
Imaginary frequencies		0	0	0

Table 22. Distances between magnesium and carbon of methyl group adsorbed at magnesium side, and oxygen and hydrogen of the second methyl group (the one pointing to the surface).

structure	Distances [pm]	
	C-Mg	H-O
1xC_2H_6	305	281
2xC_2H_6	309	277
	308	277
4xC_2H_6	312	282
	315	315
	309	280
	328	286

Introduction of additional ethane molecules leads to increase in distance between the surface and the molecules (Table 22). This fact is reflected in the decrease of surface – layer interaction, going down from 15.27 kJ/mol for a single molecule to 12.30 kJ/mol for four ethane molecules. The C-Mg distances for ethane monolayer are significantly longer with comparison to ones in the methane monolayer (Table 14), nevertheless the surface – layer interaction is higher (by ~2 kJ/mol) for methane structure.

The zero point vibrational energy correction (Table 23) for the monolayer structure slightly smaller than in case of methane calculations (Table 16).

Table 23. Thermodynamic functions in kJ/mol, for different loading of ethane molecules on the MgO(001) surface, $\boldsymbol{T = 75\ \mathrm{K}}$ (desorption temperature), $\boldsymbol{p = 0.1\ \mathrm{MPa}}$, $\boldsymbol{848 \times 848 \times 2500\ \mathrm{pm}^3}$ supercell.

# molecules	1	2	4
ΔE_{el}	-15.03	-18.95	-23.83
ΔE_{ZPV}	2.28	2.03	3.10
ΔU	-12.63	-16.80	-21.10
ΔG	-1.63	-5.61	-8.65
ΔH	-13.26	-17.43	-21.72
$-T\Delta S$	11.61	11.81	13.07
$\Delta U - \Delta E_{\mathrm{el}} - \Delta E_{\mathrm{ZPV}}$	0.12	0.12	-0.37
$\Delta U - \Delta E_{\mathrm{el}}$	2.40	2.15	2.73
$\Delta H - \Delta E_{\mathrm{el}}$	1.77	1.53	2.11
$\Delta G - \Delta E_{\mathrm{el}}$	13.40	13.34	15.18

5.9. Propane at the MgO(001) Surface

The last step of our study was an investigation of the layer formation for propane. Because with the present implementation there was no possibility to use different rigid bodies within one optimization, propane molecules in staggered configuration were modelled as whole rigid bodies. The adsorption was studied with four different coverages; from one molecule up to four molecules. The results obtained for three and four molecules are preliminary. More thorough investigation, involving more GA cycles, might be required.

For both coverages, with one (Figure 23) and two (Figure 24) molecules, propane is adsorbed with two methyl groups, each group in the dipod configuration. This way each molecule maximizes its interaction with the surface. For a single propane molecule, the methyl groups sit at the top of adjacent magnesium sites. In case of two propane molecules, the methyl groups of each molecule try to adsorb at the adjacent magnesium sites, but are slightly distorted. This happens most probably due to repulsion between hydrogen atoms of adjacent molecules. Each of the molecules adsorbs with one methyl group higher and the other one lower above the surface. The structure for three molecules is different (Figure 25). This global minimum has two molecules adsorbed with one methylene group, in the dipod configuration, each; and with one molecule adsorbed with one methyl group, also in the dipod configuration. All molecules are adsorbed at the further magnesium sites. The diversity of the molecule configurations is even bigger for four propane molecules global minimum (Figure 26). This minimum has one molecule adsorbed with one methyl group in monopod configuration (this CH_3 group is leaning towards adjacent magnesium site to form something similar to the dipod configuration); one adsorbed with two groups (methyl and methylene) in configuration similar to ethane configuration; finally, two molecules with one methyl group in the dipod configuration.

Investigation of a monolayer constructed by hand with four molecules in configuration similar to configuration for one and two molecules may yield better strucutre. This way interaction with the surface will be maximized, which might lead to a significant increase in interaction energy. Similarly we can ask the question, what will be the interaction energy if we take the global minimum structure for the ethane monolayer (Figure 22), and for each ethane molecule substitute one hydrogen atom (the one from the methyl group in the dipod configuration, pointing away from the surface) with another methyl group? Will the obtained monolayer structure have higher interaction energy than the one localized by the GA algorithm?

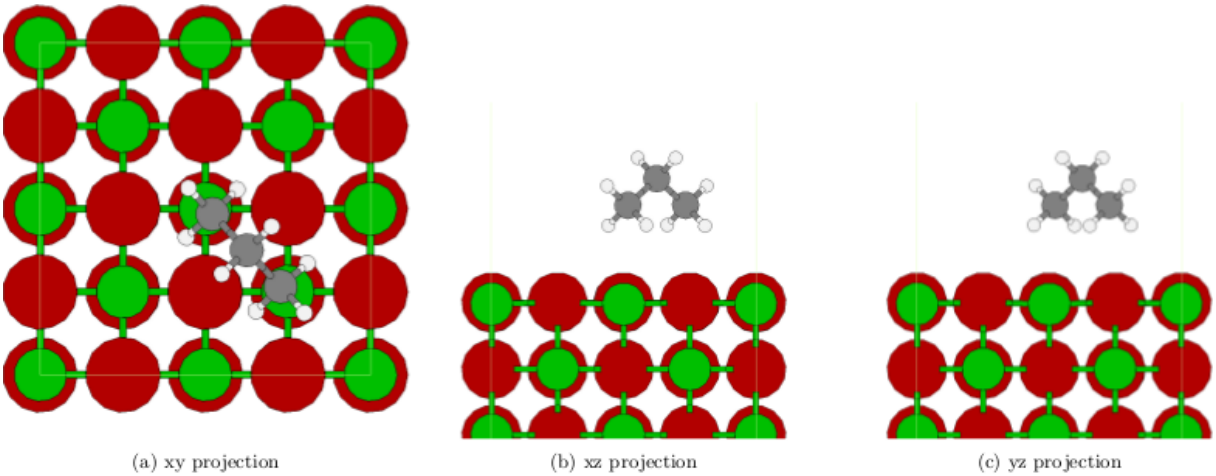

Figure 23. Global minimum for one propane molecule in staggered conformation at the MgO(001) surface (coverage $\boldsymbol{\theta = 0.25}$)

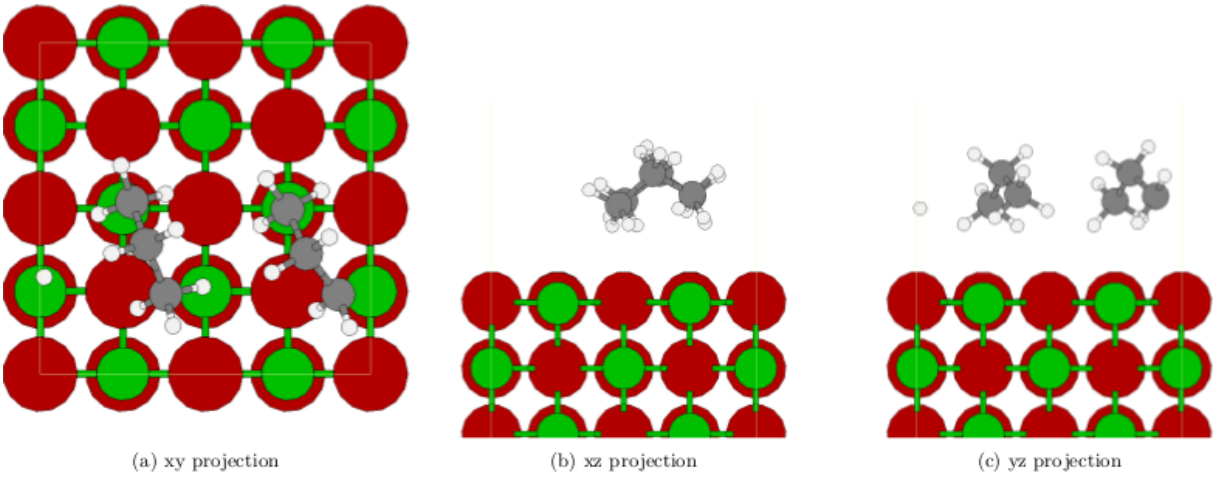

Figure 24. Global minimum for two propane molecules in staggered conformation at the MgO(001) surface (coverage $\boldsymbol{\theta = 0.50}$).

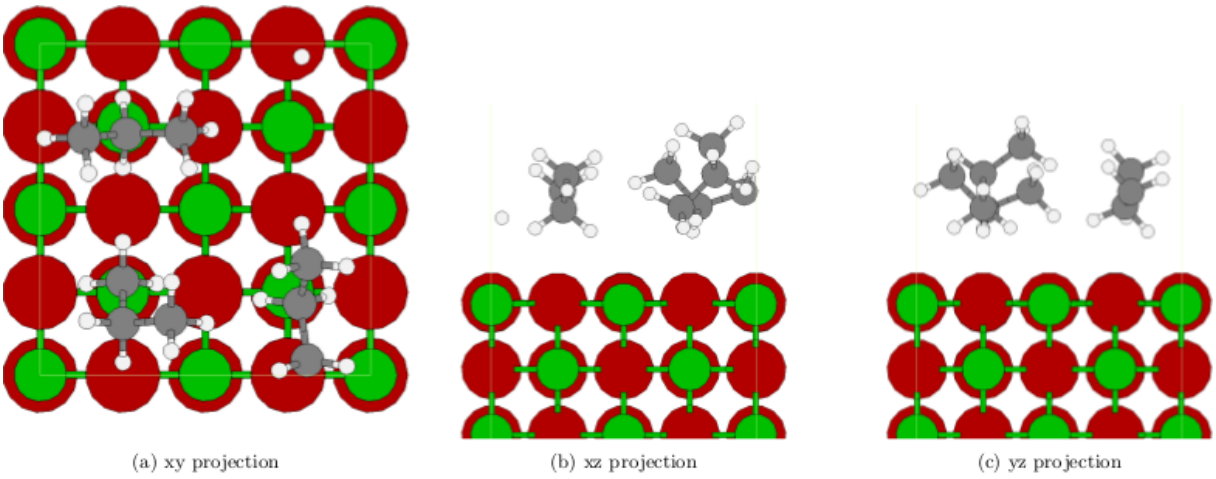

Figure 25. Global minimum for three propane molecules in staggered conformation at the MgO(001) surface (coverage $\boldsymbol{\theta = 0.75}$).

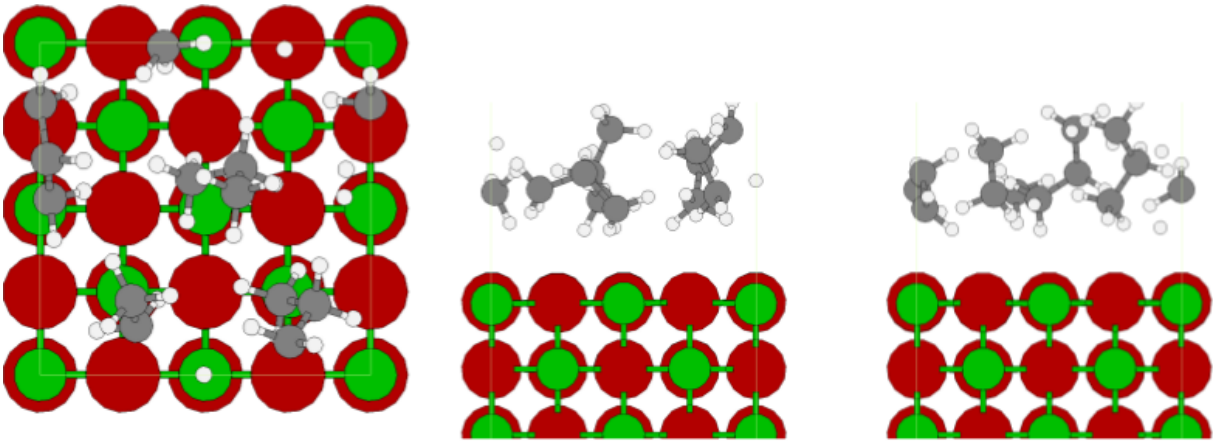

Figure 26. Global minimum for four propane molecules in staggered conformation at the MgO(001) surface (coverage $\boldsymbol{\theta = 1.00}$).

Table 24. Interaction energies per molecule for different loading of propane in staggered configuration molecules at the MgO(001) surface, all energies in kJ/mol.

	Propane staggered			
# molecules	1	2	3	4
interaction energy	-15.52	-21.75	-26.28	-27.62
interaction surface – layer	-17.60	-14.81	-14.23	-10.56
lateral interaction	2.08	-6.95	-12.05	-17.06
Imaginary frequencies	0	0	0	0

The lateral interaction for one molecule (Table 24) is positive; this is due to deformation of the propane in contact with the surface. Adding one more molecule caused lateral interaction to jump up to the value of 6.95 kJ/mol. This value is increased by about 5 kJ/mol with each additional molecule.

The energy difference between the monolayer and single molecule is substantial. The monolayer has almost double interaction energy with respect to the single molecule. At the same time the interaction energy with the surface was nearly cut in half, from 17.60 kJ/mol for

a single molecule to 10.56 kJ/mol for the monolayer. This drop in the interaction can be correlated with the carbon-hydrogen distances (Table 25). For one molecule structure there are two carbon atoms interaction with magnesium sites. The distance between them is small, 315 pm. Whereas for the monolayer structure, there is only one carbon atom interaction with the magnesium site at that small range. All other carbons atoms are considerably higher above the surface. Almost two thirds of the monolayer interaction energy comes from the lateral interaction.

There is also, of course, a question whether introducing more propane molecules can change the overall picture. Will increasing the propane coverage produce structure with higher interaction energies? One also needs to keep in mind that the final structures obtained for three and four molecules may not be global minima. Further investigation, involving more GA cycles, will be needed to confirm that these structures are indeed global minima. Also, propane was used in the staggered configuration at the GA step of the investigation. Allowing the molecule to change its conformations during GA cycles may result in structures with lower energies.

Table 25. Distances between selected atoms, all distances in pm. The C1 and C2 labels corresponds to atoms of one molecule, C1 belongs to methyl or methylene group, C2 belongs to methyl group, H-O refers to monopod configuration or ethane like configuration.

Structure	Configuration	Distances [pm]		
		C1-Mg	C2-Mg	H-O
$1xC_3H_8$	2 x $-CH_3$ dipod	315	315	
$2xC_3H_8$	2 x $-CH_3$ dipod	352	325	
	2 x $-CH_3$ dipod	363	314	
$3xC_3H_8$	$-CH_2$ dipod	315		
	$-CH_3$ dipod	309		
	$-CH_2$ dipod	302		
$4xC_3H_8$	monopod		344	248
	$-CH_3$ dipod	305		
	$-CH_3$ dipod	324		
	ethane-like	363		281

The distance of the propane molecules to the surface (Table 25) can be straightforwardly correlated to the surface-layer interaction. For the lowest coverage, with only one propane molecule, the surface-layer interaction is the greatest. This molecule interacts with both methyl groups at a very close distance to the magnesium sites. In case of two molecules structure, both molecules also interacts with two methyl groups, but this time at a larger distance to the surface's sites, hence the notable decline in surface-layer interaction. The system with three propane molecules has the same surface-layer interaction as the one with two molecules; although, in this case there are only three carbon atoms in contact with the surface, yet all of them are at the greatly smaller distance. The lowest interaction energy is observed for the highest coverage. In this case, the value of the energy drops by another 4 kJ/mol, which is due to the fact that this structure has only two propane molecules with carbon atoms substantially near magnesium adsorption sites.

The zero point vibrational energy correction (Table 26) is slightly bigger than in cases of methane (Table 16) and ethane (Table 23).

Table 26. Thermodynamic functions in kJ/mol, for different loading of propane in staggered configuration molecules at the MgO(001) surface, $\boldsymbol{T = 93\ \mathrm{K}}$ (desorption temperature), $\boldsymbol{p = 0.1\ \mathrm{MPa}}$, $\boldsymbol{848 \times 848 \times 2500\ \mathrm{pm}^3}$ supercell.

# molecules	1	2	3	4
ΔE_{el}	-15.52	-21.75	-26.28	-27.62
ΔE_{ZPV}	1.73	1.96	3.09	4.19
ΔU	-15.14	-21.23	-24.91	-25.52
ΔG	-0.43	-6.52	-9.11	-8.93
ΔH	-15.92	-22.01	-25.69	-26.29
$-T\Delta S$	15.47	15.48	16.56	17.35
$\Delta U - \Delta E_{\mathrm{el}} - \Delta E_{\mathrm{ZPV}}$	-1.35	-1.44	-1.72	-2.09
$\Delta U - \Delta E_{\mathrm{el}}$	0.38	0.52	1.37	2.10
$\Delta H - \Delta E_{\mathrm{el}}$	-0.40	-0.26	0.60	1.33
$\Delta G - \Delta E_{\mathrm{el}}$	15.09	15.24	17.17	18.69

5.10. Comparison with Experiment

The computational results were compared with coverage-dependent desorption kinetics data obtained from temperature-programmed desorption (TPD) experiments for methane, ethane, and propane on high-quality MgO(100) film (Tait *et al.*[167]). The TPD experiments were conducted under ultrahigh-vacuum conditions (base pressure$\sim 1 \times 10^{-10}$ Torr *i.e.* $\sim 1.3 \times 10^{-14}$ MPa) by ramping the temperature of the sample at given heating rates and measuring the desorption of the molecule by quadrupole mass spectrometer (QMS). For each molecule, TPD experiments were conducted at many initial coverages, and at several other heating rates for each molecule.

In each TPD spectra two distinct peaks corresponding to desorption of the multilayer (low temperature) and desorption of the first monolayer (higher temperature) were observed. Calculation of the coverage-dependent desorption energy was made for each molecule by the inversion-optimization analysis[191] to obtain the "best-fit" prefactor for desorption to simulate TPD data. Using this best-fit prefactor the desorption energy versus coverage for each of the molecules was calculated for each coverage from range of 0–0.8 ML (monolayers); and fitted with a solid lines (energy curves) representing an empirical equation that describes coverage dependence of desorption energy (E_{d})

$$E_{\mathrm{d}}(\theta) = E_0 + \gamma\theta + E_{\mathrm{def}} e^{\theta/\theta_{\mathrm{def}}} \qquad (54)$$

The first term E_0 represents the desorption energy obtained by extrapolating the linear region of the coverage-dependent energy curve back to the limit of zero coverage (*i.e. zero-limit desorption energy*), and can be interpreted as the activation energy for desorption of an isolated molecule from a MgO(100) terrace site (*i.e.*, in the absence of defect sites and adsorbate-adsorbate interactions). The factor γ is the increase in desorption energy due to lateral interactions between adsorbates (*i.e.* lateral interaction factor); E_{def} factor is related to the difference between the adsorption energy of an alkane molecule adsorbed on a MgO terrace site compared to one adsorbed at a defect site (the energy difference between E_0 and the measured desorption energy at zero coverage, $E_{\mathrm{d}}(0)$).

The θ_{def} factor corresponds to the rate at which the influence of defect sites on the energy decays with increasing coverage (θ). It is related to density (fractional area) of defect sites on the surface, and their influence on the adsorption of near-by molecules.

The interaction energies obtained from calculations cannot be directly compared with experimental desorption energies. The temperature-programmed desorption experiments[167] yield Arrhenius activation energies, E_{A}, which differ from the enthalpies at the desorption temperature, ΔH_{T}, by RT.[2]

$$\Delta H_{\mathrm{T}} = E_{\mathrm{A}} - RT \tag{55}$$

By subtracting the thermal energies, ΔH_{th}, (thermal contributions to change in enthalpy due to non-zero temperature at which the experiment was performed)

$$\Delta H_0 = \Delta H_{\mathrm{T}} - \Delta H_{\mathrm{th}} \tag{56}$$

and zero-point vibrational energy (ZPVE) one obtain desorption energy (ΔE),

$$\Delta E = \Delta H_0 - \Delta E_{\mathrm{ZPV}} \tag{57}$$

which can be then compared with calculated interaction energies for the energy minimum structures.

Table 27. Experimental Arrhenius desorption barriers, obtained from equation (54), and temperatures (Tait *et al.*[167]).

Alkane	Desorption energy ($\theta = 1$) [kJ/mol]	Desorption energy (Zero coverage limit) [kJ/mol]	Desorption temperature [K]
Methane	12.63	11.1	47
Ethane	22.76	21.3	75
Propane	29.40	28.0	93

Theoretically corrected experimental desorption energy for methane (Table 28) are very close to calculated interaction energies. For the small unit cell (Table 4), which corresponds to monolayer coverage in the PAR configuration, the interaction energy is 14.02 kJ/mol. The desorption energies from the experiment is 14.82 kJ/mol. Given the fact that the PAR structure is not the global minimum, the difference of 0.80 kJ/mol is relatively small. For the zero-

coverage limit, the discrepancy between experiment (13.29 kJ/mol) and theory (16.76 kJ/mol, monopod) is significantly bigger, and amounts to 3.47 kJ/mol. However, if we assume the isolated dipod configuration (12.63 kJ/mol) to be the global minimum, the difference decreases to only 0.66 kJ/mol. The interaction energy obtained for the methane monolayer ROT structure (14.82 kJ/mol, Table 12) matches perfectly to the theoretically corrected experimental desorption energies.

Table 28. Experimental desorption energies derived from equations (55), (56), and (57). Input values taken from table 27. Thermodynamic contributions were obtained from vibrational analysis of global minima structures corresponding to monolayer coverage. All energies are in kJ/mol. Experimental data obtained by Tait *et al.*[167]

	monolayer coverage			Zero-coverage limit		
	methane	ethane	propane	methane	ethane	propane
Arrhenius desorption	12.63	22.76	29.40	11.10	21.30	28.00
Temperature [K]	47	75	93	47	75	93
RT value	0.39	0.62	0.77	0.39	0.62	0.77
thermal energy correction	0.98	1.00	0.91	0.98	1.00	0.91
ZPVE correction	-3.56	-3.10	-4.19	-3.56	-3.10	-4.19
desorption energies	14.82	24.24	31.91	13.29	22.78	30.51

The difference between low coverage theoretical energy and zero-coverage limit extrapolated from experimental data might suggest that the extrapolation procedure fails. It is impossible to measure experimentally interaction energy of only one molecule, therefore the value have to be extrapolated from the coverage-dependent energy. Nevertheless, the extrapolation might not necessarily capture the nature of single molecule interaction. In presence of other molecules, single molecule may be prevented from assuming certain positions, due to the strong lateral interaction. The difference with respect to the zero-coverage limit estimation may be due to formation of methane islands, with methane molecules in the dipod configuration. To verify

which position of a single molecule is preferable, further calculation, *e.g.* cluster or embedded, need to be done.

The theoretical adsorption energy for the ethane monolayer (23.83 kJ/mol, Table 21) is close to the experimental adsorption energy (24.24 kJ/mol, Table 28). However, the zero-coverage limit desorption energy, 22.78 kJ/mol (Table 28), is significantly higher than theoretical value for an isolated ethane molecule, 15.03 kJ/mol (Table 21). Stronger interaction for densely packed ethane molecules suggests that for lower coverage than monolayer, ethane will form islands of interacting molecules and bare surface. Therefore, the extrapolation procedure will not predict the interaction energy of a single, isolated molecule.

The experimental desorption energy for the propane monolayer, 31.91 kJ/mol (Table 28), is relatively close to the calculated desorption energy – 27.62 kJ/mol (Table 24). Yet, the difference of 4.29 kJ/mol is significantly larger when compared with previous results for ethane and methane. In case of zero-limit coverage, the discrepancy between experimental and theoretical values is tremendous. The desorption energy, 30.51 kJ/mol (Table 28), is approximately double the theoretical value 15.52 kJ/mol (Table 24).

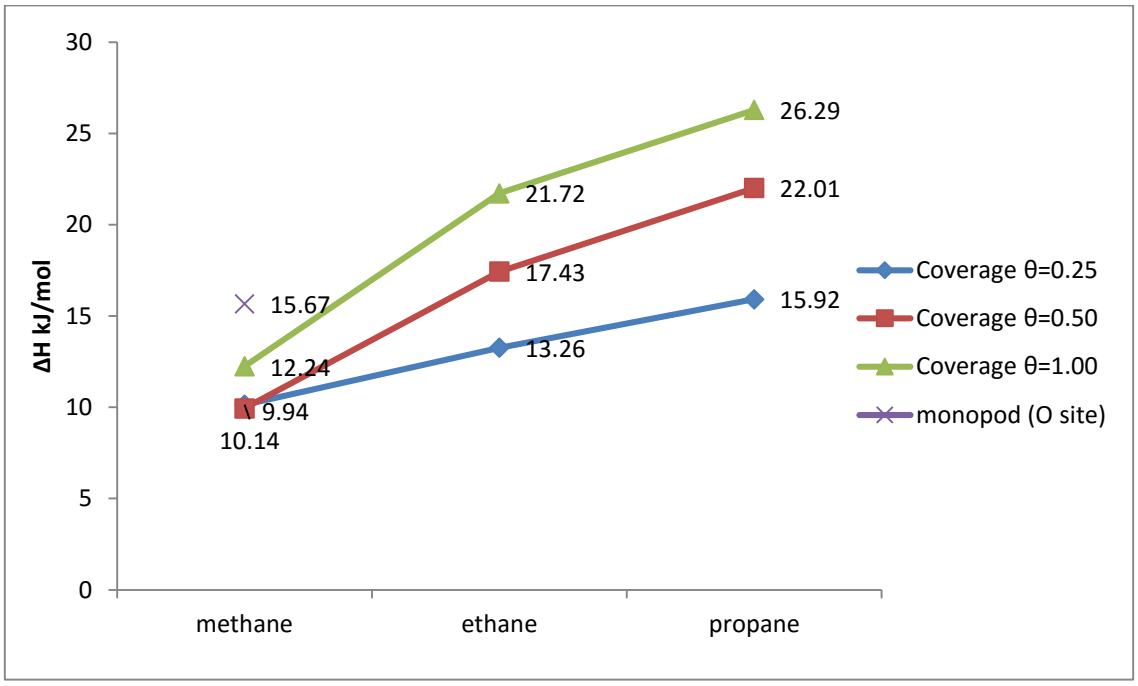

Figure 27. Change in enthalpy as a function of hydrocarbon chain length for different coverage of molecules.

Figure 27 shows the change in the adsorption enthalpy with increasing number of carbon atoms for different coverage. The enthalpy change dependence on the coverage is the smallest for methane. Exception is the single molecule coverage which has a significantly higher enthalpy change (15.67 kJ/mol, Table 16). For ethane and propane molecules loading on the surface has much more pronounced influence. Coverage shift for ethane molecules from 0.25 to 1.00 increases the adsorption enthalpy by ~8 kJ/mol (from 13.26 to 21.72 kJ/mol, Table 23). The same coverage increase for propane results in the adsorption enthalpy increase by ~10 kJ/mol (from 15.92 to 26.29 kJ/mol, Table 26). This tendency can be explained with dominating role of the lateral interaction for higher coverages (Table 29). For all three molecules, the interaction energy increases with the coverage of molecules. Difference between molecules interaction is most pronounced for the monolayer coverage. For the monolayer coverage, surface – layer interactions are comparable, but lateral interaction for ethane is 2.4 and for propane 3.5 times stronger than for methane.

Table 29. Interaction energies for methane, ethane, and propane with different coverage.

	Coverage	Methane	Ethane	Propane
Interaction energy	0.25	-16.76	-15.03	-15.52
	0.50	-12.82	-18.95	-21.75
	1.00	-14.82	-23.83	-27.62
Lateral interaction energy	0.25	0.02	0.24	2.08
	0.50	-2.15	-5.03	-6.95
	1.00	-4.85	-11.53	-17.06
Surface – layer interaction energy	0.25	-16.78	-15.27	-17.60
	0.50	-10.68	-13.93	-14.81
	1.00	-9.97	-12.30	-10.56

6. Summary

The present work was concentrated on developing the Rigid Body Genetic Algorithm (RBGA), and applying it to investigate the hydrocarbon adsorption on the MgO(001) surface. The RBGA method is a modified hybrid genetic algorithm with rigid body optimization at the local optimization step. The modification allows for a vast simplification of the optimization problem, and, in turn, to search a large number of possible configuration. The key assumption of the method is that individual parts of the system (rigid bodies) do not change their internal configuration throughout the global optimization. Therefore, this method is a perfect tool to study phenomena like adsorption, where all the subsystems – surface and individual molecules – preserve their internal structure. The algorithm allows to obtain global minima, which then can be fully optimized and to account for deformations due to the relaxation of the surface and adsorbate molecules.

The first step of this work was the design and implementation of the RBGA method. The RBGA method, as a hybrid genetic algorithm, uses local relaxation steps. This ensures that structures with high energies but within basin of attraction of a low lying minima can contribute to the population and ensuring faster convergence rate of the global search. In the local optimization step the RBGA method uses the rigid body approximation. To program the method, modifications to existing quantum mechanics programs had to be done. The rigid body optimization scheme was developed within the QMPOT code which performs the local optimization. The interface to the QMPOT program was written as a part of the DODO hybrid genetic algorithm program to enable usage of the rigid body optimization as a local relaxation step. Finally, the SURFACE program – program responsible for performing genetic operations (initialization, crossover, mutation) – was modified to work with the rigid body approximation, *i.e.* to make sure that during any of these operations the rigid bodies are kept intact.

The performance of the RBGA method was evaluated on the test system. The system consisted of four methane molecules on the magnesium oxide surface; a 2A × 2A slab model (A is twice the Mg-O distances) was employed. The concentration of the methane molecules

corresponded to the monolayer coverage. The performance was tested by running three independent RBGA optimizations. In all three runs the ROT configuration was located as the global minimum. This result is in an agreement with the previous theoretical investigations.

The outcome of the test confirmed that the RBGA is able to localize global minimum for the methane monolayer structure at the MgO(001) surface, and validated the usage of the method for further studies of the hydrocarbon adsorption.

In the second step of the study, a vast number of hydrocarbon-magnesium oxide systems, with different coverage, were investigated. Three system: methane, ethane, and propane, were of interest. Each hydrocarbon molecule was studied with multiple coverage loadings. All global minima from the RBGA calculations were fully optimized. Since a lot of obtained structures for methane systems possessed imaginary frequencies (*i.e.* they were not minima), an additional step – Normal Mode optimization – was introduced. Unfortunately, some imaginary frequencies were still present after this refinement. Lastly, the obtained desorption energies were corrected with the ZPVE and thermal energies. Final desorption energies are in a very good agreement with the experiment.

The methane adsorption at the MgO surface was found to be a very difficult case for the RBGA optimization. During local optimization, the structure of the system can change significantly; a small change in energy can correspond to huge structure rearrangements. The PES surface seems to be very shallow, with low barriers – causing the methane molecules to behave as spheres rolling on the flat surface of MgO(001).

In methane adsorption studies, each methane molecule was treated as a separate rigid body. For a single molecule adsorption two models of the surface were used, smaller and larger, corresponding to monolayer coverage and quarter of monolayer coverage, respectively. In contradiction to expectation, the isolated molecule in monopod configuration had the lowest energy, significantly lower than the one in the dipod configuration. The dipod configuration is believed to be the most favorable one. But this might be true only for higher coverage of the methane, as the experimental findings concern the adsorption of a methane film, not a single molecule. Interaction energies of the monolayer are in a perfect match with theoretically corrected experimental desorption energies (to account for ZPVE and thermal energies at the temperature of the experiment). Assuming the dipod to be the preferable configuration for the isolated molecule the match between experimental and theoretical values is also very close (within 1 kJ/mol). The same number for methane in the monopod configuration adsorbed at oxygen site is bigger (~3.50 kJ/mol).

The RBGA method had no problem to localize global minima up to five methane molecules per unit cell. For six molecules structures, the crossing over procedure yielded unreasonable structures – with five and six methane molecules in the first layer. These structures optimized to higher order stationary points, which were then difficult to optimize to minima at the full optimization step.

The global minimum obtained for five methane molecules coverage revealed a probable formation pattern of the second layer. The structure consisted of the first layer in the ROT configuration, and additional methane on top of it. This suggested a way to validate results obtained for six methane molecules. To do so, the ROT structure was used as a "surface", and two methane molecules were deployed on top of it. The studies confirmed that the RBGA did not locate the global minimum for the six methane system – the structure obtained by introducing two molecules on top of the ROT structure produced a lower energy configuration. Finally, by introducing four methane molecules on top of the ROT structure, the formation of the second methane layer was investigated.

In case of ethane adsorption, each molecule was treated as two separate rigid bodies – methyl groups. This allowed for additional conformational freedom within the ethane molecule, *i.e.* rotational freedom around the C-C bound. The drawback of this approach was that the initialization and mating steps produced structures with unreasonable geometries with very high energy (like CH_4-CH_2 adducts and CH_3-CH_3 structure with inverted one of the methyl groups). Nevertheless, the RBGA method was able to localize global minima with reasonable geometries, and the obtained monolayer structure yields an interaction energy very close to the experimental value. This comparison for the isolated molecules is ambiguous. The desorption energy, in that case, was almost half the experimental zero-coverage limit desorption energy, but on the other hand, very close to other theoretical CCSD(T):MP2//MP2:PBE+D studies. The energy difference between experiment and theory for ethane monolayer is very small (~0.5 kJ/mol). For an isolated molecule this difference is much higher (almost 8 kJ/mol).

Since the implementation of the RBGA method does not allow for optimization of different rigid bodies, the propane molecule had to be treated as one, complete rigid body. Propane in staggered configuration was used. The interaction energy for the propane monolayer is close to the experimental desorption energy, but is significantly larger than in previous cases (4.29 kJ/mol). This difference can be accounted for the fact that the results for propane are preliminary, and the global minimum may not have been localized yet. Further investigation will be required to answer that question.

Comparing molecule–surface interaction for isolated molecules, the larger the molecule, the greater is its interaction with the surface. This is most probably due to a larger surface of the molecule that can interact with the substrate. At the same time carbon–magnesium site distances elongate with growth of the hydrocarbon chain. The comparison of the interaction energies for isolated molecules suggests that the extrapolation method used to get zero-coverage limit from the experimental results may be missing some key element. In contrast, the surface-layer interaction for monolayer structures is fairly constant (between 10 and 12 kJ/mol). However, the lateral interactions are increasing with the size of molecules.

The studies showed that the RBGA method can be of great advantage in simplifying adsorption optimization problem. The code performs well, and some of the results are in very good agreement with the experiment. Yet, still some questions are unanswered, especially in case of propane adsorption. Further studies will be needed to answer them. In addition, a correction to the RBGA code can be made to increase algorithm performance and convergence. Better performance at the mating step for the methane, *i.e.* keep better distances between molecules – to avoid squeezing of molecules in the first layer, is needed. Additionally, a mechanism to keep ethane molecules as a whole, *i.e.* to introduce an "internal" degrees of freedom (the rotation around C-C bond), could help to avoid unreasonable geometries with two methyl groups apart. Finally, introducing different rigid bodies would enable to model propane molecules as two methyl groups and one methylene group.

7. Zusammenfassung

Die vorliegende Arbeit behandelt die Entwicklung des genetischen Starrkörper-Algorithmus (rigid body genetic algorithm, RGBA), und seine Anwendung zur Untersuchung der Kohlenwasserstoff-Adsorption auf der MgO (001) Oberfläche. Die RBGA Methode ist ein modifizierter hybrid-genetischer Algorithmus mit Starrkörper-Optimierung im lokalen Optimierungsschritt. Diese Modifikation führt zu einer großen Vereinfachung des Optimierungsproblems und ermöglicht damit, eine große Anzahl von möglichen Konfigurationen zu analysieren. Die zentrale Annahme der Methode ist, dass die einzelnen Teile des Systems (starrer Körper) während der gesamten globalen Optimierung nicht ihre interne Konfiguration ändern. Daher ist diese Methode ein geeignetes Werkzeug, um Phänomene wie Adsorption zu studieren, in dem alle Teilsysteme - Oberfläche und einzelne Moleküle - ihre interne Struktur bewahren. Der Algorithmus ermöglicht das Auffinden der globalen Minima für die Starrkörper, die dann im nächsten Schritt vollständig optimiert („relaxiert“) werden, um Verformungen aufgrund der Entspannung der Oberfläche und des Adsorbats auszumachen.

Der erste Schritt dieser Arbeit war die Konzeption und Umsetzung der RBGA Methode. Als hybrid-genetischen Algorithmus verwendet die RBGA Methode lokale Relaxations-schritte. Das stellt sicher, dass Strukturen mit hohen Energien, aber innerhalb Becken der Anziehung eines tief liegenden Minimums zur Bevölkerung beitragen können und führt so zu einer schnelleren Konvergenzrate der globalen Suche. In der lokalen Optimierung verwendet der RBGA die Starrkörper-Näherung. Um das Verfahren zu programmieren, mussten Änderungen der bestehenden Quantenmechanikprogramme durchgeführt werden. Das Starrkörper-Optimierungsschema wurde im Rahmen des QMPOT Codes entwickelt, welcher die lokale Optimierung ausführt. Die Schnittstelle zum QMPOT Programm ist als Teil des hybrid-henetischen Algorithmus-Programms DODO geschrieben um die Nutzung der Rigid Body Optimierung als lokale Relaxations-schritt zu ermöglichen. Schließlich ist das Programm SURFACE – ein Programm zur Durchführung genetischer Operationen (Initialisierung,

Crossover, Mutation) – modifiziert worden, um mit der Starrkörper-Näherung zu arbeiten, das heisst, um sicherzustellen, dass bei einer dieser Operationen die Starrkörper intakt bleiben.

Die Leistung des RBGA Methode wurde bei einem Testsystem untersucht. Das System bestand aus vier Methan-Molekülen auf der Magnesiumoxid-Oberfläche; ein 2A × 2A Oberflächenmodell (A ist zweimal der Mg-O-Abstand) wurde eingesetzt. Die Konzentration der Methanmoleküle entsprach einer Monolage. Die Performance wurde durch Ausführen von drei unabhängigen RBGA Optimierungen getestet. In allen drei Versuchen wurde die ROT-Konfiguration als das globale Minimum identifiziert. Dieses Ergebnis ist in Übereinstimmung mit bisherigen theoretischen Untersuchungen.

Dieses Testergebnis bestätigt, dass der RGBA in der Lage ist, das globale Minimum für die Methan Monolage auf der MgO(001) Oberfläche zu finden und hat die Anwendung dieser Methode für weitere Untersuchungen zur Kohlenwasserstoff-Asorption validiert.

Im zweiten Schritt der Untersuchung wurden eine große Anzahl von Kohlenwasserstoff-Magnesiumoxid-Systemen mit verschiedenen Beladungen untersucht. Drei Systeme: Methan, Ethan und Propan sind von Interesse. Jedes Kohlenwasserstoffmolekül wurde mit unterschiedlicher Beladung untersucht. Alle globalen Minima der RBGA Berechnungen wurden vollständig optimiert. Da viele der für Methan erhaltenen Strukturen imaginären Frequenzen aufwiesen (was bedeutet, dass sie keine Minima waren), wurde ein zusätzlicher Schritt, die Normalmoden-Optimierung, eingeführt. Leider waren einige imaginäre Frequenzen auch nach dieser Verfeinerung vorhanden. Schließlich wurden die erhaltenen Desorptionsenergien mit ZPVE und thermischen Energien korrigiert. Endgültige Desorptionsenergien sind in einer sehr guten Übereinstimmung mit dem Experiment.

Für die Adsorption von Methan in der MgO-Oberfläche wurde gefunden, dass sie ein sehr schwieriger Fall für die RGBA-Optimierung ist. Während der lokalen Optimierung kann sich die Struktur des Systems erheblich verändern; einer kleinen Änderung in der Energie können riesige Struktur Umlagerungen entsprechen. Die Potentialfläche scheint sehr flach zu sein, mit niedrigen Barrieren – was bewirkt, dass die Methanmoleküle sich verhalten wie Kugeln, die auf der flachen MgO(001) Oberfläche rollen.

In Methanadsorptions-Studien wurde jedes Methanmolekül als separater starrer Körper behandelt. Für ein Molekül Adsorption zwei Modelle der Oberfläche verwendet wurden, kleinere und größere entsprechend Abdeckung und Viertelmonoschichtbedeckung Monoschicht auf. Entgegen der Erwartungen, hatte das isolierte Molekül in Monopod-Konfiguration die niedrigste Energie, deutlich niedriger als die in der dipod Konfiguration. Von

der dipod Konfiguration wird angenommen, dass sie die energetisch günstigste sei. Aber dies könnte nur für höhere Methan-Beladungen zutreffen, da die experimentellen Befunde die Adsorption eines Methan-Films betreffen, nicht eines einzelnen Moleküls. Wechselwirkungsenergien der Monoschicht sind in einer seher guter Übereinstimmung mit theoretisch korrigierten, experimentellen Desorptionsenergien (zur Berücksichtigung von ZPVE und thermischen Energien bei der Temperatur des Experiments). Unter der Annahme, dass die dipod Konfiguration für das isolierte Molekül bevorzugt ist, ist die Übereinstimmung zwischen experimentellen und theoretischen Werten mit einer Abweichung von weniger als 1 kJ/mol auch sehr gut. Dieser Wert ist für die Monopod-Konfiguration-bei-Sauerstoff Konfiguration größer (~3,50 kJ/mol).

Die RBGA Verfahren hatte kein Problem, globale Minima bis zu fünf Methanmolekülen pro Einheitszelle zu lokalisieren. Für Strukturen mit sechs Molekülen Strukturen, ergab das Crossover-Verfahren unvernünftige Strukturen - mit fünf und sechs Methanmolekülen in der ersten Schicht. Diese Strukturen wurden zu stationären Punkten höherer Ordnung optimiert, für die dann die Optimierung zu Minima im vollständigen Optimisierungsschritt schwer war.

Das globale Minimum für die Beladung mit fünf Methanmolekülen pro Einheitszelle ergab ein wahrscheinliches Bildungsmuster der zweiten Schicht. Die Struktur besteht aus der ersten Schicht in der ROT-Konfiguration und einem zusätzlichen Methanmolekül darauf. Dies legte einen Weg nahe, um die Ergebnisse, die für sechs Methanmoleküle erhalten wurden, zu validieren. Um dies zu tun, wurde die ROT-Struktur als "Oberfläche" verwendet wird, und zwei Methanmoleküle wurden auf ihm eingesetzt. Die Studien bestätigten, dass die RBGA nicht das globale Minimum für die sechs Methan System fand - die durch die Einführung von zwei Molekülen auf der Oberseite des ROT-Struktur erhalte Struktur erzeugt eine Konfiguration tieferer Energie. Schließlich wird durch die Einführung von vier Methanmolekülen auf der Oberseite der ROT Struktur die Bildung der zweiten Methanschicht auf der ersten untersucht.

Im Falle der Ethan-Adsorption wurde jedes Molekül als zwei getrennte starre Körper – die Methylgruppen - behandelt. Diese zusätzliche konformative Freiheit inerhalb des Ethan-Molekül erlaubt zum Beispiel eine Rotation um die C-C-Bindung. Der Nachteil dieser Vorgehensweise war, dass die Initialisierung und Paarungsschritt erzeugten Strukturen mit ungünstig Geometrien und sehr hoher Energie (wie CH_4-CH_2 Addukte und CH_3-CH_3 Struktur mit eine der Methylgruppen einer anderen Richtung orientiert). Dennoch war das RBGA Verfahren in der Lage, globale Minima mit vernünftigen Geometrien zu lokalisieren, und die

erhalte Monolage weist eine Wechselwirkungsenergie sehr nah an dem experimentellen Wert auf. Dieser Vergleich sieht für das isolierte Molekül weniger gut aus. Die Desorptionsenergie war in diesem Fall fast die Hälfte der experimentellen Grenz-Desoorptionsenergie für Nullbedeckung, jedoch, nahe an anderen theoretischen CCSD(T):MP2//MP2:PBE+D-Studien. Die Energiedifferenz zwischen Experiment und Theorie ist für die Ethan-Monoschicht sehr klein (~0,5 kJ/mol). Für ein isoliertes Molekül ist dieser Unterschied wesentlich höher (fast 8 kJ/mol).

Da die Implementierung der RBGA Methode die Optimierung von freidefinierten starren Körpern nicht ermöglicht, musste das Propan-Molekül als ein ganzer starrer Körper behandelt werden. Die gestaffelte Konfiguration wurde für das Propanmolekül angenommen. Die errechnete Wechselwirkungsenergie der Propan-Einzelschicht liegt in der Nähe der experimentellen Desorptionsenergie, aber die Diskrepanz ist wesentlich größer als in den vorangegangenen Fällen (4.29 kJ/mol). Dieser Unterschied kann auf die Tatsache, dass die Ergebnisse für Propan noch vorläufig sind und das globale Minimum ist möglicherweise noch nicht lokalisiert werden konnte, zurückgeführt werden. Weitere Untersuchungen werden benötigt, um diese Frage zu beantworten.

Vergleicht man Molekül-Oberflächen-Wechselwirkungsenergien für isolierte Moleküle, so gilt: je größer das Molekül, umso größer ist seine Wechselwirkung mit der Oberfläche. Dies ist höchstwahrscheinlich auf eine größere Oberfläche des Moleküls, die mit dem Substrat interagieren kann zurückzuführen. Gleichzeitig wachsen Kohlenstoff-Magnesium-Abstände mit der Länge der Kohlenwasserstoffkette. Der Vergleich der Wechselwirkungsenergien für isolierte Moleküle legt nahe, dass der Extrapolationsmethode zur Nullbedeckungs-Grenze aus den Versuchsergebnissen ein wichtiges Element fehlt. Im Gegensatz dazu ist die Oberflächen-Schicht-Wechselwirkung für Monopod-Strukturen ziemlich konstant (zwischen 10 und 12 kJ/mol). Jedoch sind die seitlichen (lateralen) Wechselwirkungen mit der Größe der Moleküle erhöht.

Die Studien zeigten, dass die RBGA Methode von großem Vorteil bei der Vereinfachung von Adsorptions-Optimierungsproblemen sein kann. Der Code funktioniert gut, und einige der Ergebnisse sind in sehr guter Übereinstimmung mit dem Experiment. Doch noch sind einige Fragen offen, insbesondere im Falle der Propan-Adsorption. Weitere Studien sind erforderlich, um sie zu beantworten. Darüber hinaus kann eine Korrektur der RBGA Code, um die Leistung und Konvergenzalgorithmus zu erhöhen, angestrebt werden. Bessere Leistung beim Paarungsschritt für das Methan – nämlich, bessere Abstände zwischen den Molekülen zu

halten, um das Zusammendrücken der Moleküle in der ersten Schicht zu verhindern, ist erforderlich. Zusätzlich könnte ein Mechanismus, um Ethan-Moleküle als Ganzes zu halten, das heisst, „interne" Freiheitsgrade (Rotation um die CC-Bindung) einzuführen, dazu beitragen, Geometrien mit zwei getrennt Methylgruppen zu vermeiden. Schließlich würde die Einführung unterschiedlicher, freidefinierbarer starrer Körper die Modellierung von Propan-Molekülen als zwei Methylgruppen und eine Methylengruppe ermöglichen.

8. Outlook

The Rigid Body Genetic algorithm (RBGA) has proven to be a useful global optimization tool, which can be easily applied to investigate weak bound systems, like adsorption processes. Unfortunately, the RBGA method also has its short-comings. Refinement and development work is required to make this method more robust and widely applicable for a range of different theoretical problems. Moreover, the so far achieved results concerning adsorption of hydrocarbons could be further refined using more advance theoretical tools; and additional global optimization could be performed.

From implementation point of view, addition of new functionality to resolve existing problems and address new ones is possible and straightforward. In application to adsorption of hydrocarbons at the magnesium oxide surface, one of the obstacles is the process of initiating and recombining structures, which yields to unreasonable geometries or structures with few imaginary frequencies. An important improvement to the code would be introduction of possibility to define and optimize different types of rigid bodies. The introduction of different rigid bodies would, for example, permit to examine phases consisting of different molecules, which co-exist on the surface or within a bulk. It would also allow investigating the coverage of the surface when different types of molecules are present. Another useful upgrade to the code would be to enable usage of the RBGA optimization for the bulk systems, like zeolites or MOFs (Metal-Organic Frameworks). The possibility to introduce optimization within bulk systems would facilitate investigation of oxidative coupling of methane (OCM) in aluminosilicate zeolite, like Chabazite; and also would available to study molecular crystals, like crystallization processes or rearrangement of molecules in organic semiconductors. In this approach each molecule would be treated as a separate rigid body. The genetic algorithm can be modified to additionally introduce internal degrees of freedom (DoF) within each rigid body. The internal DoF would introduce extra flexibility during optimization. This could prevent the algorithm from generating senseless geometries, especially for cases when we want to treat subparts of a given molecule separately. In turn, this would speed up optimization for these systems.

Modification to the RBGA code to enable working with clusters would make possible investigation of molecular clusters, or complex processes like protein solvation, protein folding (with each amino acid or group of amino acids as single rigid body), dynamics and formation of biological membranes like liposomes, micelles, lipid bilayers (each lipid molecule as a rigid body). An interesting variation of the RBGA algorithm would be a merge between rigid body optimization and unconstrained optimization. The composition of the molecule with predefined parts (like double or triple carbon bonds, benzene rings) could be specified. The specification could be used by the RBGA to construct molecule (or molecules) with predefined functional groups and single atoms. The genetic algorithm could be used to predict possible structures of molecules basing on partial information about their geometry. This might be useful to predict geometries of organic molecules, when only some information can be extracted from experiments. Alternation of this idea could be used to design drugs and enzyme inhibitors in cavity of the enzyme active site. With a given structure of the active site, treated as a rigid body, and a predefined set of organic fragments – also rigid bodies – the algorithm could search for a structure with desired character of the interaction with the protein. This algorithm could also use single atoms as well.

From the hydrocarbon adsorption at the magnesium oxide surface point of view, there are still unanswered questions to which further studies could bring interesting insights. Some of them could be easily and fast addressed, by using so far obtained structures, and recalculating them with different levels of theory, especially to account for the incorrectness due to usage of empirical dispersion correction. Other questions which would include more time consuming investigation, with more RGBA optimizations, could be answered in a longer time perspective. Obtained interaction energies could be validated using higher level of theory. The energies could be recalculated using the Periodic Local MP2 (periodic LMP2) code to account for electron correlation. Structures for single molecule coverage could be optimized with a cluster model while employing higher level of theory methods like MP2 or CC to account for long range interactions and to obtain better estimation of the interaction energies. To investigate dynamic of the monolayer structures MD, calculations can be employed. More RBGA global optimization cycles are certainly required for higher loading of propane at the MgO(001) surface. The formation of third layer of methane, and second layers of ethane, and propane could be investigated the same way we investigated creation of the methane second layer. In addition, introducing point defects or irregularities to the surface, like steps, could give a valuable insight on what is their influence on the formation of hydrocarbon layers. A larger surface model could be used to see if it is possible to obtain stable coverage of separate methane

molecules in the monopod at oxygen site within lower loading of the gas. Other types of rigid bodies on the MgO(001), like carbon oxide, could also be an interesting study case for the RBGA algorithm.

9. Bibliography

Bibliography EndNote

1. Jensen, F. *Introduction to computational chemistry*. Wiley. com: 2007.

2. Cramer, C. J. *Essentials of computational chemistry: theories and models*. Wiley. com: 2005.

3. Kresse, G.; Furthmüller, J. Software VASP, Vienna (1999). *Phys. Rev. B* 1996, 54, 169.

4. Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.* 1989, 162, 165-169.

5. Frisch, M.; Trucks, G.; Schlegel, H. B.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. Gaussian 09, revision A. 02; Gaussian. *Inc., Wallingford, CT* 2009, 270, 271.

6. M.F. Guest, I. J. B., H.J.J. van Dam, P. Sherwood, J.M.H. Thomas, J.H. van Lenthe, R.W.A Havenith, J. Kendrick. The GAMESS-UK electronic structure package: algorithms, developments and applications. *Mol. Phys.* 2005, 103, 719-747.

7. Werner, H.-J.; Knowles, P.; Lindh, R.; Manby, F.; Schütz, M.; Celani, P.; Korona, T.; Rauhut, G.; Amos, R.; Bernhardsson, A. MOLPRO, version 2006.1, a package of ab initio programs. 2006.

8. Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P. å.; Neogrády, P.; Pedersen, T. B.; Pitoňák, M.; Reiher, M.; Roos, B. O. MOLCAS 7: The next generation. *J. Comput. Chem.* 2010, 31, 224-247.

9. Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P. MOLCAS: a program package for computational chemistry. *Comput. Mater. Sci.* 2003, 28, 222-239.

10. Dovesi, R.; Saunders, V.; Roetti, R.; Orlando, R.; Zicovich-Wilson, C.; Pascale, F.; Civalleri, B.; Doll, K.; Harrison, N.; Bush, I. CRYSTAL09 (CRYSTAL09 user's manual). *University of Torino, Torino, Italy* 2009.

11. Te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* 2001, 22, 931-967.

12. Hratchian, H. P.; Schlegel, H. B. *Finding minima, transition states, and following reaction pathways on ab initio potential energy surfaces*. Elsevier: Amsterdam: 2005.

13. Wales, D. *Energy landscapes: applications to clusters, biomolecules and glasses*. Cambridge University Press: 2003.

14. Schlegel, H. B. Exploring potential energy surfaces for chemical reactions: an overview of some practical methods. *J. Comput. Chem.* 2003, 24, 1514-1527.

15. Schlegel, H. B. Geometry optimization. *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2011, 1, 790-809.

16. Schlegel, H. B. *Geometry optimization on potential energy surfaces*. 1995; Vol. 2, p 459-500.

17. Schlegel, H. B. Optimization of equilibrium geometries and transition structures. *J. Comput. Chem.* 1982, 3, 214-218.

18. Liotard, D. A. Algorithmic tools in the study of semiempirical potential surfaces. *Int. J. Quantum Chem* 1992, 44, 723-741.

19. Head, J. D.; Weiner, B.; Zerner, M. C. A survey of optimization procedures for stable structures and transition states. *Int. J. Quantum Chem* 1988, 33, 177-186.

20. Bell, S.; Crighton, J. S. Locating transition states. *J. Chem. Phys.* 1984, 80, 2464.

21. Pulay, P. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules: I. Theory. *Mol. Phys.* 1969, 17, 197-204.

22. Bishop, D. M.; Randic, M. Ab initio calculation of harmonic force constants. *J. Chem. Phys.* 1966, 44, 2480.

23. Gerratt, J.; Mills, I. M. Force Constants and Dipole-Moment Derivatives of Molecules from Perturbed Hartree–Fock Calculations. I. *J. Chem. Phys.* 1968, 49, 1719.

24. Born, M.; Oppenheimer, R. Zur quantentheorie der molekeln. *AnP* 1927, 389, 457-484.

25. Helgaker, T. Optimization of minima and saddle points. In *Lecture Notes in Quantum Chemistry*, Springer: 1992; pp 295-324.

26. Wilson, E. B.; Decius, J.; Cross, P. C. Molecular vibrations. 1955.

27. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab initio molecular orbital theory*. Wiley New York et al.: 1986; Vol. 33.

28. Pulay, P.; Fogarasi, G. Geometry optimization in redundant internal coordinates. *J. Chem. Phys.* 1992, 96, 2856.

29. Fogarasi, G.; Zhou, X.; Taylor, P. W.; Pulay, P. The calculation of ab initio molecular geometries: efficient optimization by natural internal coordinates and empirical correction by offset forces. *J. Am. Chem. Soc.* 1992, 114, 8191-8201.

30. Baker, J.; Kessi, A.; Delley, B. The generation and use of delocalized internal coordinates in geometry optimization. *J. Chem. Phys.* 1996, 105, 192.

31. Baker, J.; Kinghorn, D.; Pulay, P. Geometry optimization in delocalized internal coordinates: An efficient quadratically scaling algorithm for large molecules. *J. Chem. Phys.* 1999, 110, 4986.

32. von Arnim, M.; Ahlrichs, R. Geometry optimization in generalized natural internal coordinates. *J. Chem. Phys.* 1999, 111, 9183.

33. Peng, C.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J. Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J. Comput. Chem.* 1996, 17, 49-56.

34. Farkas, Ö.; Schlegel, H. B. Methods for geometry optimization of large molecules. I. An O (N) algorithm for solving systems of linear equations for the transformation of coordinates and forces. *J. Chem. Phys.* 1998, 109, 7100.

35. Baker, J.; Pulay, P. Geometry optimization of atomic microclusters using inverse-power distance coordinates. *J. Chem. Phys.* 1996, 105, 11100.

36. Billeter, S. R.; Turner, A. J.; Thiel, W. Linear scaling geometry optimisation and transition state search in hybrid delocalised internal coordinates. *Phys. Chem. Chem. Phys.* 2000, 2, 2177-2186.

37. Paizs, B.; Baker, J.; Suhai, S.; Pulay, P. Geometry optimization of large biomolecules in redundant internal coordinates. *J. Chem. Phys.* 2000, 113, 6566.

38. Kudin, K. N.; Scuseria, G. E.; Schlegel, H. B. A redundant internal coordinate algorithm for optimization of periodic systems. *J. Chem. Phys.* 2001, 114, 2919.

39. Bučko, T.; Hafner, J.; Ángyán, J. G. Geometry optimization of periodic systems using internal coordinates. *J. Chem. Phys.* 2005, 122, 124508.

40. Andzelm, J.; King-Smith, R.; Fitzgerald, G. Geometry optimization of solids using delocalized internal coordinates. *Chem. Phys. Lett.* 2001, 335, 321-326.

41. Eckert, F.; Pulay, P.; Werner, H. J. Ab initio geometry optimization for large molecules. *J. Comput. Chem.* 1997, 18, 1473-1483.

42. Baker, J.; Pulay, P. Efficient Geometry Optimization of Molecular Clusters. *J. Comput. Chem.* 2000, 21, 69-76.

43. Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. Systematic ab initio gradient calculation of molecular geometries, force constants, and dipole moment derivatives. *J. Am. Chem. Soc.* 1979, 101, 2550-2560.

44. Schlegel, H. B. Estimating the hessian for gradient-type geometry optimizations. *Theor. Chim. Acta.* 1984, 66, 333-340.

45. Fletcher, R. *Practical methods of optimization*. John Wiley & Sons: 2013.

46. Gill, P. E.; Murray, W.; Wright, M. H. *Practical optimization*. 1981.

47. Wales, D. J. Locating stationary points for clusters in cartesian coordinates. *J. Chem. Soc., Faraday Trans.* 1993, 89, 1305-1313.

48. Wales, D. J. Basins of attraction for stationary points on a potential-energy surface. *J. Chem. Soc., Faraday Trans.* 1992, 88, 653-657.

49. Mezey, P. G. Topology of energy hypersurfaces. *Theor. Chim. Acta.* 1982, 62, 133-161.

50. Mezey, P. G. The topology of energy hypersurfaces II. Reaction topology in euclidean spaces. *Theor. Chim. Acta.* 1983, 63, 9-33.

51. Nelder, J. A.; Mead, R. A simplex method for function minimization. *Comput. J.* 1965, 7, 308-313.

52. Lewis, R. M.; Torczon, V.; Trosset, M. W. Direct search methods: then and now. *J. Comput. App. Math.* 2000, 124, 191-207.

53. Kolda, T. G.; Lewis, R. M.; Torczon, V. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Rev.* 2003, 45, 385-482.

54. García-Palomares, U. M.; Rodríguez, J. F. New sequential and parallel derivative-free algorithms for unconstrained minimization. *SIAM J. Optimiz.* 2002, 13, 79-96.

55. Han, L.; Neumann, M. Effect of dimensionality on the Nelder–Mead simplex method. *Optim. Method. Softw.* 2006, 21, 1-16.

56. Lagarias, J. C.; Reeds, J. A.; Wright, M. H.; Wright, P. E. Convergence properties of the Nelder--Mead simplex method in low dimensions. *SIAM J. Optimiz.* 1998, 9, 112-147.

57. Head, J. D.; Zerner, M. C. A Broyden—Fletcher—Goldfarb—Shanno optimization procedure for molecular geometries. *Chem. Phys. Lett.* 1985, 122, 264-270.

58. Head, J. D. Partial optimization of large molecules and clusters. *J. Comput. Chem.* 1990, 11, 67-75.

59. Baker, J.; Hehre, W. J. Geometry optimization in cartesian coordinates: The end of the Z-matrix? *J. Comput. Chem.* 1991, 12, 606-610.

60. Baker, J. Techniques for geometry optimization: A comparison of Cartesian and natural internal coordinates. *J. Comput. Chem.* 1993, 14, 1085-1100.

61. Baker, J.; Chan, F. The location of transition states: A comparison of Cartesian, Z-matrix, and natural internal coordinates. *J. Comput. Chem.* 1996, 17, 888-904.

62. Powell, M. J. D. In *Nonlinear optimization 1981*, NATO Conference Series. Series II: Systems Science, Proceedings of the NATO Advanced Research Institute, held at Cambridge (UK), July, 1981, London: Academic Press, 1982, edited by Powell, MJD, 1982.

63. Scales, L. *Introduction to non-linear optimization*. Springer-Verlag New York, Inc.: 1985.

64. Hooke, R.; Jeeves, T. A. ``Direct Search"Solution of Numerical and Statistical Problems. *J. ACM* 1961, 8, 212-229.

65. Torczon, V. On the convergence of pattern search algorithms. *SIAM J. Optimiz.* 1997, 7, 1-25.

66. Schlegel, H. B. Optimization of equilibrium geometries and transition structures. In *Advances in Chemical Physics; Ab Initio Methods in Quantum Chemistry-I*, 1987; pp 249-286.

67. Fletcher, R.; Reeves, C. M. Function minimization by conjugate gradients. *Comput. J.* 1964, 7, 149-154.

68. Polak, E. *Computational methods in optimization: a unified approach*. Access Online via Elsevier: 1971; Vol. 77.

69. Hoare, M.; McInnes, J. Morphology and statistical statics of simple microclusters. *AdPhy* 1983, 32, 791-821.

70. Hoare, M.; McInnes, J. Statistical mechanics and morphology of very small atomic clusters. *Faraday Discuss. Chem. Soc.* 1976, 61, 12-24.

71. Doye, J. P.; Wales, D. J. Saddle points and dynamics of Lennard-Jones clusters, solids, and supercooled liquids. *J. Chem. Phys.* 2002, 116, 3777.

72. Stillinger, F. H.; Weber, T. A. Packing structures and transitions in liquids and solids. *Science* 1984, 225, 983-989.

73. Northby, J. Structure and binding of Lennard-Jones clusters: 13≤ N≤ 147. *J. Chem. Phys.* 1987, 87, 6166.

74. Wille, L. T.; Vennik, J. Computational complexity of the ground-state determination of atomic clusters. *J. Phys. A: Math. Gen.* 1985, 18, L419.

75. Papadimitriou, C. H. Computational complexity. 2003.

76. Hamacher, K.; Wenzel, W. Scaling behavior of stochastic minimization algorithms in a perfect funnel landscape. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids,* 1999, 59, 938.

77. Kirkpatrick, S. Optimization by simulated annealing: Quantitative studies. *J. Stat. Phys.* 1984, 34, 975-986.

78. Kirkpatrick, S.; Jr., D. G.; Vecchi, M. P. Optimization by simmulated annealing. *Science* 1983, 220, 671-680.

79. Flannery, B. P.; Press, W. H.; Teukolsky, S. A.; Vetterling, W. Numerical recipes in C. *Press Syndicate of the University of Cambridge, New York* 1992.

80. Kalos, M. H.; Whitlock, P. A. *Monte carlo methods*. John Wiley & Sons: 2008.

81. Berg, B. A.; Neuhaus, T. Multicanonical algorithms for first order phase transitions. *Phys. Lett. B* 1991, 267, 249-253.

82. Lee, J. New Monte Carlo algorithm: entropic sampling. *Phys. Rev. Lett.* 1993, 71, 211.

83. Hesselbo, B.; Stinchcombe, R. B. Monte Carlo simulation and global optimization without parameters. *Phys. Rev. Lett.* 1995, 74, 2151.

84. Lyubartsev, A.; Martsinovski, A.; Shevkunov, S.; Vorontsov-Velyaminov, P. New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles. *J. Chem. Phys.* 1992, 96, 1776.

85. Marinari, E.; Parisi, G. Simulated tempering: a new Monte Carlo scheme. *Europhys. Lett.* 1992, 19, 451.

86. Hansmann, U. H. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* 1997, 281, 140-150.

87. Wenzel, W.; Hamacher, K. Stochastic tunneling approach for global minimization of complex potential energy landscapes. *Phys. Rev. Lett.* 1999, 82, 3003.

88. Barhen, J.; Protopopescu, V.; Reister, D. TRUST: A deterministic algorithm for global optimization. *Science* 1997, 276, 1094-1097.

89. Cetin, B.; Barhen, J.; Burdick, J. Terminal repeller unconstrained subenergy tunneling (TRUST) for fast global optimization. *J. Optimiz. Theory App.* 1993, 77, 97-126.

90. Frantz, D.; Freeman, D. L.; Doll, J. Reducing quasi-ergodic behavior in Monte Carlo simulations by J-walking: Applications to atomic clusters. *J. Chem. Phys.* 1990, 93, 2769.

91. Serra, P.; Stanton, A. F.; Kais, S.; Bleil, R. E. Comparison study of pivot methods for global optimization. *J. Chem. Phys.* 1997, 106, 7170.

92. Rossi, G.; Ferrando, R. Global optimization by excitable walkers. *Chem. Phys. Lett.* 2006, 423, 17-22.

93. Amara, P.; Hsu, D.; Straub, J. E. Global energy minimum searches using an approximate solution of the imaginary time Schrödinger equation. *J. Phys. Chem.* 1993, 97, 6715-6721.

94. Sylvain, M.; Somorjai, R. Novel approach for computing the global minimum of proteins. 2. One-dimensional test cases. *J. Phys. Chem.* 1991, 95, 4147-4152.

95. Somorjai, R. Novel approach for computing the global minimum of proteins. 1. General concepts, methods, and approximations. *J. Phys. Chem.* 1991, 95, 4141-4146.

96. Anderson, J. B. A random-walk simulation of the Schrödinger equation: H. *J. Chem. Phys.* 1975, 63, 1499.

97. Finnila, A.; Gomez, M.; Sebenik, C.; Stenson, C.; Doll, J. Quantum annealing: A new method for minimizing multidimensional functions. *Chem. Phys. Lett.* 1994, 219, 343-348.

98. Andricioaei, I.; Straub, J. E. Finding the needle in the haystack: Algorithms for conformational optimization. *Comput. Phys.* 1996, 10, 449.

99. Levy, A.; Montalvo, A. The tunneling algorithm for the global minimization of functions. *SIAM J. Sci. Stat. Comp.* 1985, 6, 15-29.

100. Shalloway, D. Application of the renormalization group to deterministic global minimization of molecular conformation energy functions. *J. Global Optim.* 1992, 2, 281-311.

101. Doye, J. P.; Leary, R. H.; Locatelli, M.; Schoen, F. Global optimization of morse clusters by potential energy transformations. *INFORMS J. Comput.* 2004, 16, 371-379.

102. Besold, G.; Risbo, J.; Mouritsen, O. G. Efficient Monte Carlo sampling by direct flattening of free energy barriers. *Comput. Mater. Sci.* 1999, 15, 311-340.

103. Hansmann, U. H.; Okamoto, Y. The generalized-ensemble approach for protein folding simulations. *Ann. Rev. Comp. Phys.* 1999.

104. Hansmann, U. H. Generalized ensemble techniques and protein folding simulations. *Comput. Phys. Commun.* 2002, 147, 604-607.

105. Hansmann, U. H.; Okamoto, Y. New Monte Carlo algorithms for protein folding. *Curr. Opin. Struct. Biol.* 1999, 9, 177-183.

106. Huber, T.; Torda, A. E.; van Gunsteren, W. F. Local elevation: a method for improving the searching properties of molecular dynamics simulation. *J. Comput. Aided Mol. Des.* 1994, 8, 695-708.

107. Grubmüller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids,* 1995, 52, 2893.

108. Glover, F. Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research* 1986, 13, 533-549.

109. Glover, F. Tabu search—part I. *INFORMS J. Comput.* 1989, 1, 190-206.

110. Glover, F. Tabu search—part II. *INFORMS J. Comput.* 1990, 2, 4-32.

111. Laio, A.; Parrinello, M. Escaping free-energy minima. *P. Natl. Acad. Sci.* 2002, 99, 12562-12566.

112. Wille, L. T.; Hansmann, U. H. Global optimization by energy landscape paving. *Phys. Rev. Lett.* 2002.

113. Wales, D. J.; Doye, J. P. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem. A* 1997, 101, 5111-5116.

114. Wales, D. J.; Scheraga, H. A. Global optimization of clusters, crystals, and biomolecules. *Science* 1999, 285, 1368-1372.

115. Leary, R. H. Global optimization on funneling landscapes. *J. Global Optim.* 2000, 18, 367-383.

116. Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* 1977, 23, 187-199.

117. Blundell, T.; Sibanda, B.; Sternberg, M.; Thornton, J. Knowledge-based prediction of protein structures. *Nature* 1987, 326, 26.

118. Goldberg, D. E.; Holland, J. H. Genetic algorithms and machine learning. *Mach. Learn.* 1988, 3, 95-99.

119. Goldberg, D. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley: 1989.

120. Davis, L. Handbook of genetic algorithms. 1991.

121. Michalewicz, Z. *Genetic algorithms+ data structures= evolution programs*. springer: 1996.

122. Goldberg, D. E.; Voessner, S. Optimizing global-local search hybrids. *Urbana* 1999, 51, 61801.

123. Goldberg, D. E.; Deb, K.; Clark, J. H. Genetic algorithms, noise, and the sizing of populations. *Complex systems* 1991, 6, 333-362.

124. Judson, R. Genetic Algorithms and Their Use in Chemistry. *Rev. Comput. Chem.* 10, 1-73.

125. Leardi, R. Genetic algorithms in chemometrics and chemistry: a review. *J. Chemom.* 2001, 15, 559-569.

126. Hibbert, D. B. Genetic algorithms in chemistry. *Chemometrics Intellig. Lab. Syst.* 1993, 19, 277-293.

127. Leardi, R. Genetic algorithms in chemistry. *J. Chromatogr. A* 2007, 1158, 226-233.

128. Niazi, A.; Leardi, R. Genetic algorithms in chemometrics. *J. Chemom.* 2012, 26, 345-351.

129. Deaven, D.; Ho, K. Molecular Geometry Optimization with a Genetic Algorithm. *Phys. Rev. Lett.* 1995, 75, 288-291.

130. Chuang, F. C.; Ciobanu, C. V.; Shenoy, V. B.; Wang, C. Z.; Ho, K. M. Finding the reconstructions of semiconductor surfaces via a genetic algorithm. *Surf. Sci.* 2004, 573, L375-L381.

131. Włodarczyk, R.; Sierka, M.; Kwapień, K.; Sauer, J.; Carrasco, E.; Aumer, A.; Gomes, J. F.; Sterrer, M.; Freund, H.-J. Structures of the Ordered Water Monolayer on MgO(001). *The Journal of Physical Chemistry C* 2011, 115, 6764-6774.

132. Johnston, R. L. Evolving better nanoparticles: Genetic algorithms for optimising cluster geometries. *Dalton Trans.* 2003, 4193.

133. Holland, J. H. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press: 1975.

134. Hartke, B. Global geometry optimization of clusters using genetic algorithms. *J. Phys. Chem.* 1993, 97, 9973-9976.

135. Hartke, B. Global geometry optimization of clusters using a growth strategy optimized by a genetic algorithm. *Chem. Phys. Lett.* 1995, 240, 560-565.

136. Niesse, J. A.; Mayne, H. R. Minimization of small silicon clusters using the space-fixed modified genetic algorithm method. *Chem. Phys. Lett.* 1996, 261, 576-582.

137. Xiao, Y.; Williams, D. E. Genetic algorithm: a new approach to the prediction of the structure of molecular clusters. *Chem. Phys. Lett.* 1993, 215, 17-24.

138. Zeiri, Y. Prediction of the lowest energy structure of clusters using a genetic algorithm. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids,* 1995, 51, R2769-R2772.

139. Daven, D.; Tit, N.; Morris, J.; Ho, K. Structural optimization of Lennard-Jones clusters by a genetic algorithm. *Chem. Phys. Lett.* 1996, 256, 195-200.

140. Hartke, B. Global optimization. *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2011, 1, 879-887.

141. Hartke, B. Application of evolutionary algorithms to global cluster geometry optimization. In *Applications of Evolutionary Computation in Chemistry*, Springer: 2004; pp 33-53.

142. Bush, T.; Catlow, C. R. A.; Battle, P. Evolutionary programming techniques for predicting inorganic crystal structures. *J. Mater. Chem.* 1995, 5, 1269-1272.

143. Woodley, S.; Battle, P.; Gale, J.; Catlow, C. A. The prediction of inorganic crystal structures using a genetic algorithm and energy minimisation. *Phys. Chem. Chem. Phys.* 1999, 1, 2535-2542.

144. Bazterra, V. E.; Ferraro, M. B.; Facelli, J. C. Modified genetic algorithm to model crystal structures. I. Benzene, naphthalene and anthracene. *J. Chem. Phys.* 2002, 116, 5984-5991.

145. Oganov, A. R.; Glass, C. W. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *J. Chem. Phys.* 2006, 124, 244704.

146. Abraham, N.; Probert, M. A periodic genetic algorithm with real-space representation for crystal structure and polymorph prediction. *Phys. Rev. B: Condens. Matter* 2006, 73, 224104.

147. Briggs, R. M.; Ciobanu, C. V. Evolutionary approach for finding the atomic structure of steps on stable crystal surfaces. *Phys. Rev. B: Condens. Matter* 2007, 75, 195415.

148. Zhu, Q.; Li, L.; Oganov, A. R.; Allen, P. B. Evolutionary method for predicting surface reconstructions with variable stoichiometry. *Phys. Rev. B: Condens. Matter* 2013, 87, 195317.

149. Goldstein, H. *Classical Mechanics*. Addison-Wesley: 1980.

150. Chakrabarti, D.; Wales, D. J. Simulations of rigid bodies in an angle-axis framework. *Phys. Chem. Chem. Phys.* 2009, 11, 1970-6.

151. Schröder, K.-P. A method for the separate computation of intermolecular vibrational frequencies with application on the H_2O - HF and $(H2O)_n$ (n= 2–6) complexes. *Chem. Phys.* 1988, 123, 91-101.

152. Sierka M., S. J. Finding transition structures in extended systems - A strategy based on a combined quantum mechanics–empirical valence bond approach. *J. Chem. Phys.* 2000, 112, 6983.

153. Lunsford, J. H. The Catalytic Oxidative Coupling of Methane. *Angew. Chem. Int. Ed.* 1995, 34, 970-980.

154. Klabunde, K. J.; Nieves, I. Interaction of activated magnesium oxide surfaces with spin traps. *J. Phys. Chem.* 1988, 92, 2521-2525.

155. Press, W. Structure and Phase Transitions of Solid Heavy Methane (CD). *J. Chem. Phys.* 1972, 56, 2597.

156. Landolt-Börnstein. Magnesium oxide (MgO) crystal structure, lattice parameters, thermal expansion. In *II-VI and I-VII Compounds; Semimagnetic Compounds Group III Condensed Matter*, 1999; Vol. 41B, pp 1-6.

157. Coulomb, J.; Madih, K.; Croset, B.; Lauter, H. Evidence of a Square Two-Dimensional Solid of Methane Physisorbed on the (100) Surface of Magnesium Oxide. *Phys. Rev. Lett.* 1985, 54, 1536-1538.

158. Jung, D.; Cui, J.; Frankl, D.; Ihm, G.; Kim, H. Y.; Cole, M. He atom near methane-plated MgO: Interaction and scattering. *Phys. Rev. B: Condens. Matter* 1989, 40, 11893-11901.

159. Larese, J. Neutron scattering studies of the structure and dynamics of methane absorbed on MgO (100) surfaces. *Physica B* 1998, 248, 297-303.

160. Madih, K.; Croset, B.; Coulomb, J.; Lauter, H. Thin methane film growing mode on MgO/100 surface. *Europhys. Lett.* 1989, 8, 459.

161. Jung, D.; Cui, J.; Frankl, D. Dynamics and kinetics of monolayer CH4 on MgO(001) studied by helium-atom scattering. *Phys. Rev. B: Condens. Matter* 1991, 43, 10042-10050.

162. Larese, J. Z.; Hastings, J. M.; Passell, L.; Smith, D.; Richter, D. Rotational tunneling of methane on MgO surfaces: A neutron scattering study. *J. Chem. Phys.* 1991, 95, 6997.

163. Larese, J.; Asmussen, B.; Adams, M.; Carlile, C.; Martin, D.; Ferrand, M. Rotational tunneling studies of methane films adsorbed on MgO: Crossover from two-to-three dimensions? *Physica B* 1996, 226, 221-223.

164. Larese, J.; y Marero, D.; Sivia, D.; Carlile, C. Tracking the Evolution of Interatomic Potentials with High Resolution Inelastic Neutron Spectroscopy. *Phys. Rev. Lett.* 2001, 87.

165. Gay, J.; Stocker, P.; Degenhardt, D.; Lauter, H. Rotational diffusion of methane molecules adsorbed on MgO(100). *Phys. Rev. B: Condens. Matter* 1992, 46, 1195-1197.

166. Bruch, L. W.; Larese, J. Z. Initial stages of square lattice stacks of CH_{4}/MgO(001). *Phys. Rev. B: Condens. Matter* 2012, 85.

167. Tait, S. L.; Dohnálek, Z.; Campbell, C. T.; Kay, B. D. n-alkanes on MgO (100). II. Chain length dependence of kinetic desorption parameters for small n-alkanes. *J. Chem. Phys.* 2005, 122, 164708.

168. Deprick, B.; Julg, A. A theoretical study of adsorption of CH4 on the (100) faces of MgO and NaCl. *New J. Chem.* 1987, 11, 299-302.

169. Girard, C.; Girardet, C. Potential energy calculations for argon and methane adsorbed on MgO (001) substrate. *Chem. Phys. Lett.* 1987, 138, 83-89.

170. Alavi, A. Molecular-dynamics simulation of methane adsorbed on MgO: Evidence for a Kosterlitz-Thouless transition. *Mol. Phys.* 1990, 71, 1173-1191.

171. Ferrari, A.; Huber, S.; Knözinger, H.; Neyman, K.; Rösch, N. FTIR spectroscopic and density functional model cluster studies of methane adsorption on MgO. *J. Phys. Chem. B* 1998, 102, 4548-4555.

172. Todnem, K.; Borve, K. J.; Nygren, M. Molecular adsorption of methane and methyl onto MgO(100) - An embedded-cluster study. *Surf. Sci.* 1999, 421, 296-307.

173. Drummond, M.; Sumpter, B.; Shelton, W.; Larese, J. Density functional investigation of the adsorption of a methane monolayer on an MgO(100) surface. *Phys. Rev. B: Condens. Matter* 2006, 73.

174. Drummond, M. L.; Sumpter, B. G.; Shelton, W. A.; Larese, J. Z. Electronic structure investigation of surface-adsorbate and adsorbate-adsorbate interactions in multilayers of CH_4 on MgO (100). *Journal of Physical Chemistry C* 2007, 111, 966-976.

175. Drummond, M. L.; Sumpter, B. G.; Shelton, W. A.; Larese, J. Z. In *Adsorption of Methane on the (100) Surface of MgO: Insight into Surface-Adsorbate and Adsorbate-Adsorbate Interactions from First-Principles Calculations*, MRS Proceedings, Cambridge Univ Press: 2006.

176. Trevethan, T.; Shluger, A. Building blocks for molecular devices: Organic molecules on the MgO (001) surface. *The Journal of Physical Chemistry C* 2007, 111, 15375-15381.

177. Pisani, C.; Maschio, L.; Casassa, S.; Halo, M.; Schutz, M.; Usvyat, D. Periodic local MP2 method for the study of electronic correlation in crystals: Theory and preliminary applications. *J. Comput. Chem.* 2008, 29, 2113-24.

178. Tosoni, S.; Sauer, J. Accurate quantum chemical energies for the interaction of hydrocarbons with oxide surfaces: CH(4)/MgO(001). *Phys. Chem. Chem. Phys.* 2010, 12, 14330-40.

179. Boese, D.; Sauer, J. Accurate adsorption Energies for Small Molecules on Oxide Surfaces: CH4/MgO(001) and C2H6/MgO(001). *manuscript 2013*.

180. Alessio, M. Personal communication.

181. Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* 1996, 6, 15-50.

182. Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B: Condens. Matter* 1996, 54, 11169.

183. Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* 1996, 77, 3865-3868.

184. Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]. *Phys. Rev. Lett.* 1997, 78, 1396-1396.

185. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B: Condens. Matter* 1994, 50, 17953.

186. Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B: Condens. Matter* 1999, 59, 1758.

187. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* 2006, 27, 1787-1799.

188. Piccini, G.; Sauer, J. Quantum Chemical Free Energies: Structure Optimization and Vibrational Frequencies in Normal Modes. *Journal of Chemical Theory and Computation* 2013, 9, 5038-5045.

189. Bouř, P.; Keiderling, T. A. Partial optimization of molecular geometry in normal coordinates and use as a tool for simulation of vibrational spectra. *J. Chem. Phys.* 2002, 117, 4126-4132.

190. Perdew, J. P.; Burke, K.; Wang, Y. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys. Rev. B: Condens. Matter* 1996, 54, 16533.

191. Tait, S. L.; Dohnálek, Z.; Campbell, C. T.; Kay, B. D. n-alkanes on MgO(100). I. Coverage-dependent desorption kinetics of n-butane. *J. Chem. Phys.* 2005, 122, 164707.

10.Publications

Kinetic Study of the Reaction of Vanadium and Vanadium-Titanium Oxide Cluster Anions with SO_2

E. Janssens, S. Lang, M. Brümmer, A. Niedziela, G. Santambrogio, K. Asmis, J. Sauer

Phys. Chem. Chem. Phys 14 (**2012**) 14344-14353

Structural Variability in Transition Metal Oxide Clusters: Gas Phase Vibrational Spectroscopy of $V_3O_{6-8}^+$

K. R. Asmis, T. Wende, M. Brümmer, O. Gause, G. Santambrogia, E. C. Stanca-Kaposta, J. Döbler, A. Niedziela, J. Sauer

Phys. Chem. Chem. Phys. 14 (**2012**) 9377-9388

11. Acknowledgements

I would like to thank Prof. Joachim Sauer for giving me the opportunity to work in his group and I would like to thank Prof. Marek Sierka for convincing him to do so.

Special thanks go to Dr. Bischoff and Dr. Oncak for commenting and proofreading the thesis, making it better than it would have been otherwise.

I would like to thank Frank Fischer for help with translating the summary.

Many thanks go to group members for support on specific software: Asbjörn Burow (Turbomole), Dr. Joachim Paier (VASP), Radosław Włodarczyk (DODO, Surfaces), GiovanniMaria Piccini (NMopt, Thermo, Anharm).

I want to recognize Cornelia Krell and Kristin Kuschnerik for help with bureaucracy and formalities. With their support I was able to concentrate on my scientific work. Likewise, without Thomas Dargel's IT support no real science would be possible conduct.

I would also like to thank the IMPRS school for giving me the financial support to attend external courses and conferences; and an opportunity to meet professors, professionals, and peers from different fields of chemistry.

For general discussion and on MgO special acknowledgments go to: Dr. Florian Bischoff, Dr. Asbjorn Burow, Dr. Karolina Kwapien, Dr. Sabrina Sicolo, Dr. Daniel Boese, GiovanniMaria Piccini, Radosław Włodarczyk, Frank Fischer, Matthias Baldofski, Maristella Alessio

I would like thank all current and former members of the quantum chemistry group and IMPRS school for giving me the chance to be the part of the international scientific community, to learn from you, to broaden and share my knowledge.

12. Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin,

Andrzej Niedziela