Theoretical study of anisotropic effects in crystal growth on the example of concrete and calcium oxalate monohydrate

Vilhelm Ekberg

February 7, 2019

Supervisor: Valera Veryazov **Examiner:** Jan Forsman

Theoretical Chemistry
Department of Chemistry
Lund University
SE-221 00 Lund, Sweden

Abstract

While many crystals, both natural and artificial, grow isotropically, there are examples of crystals growing anisotropically. The aim of this theoretical study was to confirm the growth modes displayed by three different crystals, representing 3D, 2D and 1D structures. This was done by using energetic calculations with both molecular mechanics, with $\text{ReaxFF}_{\text{SiO}}$ as the force field, and semi-empirical methods, with PM6 as the method. To start with, the 3D structure was represented by CaO, and both the $\text{ReaxFF}_{\text{SiO}}$ and PM6 results confirmed the isotropic growth mode of this rock salt structure. As for the 2D structure, nanoplatelets of calcium silicate hydrate (C-S-H), which make up the bulk phase in concrete, were investigated. These were modelled as C-S-H clusters with approximate cylinder shapes and adjustable chemical bonds and compositions. The $\text{ReaxFF}_{\text{SiO}}$ calculations showed that the energetic gain was higher when the radius was increased than when the height was increased. This was in agreement with the platelet shape of real C-S-H nanoparticles. However, the high computational costs of semi-empirical methods meant that the PM6 results did not represent a large enough cylinder size span to confirm the $\text{ReaxFF}_{\text{SiO}}$ results. Lastly, the 1D structure was represented in the form of calcium oxalate monohydrate (COM), which may form elongated shapes such as raphides and styloids. Due to the $\text{ReaxFF}_{\text{SiO}}$ force field not being optimized for COM crystals, only PM6 calculations were performed. However, the computational demands of semi-empirical methods again limited the number of PM6 data points considerably. Therefore, it could not be concluded if the 1D growth mode was the most energetically favourable.

Keywords: anisotropic crystal growth, COM, C-S-H, PM6, $\text{ReaxFF}_{\text{SiO}}$

Acknowledgements

I wish to acknowledge my supervisor Dr Valera Veryazov, who has provided me with invaluable support. Thanks to him, I have been brought through every aspect of the project, from literature search via execution to conclusion. He has supplied me with all the necessary theoretical and practical tools for producing the results as well as interpreting and understanding them. Furthermore, he has taken a keen interest in my personal development as a researching student as well as promoted my understanding of the science world. As the brilliant mastermind of the study of anisotropic crystal growth of C-S-H in this project, his help cannot be overestimated.

I also want to extend my sincerest gratitudes to Dr. Alexei Abrikossov, who, among many things, has produced the database necessary for me to yield my results. He has also performed crucial bug fixes and introduced me to the relevant computing softwares. An essential companion of Dr. Valera Veryazov in the study of both C-S-H and C-A-S-H, he too has generously provided me with the required support for completing the project. Furthermore, the work of Marc Alías Rodríguez, the author of Puck and a study relating to this, has also been invaluable in producing the foundation to this study. Since much of the structure and methodology of this study has been connected to his ambitious work, he is worthy of his share of acknowledgement. Even though we have never met, I feel connected to him through our mutual focus on the study of the C-S-H growth.

Lastly, the staff at the Lund Theoretical Chemistry department have enriched my stay here with interesting discussions about life, work and everything in between. They welcomed me as member of their group from day one, a context which I consider myself privileged to have been a part of. I also want to humbly thank my parents for encouraging me and supporting me in my studies.

List of abbreviations

COM, Calcium oxalate monohydrate
COD, Calcium oxalate dihydrate
COT, Calcium oxalate trihydrate
CCN, Cement chemist notation
C-S-H, Calcium silicate hydrate
HF-SCF, Hartree-Fock self-consistent field
MD, Molecular dynamics
MM, Molecular mechanics
NDDO, Neglect of diatomic differential overlap
PM6, Parametric method number 6
ReaxFF, Reactive force field
STO, Slater type orbital

Contents

1 Introduction

1.1 Background and objective

It is a well known fact that crystals can adapt various shapes. For instance, whereas some crystals grow to isotopic shapes, others grow anisotropically. These shapes are exemplified in everything from NaCl or diamond (isotropic) to platelet-shaped concrete nanoparticles [1] or needle-shaped *raphides* (anisotropic), which are found in plants [2]. Furthermore, crystals do not grow indefinitely, and there is a wide size span, ranging from the nanometre scale via μm and mm to very large crystals with sizes in the order of meters. For example, the dominant sizes of some plagioclase crystal samples from Makaopuhi lava lake, Kilauea Volcano, Hawaii are around a few tens of μm [3], Antarctic ice crystals with dimensions in the order of a few mm^2 have been discovered hundreds of meters down into the dome C ice core [4], while the raphides grow to lengths of tens of μm [5]. On a much larger scale, gypsum crystals with lengths of up to 12 m in the Naica mine in Mexico have been discovered [6].

Figure 1: Gypsum crystals of the Naica cave, Chihuahua, Mexiko [7].

For many inorganic crystals grown from solution, the degree of supersaturation, the driving force for growth, is controlled by programming the temperature [8]. When it comes to protein crystal growth, the supersaturation is regulated by adding components which lower the solubility and/or by removing solvent [8]. As for crystals grown from melt, the growth rate is usually regulated by the heat flow [9]. In for instance the *Czochralski method*, where the crystal seed is slowly drawn from the melt, the heat flow is determined by, among others, the design of the crystal puller the amount of power supplied to the melt [9].

When it comes to analysing crystal structures on the atomistic level, powder diffraction experiments with X-rays and neutrons are very useful. Though X-ray diffraction experiments can probe size dimensions as low as 1-10 Å, neutron diffraction can reach even lower sizes at around 0.1 Å. The scattering intensity of X-rays is larger for heavier elements (Hg) than for lighter ones (O or H, for example), while neutrons offer significantly large scattering intensities for H. [10]

Impurities in a crystal grown from solution may stem from lattice substitution, for example. This can occur if the molecular structures of the impurity and crystal solute are similar. Lattice substitution is the reason why impurities of **L**-leucine, **L**-valine or α-amino butyric acid are incorporated into **L**-isoleucine crystals. Therefore, the crystal purity can be regulated by manipulating the solubilities of the crystal solute and impurity, In the case of amino acids, this

involves tuning the pH, temperature and electrolyte or co-solvent concentrations. [11]

Another example of the need to remove crystal impurities is in the synthesis of diamonds. Nitrogen impurities are the most common impurities, and they may stem from the solvent metal or carbon source material used during the synthesis. The concentration of nitrogen impurities is reduced by adding a nitrogen getter (Al or Ti, for instance) to the solvent metal, though this promotes incorporation of metal impurities into the diamond crystal. In order to limit the inclusion of metal impurities, the growth rate of the diamond crystal must be lower compared to the growth rate when a nitrogen getter is not used. This may be achieved via the temperature gradient method, where for example the temperature difference between the carbon source and the seed crystal may be controlled. [12]

Yet another example where crystal purity is desired is in the electronics industry, where high-purity silicon is used for computer chips [13] and solar cells [14], among others. The Siemens process is the traditional purification scheme, and it starts off with metallurgical silicon as the raw material [14]. This is hydrochlorinated to form trichlorosilane, which is then purified via distillation and decomposed to form purified silicon [14]. Crystal growth is relevant not only for physical chemistry, but also physics, materials science and geophysics [15]. Furthermore, the research during recent times has increasingly been shifting focus to a microscopic view of crystal growth [16]. In this study, the nanometre scale will be relevant, and the anisotropy in the crystal qrowth can be used for synthesising anisotropic nanoparticles and colloidal building blocks [17]. Examples of methods within this industrial branch are emulsion drying, surface templating and the synthesis of metal nanoparticles from plant extracts, fungi and viruses [17]. Consequently, the industrial potential of creating very small structures with defined shapes is another reason to study anisotropic growth [17].

The anisotropic growth is of interest in this study: what is the driving force for anisotropic growth and how should it be modelled theoretically? This study will focus on the bulk phase of concrete, *calcium silicate hydrate* (C-S-H), where the nanoparticles form platelets [1, 18], as earlier mentioned. Additionally, the anisotropic growth of elongated crystals of *calcium oxalate monohydrate* (COM), raphides and also styloids [2], will be studied. Gaining insight into the driving force for anisotropic growth will help shed light on why some crystals, both natural and artificial, grow anisotropically.

1.2 Methods for growing crystals

Growing crystals is a very complicated procedure with many different parameters to consider. One way of doing so is via the earlier mentioned *Czochralski method*, in which the crystal seed is slowly drawn vertically from the melt in a heated crucible. The seed rotates slowly, in order to counteract any circular asymmetries in the heating field. With this method, it is possible to obtain crystal with very directional shapes. Another melt growth method is the *Bridgman technique*, where the sample is melted in a furnace and then slowly lowered from it while being kept in a container. As the lower region of the container becomes cold enough, the sample crystallises, though it is impossible to control the direction of growth. [9]

If crystals cannot be grown from the melt, solution growth methods are available. It might be that the molecules decompose below the melting point, or the crystal structure might be disrupted by a phase transition as it is cooled below the melting point to room temperature. For example, the crystals can be grown from a spinning seed in an aqueous solution that is under continuous stirring. However, it can take weeks to grow crystals of appreciable size. [9]

An alternative solution growth method is the *hydrothermal method*, where the solvent consists of high-temperature water vapour instead. Usually, the seed is placed in the higher part of an autoclave, while the feeding material is found in the lower part of the autoclave, which is kept at a higher temperature. When water is added into the lower part of the autoclave, superheated steam forms that allows the feeding material to convect to the seed. Again, it can take several weeks to grow crystals. Quartz crystals are grown in this way, since it undergoes a phase transition at roughly 580^oC. It is also believed that natural crystals of various minerals, including quartz, have grown under the surface of the earth in similar conditions. [9]

Figure 2: A quartz crystal from Tibet [19].

1.3 Parameters and models for crystal growth

Various factors govern the crystal growth including, for instance, the type of medium in which the crystal is suspended (melt, solution, vapor), growth rate and growth mode [20]. Examples of growth modes are *layer-by-layer*, where one layer is finished before the next is started, or *three-dimensional growth*, where several layers grow at the same time [20]. Another point of relevance is that although crystal growth is intrinsically a non-equilibrium process, local equilibrium can be assumed [16].

In *transformation-rate-limited growth*, the growth rate is limited by the rate of the atomic processes. This is the case in solution growth methods, where it is difficult to create local supersaturation high enough around the seed to promote layer growth at an appreciable rate, or glass crystallisation, where the atoms and molecules are not very mobile. There are also examples of *thermally limited growth*, where the rate of crystallization is limited by the heat flow. Here, the heat conducted via the crystal to the interface and the heat expelled by the crystals as they are crystallised must be carried away from the solid-liquid interface. This heat flow is used in both the Czochralski and Bridgman techniques in order to control the growth process. The crystal is kept at a lower temperature than the liquid via the heaters, and the rate at which the crystal is lowered out of the Bridgman furnace or drawn from the melt in the Czochralski case is carefully controlled. [9]

The growth rate can mathematically be expressed as the following product:

$$v = a\nu^+ f\mu_k \tag{1}$$

where a is related to the dimensions of the growing unit, ν^+ is the frequency of atoms entering active growth sites and f is the fraction of active grown sites out of the interface sites; it reflects

the roughness of the surface. Therefore, a higher roughness serves to increase the growth rate. Finally, μ_k represents the driving force provided by the chemical potential difference between the two phases at the interface. [9]

There is the possibility of *dendritic growth*, or tree-like growth. This occurs when the kinetics at the interface are so fast that it is not stable to have a planar growth front. Consequently, the random motions of diffusing atoms and molecules arriving at the interface dictate the growth shape. These conditions may occur, for instance, if a melt is cooled to well below its freezing point - supercooled. [9]

Figure 3: A dendritic crystal of silver as a mineral, the black colour comes from slightly oxidized silver [21].

Figure 4: Snowflake - another example of a dendritic crystal shape [22].

Furthermore, different types of models can be used to describe crystal growth, including *geometric* models for describing interfacial growth velocities unaffected by long-range influences. As for rough surfaces, *non-geometric* models are preferable. [15]

If the surrounding medium is a melt, then a *phase field* model can be used, where a thin interface separates the solid crystal phase from the the liquid phase [23]. Here, the parameter ϕ is called the phase field and changes over the interface from 0 for a bulk solid to 1 for a liquid [23]. However, the atomic picture of crystal growth is disregarded in favour of field values such as temperature and concentration [23]. This model, which also allows dendritic growth to be simulated by incorporating interfacial anisotropy, has acquired popularity over the past years [24].

The important point is that crystals can grow in 3D, 2D and 1D, and the growth mode depends on the crystal system. This is exemplified in CaO (3D growth), C-S-H (2D growth) and raphides and styloids (1D growth), the three systems to be investigated in this study. Now, background theory for these systems will be provided.

1.4 CaO (3D growth)

The first system to be studied in this project is that of CaO, an alkaline earth oxide [25]. This system crystallises in the cubic rock salt structure [25], with a lattice constant of 4.808 Å [26]. Since the CaO system grows isotropically [25], this study seeks to confirm this. This will be done by analysing the energy as a function of the CaO crystal size, which will be done for growth in 3D, 2D and 1D.

1.5 Concrete/C-S-H (2D growth)

1.5.1 General characteristics of concrete

Concrete consists of a mixture of aggregates (sand, gravel or crushed stone), Portland cement and water. When the Portland cement is mixed with water, a chemical reaction known a hydration gives rise to the characteristically hard structure of concrete. During the hydration, the water and Portland cement coat the aggregates. Therefore, the type and size of the aggregates dictate what kind of construction the concrete will be used for. The aggregates usually make up about 60-75 % of the volume, the water 15-20 % and the Portland cement 10-15 %. Furthermore, entrained air in the porous spaces may occupy 5-8 % of the volume. [27]

1.5.2 Preparation and composition of Portland cement

In order to manufacture the Portland cement, raw materials such as limestone and clay are first quarried, since they contain the main components Ca, Si and Al. These rocks are then crushed, mixed with other miscellaneous ingredients and then heated to 1480 ^{o}C in a kiln, forming clinkers. After the clinkers are cooled, they are grinded and mixed with gypsum and limestone, finally forming Portland cement. [28]

The main components in Portland cement are CaO, SiO_2, Al_2O_3 and Fe_2O_3, with typical compositions around 60 weight%, 20 weight%, 4.5 weight% and 3 weight%, respectively [29]. In order to simplify the element description in cement chemistry, the *cement chemist notation* (CCN) is used [18, 30]. For instance, CaO is denoted C, SiO_2 is denoted S, H_2O is denoted H and Al_2O_3 is denoted A [18, 30]. This greatly simplifies the denotions of the main compounds of Portland cement, which are shown in Table 1 [18, 30].

Table 1: Main compounds in Portland cement [18, 30].

Name of compound	**Chemical formula**	**CCN**
Tricalcium silicate	$3CaO \cdot SiO_2$	C_3S
Dicalcium silicate	$2CaO \cdot SiO_2$	C_2S
Tricalcium aluminate	$3CaO \cdot Al_2O_3$	C_3A
Tetracalcium Alumino-ferrite	$4CaO \cdot Al_2O_3 \cdot Fe_2O_3$	C_3AF

1.5.3 Structure of the C-S-H phase

Calcium-silicate-hydrate (C-S-H) makes up the bulk phase of hydrated cement, and it resembles a structure of tightly packed powder grains on the nanometer scale [31]. The hardness of concrete comes from the hydration, when cement powder is mixed with water, and subsequent setting of the C-S-H phase [31]. During the precipitation, C-S-H forms particles on the nanometre scale with polydisperse character, which results in a large variety of local densities [31]. Therefore, the C-S-H phase displays a high degree of disorder [31]. It is known that the nanoparticles display anisotropic growth, forming nanoplatelets with typical dimensions around 50 nm $\times$ 30 nm $\times$ 5 nm (the C-S-H nanoplatelets do not have very circular shapes) [1].

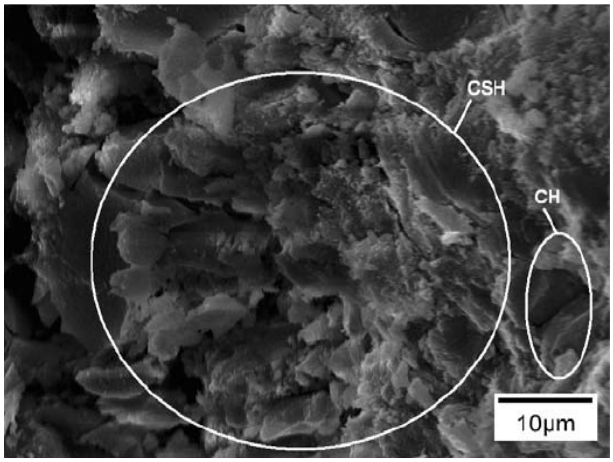


Figure 5: A microscopic view of the C-S-H phase [32].

The disordered atomic structure of C-S-H is not known exactly [33], and it is difficult to study experimentally [33, 34]. Therefore, it is not possible to find a clearly defined unit cell for the C-S-H phase; it cannot be modelled as a periodic structure. Nevertheless, XRD patterns of C-S-H in Portland cement carry close resemblance to tobermorite (C/S ratio $\sim$ 0.8) [33, 34], even though the C/S ratio in hydrated Portland cement is usually around 1.7 [33].

In order to simulate the disordered nature of the C-S-H phase, one study has proposed three different models based on the tobermorite 11 Å structure (Figure 6). All the models start off with the tobermorite 11 Å structure specified by Merlino S et. al. [35], and the following procedures were then done in order to obtain the unit cells for each model:

Model 1: The unit cell is a 3x4x1 supercell of tobermorite 11 Å, where silica monomers were randomly removed from Si2 sites (figure 6). Ca atoms were added into the interlayer in order to obtain a C/S ratio close to 1.67.

Model 2: The unit cell is a 3x5x1 supercell of tobermorite 11 Å, where four consecutive silica chains at the Si2-Si1-Si3-Si2 positions were randomly removed (Figure 6). Ca atoms were then removed in order to adjust the C/S ratio to 1.67.

Model 3: The unit cell is a 3x5x1 supercell of tobermorite 11 Å, where Si atoms were randomly removed irrespective of atomic position.

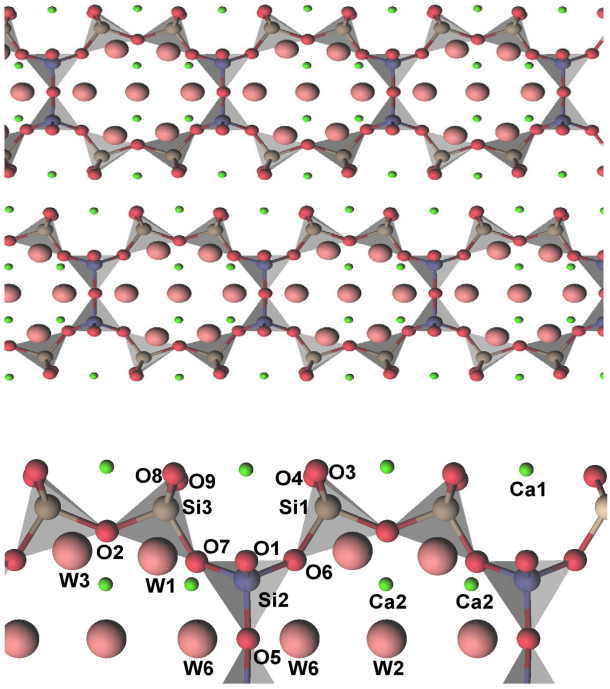


Figure 6: The ideal tobermorite 11 Å structure is shown in the upper picture. In the lower picture, dimeric silicate units (Si1, Si3), bridging silicate units (Si2) and oxygen (O5) etc. have been zoomed in on. W stands for water molecules, and silicon atoms of the bridging silicate units are blue. [33]

All three models are visualised in Figure 7, with the unit cell parameters specified in Table 2 [36]. Structures corresponding to the various models were then subjected to energy optimization with both *molecular mechanics* (MM) and semi-empirical calculations, with ReaxFF used as the force field and *parametric method number 6* (PM6) as the semi-empirical method, respectively. According to these calculations, model 1 was energetically the most stable and its density closest to the experimental value. Additionally, the density variation of model 1 with the H/S ratio corresponded to the experimental trend. Therefore, model 1 will be used in this study. [36]

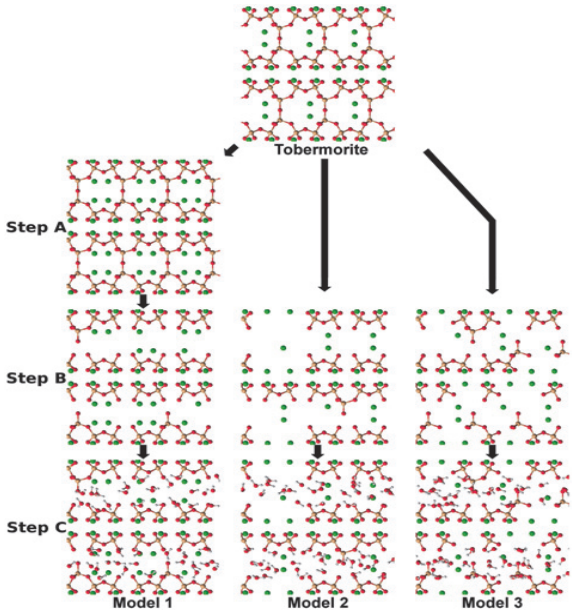


Figure 7: Unit cells from the three C-S-H models based on the tobermorite 11 Åstructure. For clarity, water molecules have been omitted from all unit cell pictures except in step C [36].

Table 2: Parameters for the unoptimized unit cells for model 1, 2 and 3 (Figure 7). The α and γ angles were both 90.0 deg. [36]

	Model 1	Model 2	Model 3
a / Å	29.48	36.84	36.84
b / Å	20.95	19.90	19.90
c / Å	22.78	22.68	22.68
β / deg	58.17	58.17	58.17

1.6 Calcium oxalate monohydrate - raphides and styloids (1D growth)

Crystals of calcium oxalate, CaC_2O_4, may be found in plants, where the purposes are, among others, defence against herbivore, calcium regulation and ion balance [2]. However, they are also found in around 70% of all kidney stones, though the exact mechanism which generates kidney stones is still under debate [37].

The most thermodynamically stable form of calcium oxalate at room temperature is *calcium oxalate monohydrate* (COM, $CaC_2O_4 \cdot H_2O$, whewellite), which forms a monoclinic crystal system. Furthermore, there are two more natural forms: *calcium oxalate dihydrate* (COD, $CaC_2O_4 \cdot (2+x)H_2O$, $x \leq 0.5$, weddellite), which forms a tetragonal crystal system, and *calcium oxalate trihydrate* (COT, $CaC_2O_4 \cdot (3-x)H_2O$, $x < 0.5$, caoxite), which forms a triclinic crystal system, though they will not be subject to studies here. [37]

Calcium oxalate crystals are categorized into five major shapes: crystal sand, raphide, druse, styloid and prismatic [2]. Raphides are made of needle-shaped crystals consisting of whewellite [2]. Hundreds to thousands of these crystals aggregate into bundles, giving rise to raphide crystals, which are found in the vacuoles in plant cells called idioblasts [2]. The raphides, of which there are four common types, can have square, H-shaped, hexagonal or octagonal cross-sections, though the ends are often pointed [2]. They extend along the c-axis, and have diameters around 2-4 μm and lengths around 10-80 μm [5]. Furthermore, the styloids represent another type of elongated crystals extending along the c-axis [2, 5]. Their ends may be pointed or squared [2],

and as an example of their dimensions, styloid COM crystals from *Eichhornia* have displayed lengths in the order of tens of μm [37].

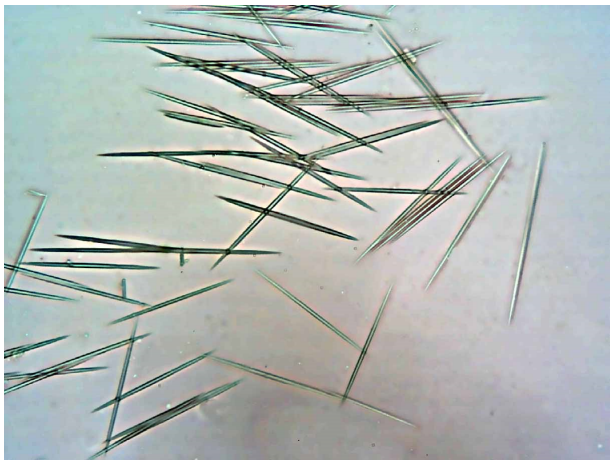


Figure 8: Raphides from liquidized variegated ivy [38].

The essence of the COM structure is two distinct, alternating layers of oxalate ions, which differ in their structural environment. Both types of oxalate ions are coordinated to Ca ions, and the O atoms of one of the oxalate ion types bind to water molecules via hydrogen bonds. While the C-C bonds in one of the oxalate ion layers are almost parallel to the b-axis, the oxalate ions in the other layer lie in the (010) plane running along the c-axis. Consequently, these two layers of oxalate ions are almost perpendicular to each other, and they are stacked along the a-axis [37]. (Figure 9)

The unit cell of COM has parameters $a = 6.250$ Å, $b = 14.471$ Å, $c = 10.114$ Å, $\beta = 109.978^o$, $Z = 8$ and symmetry space group $P2_{1/c}$ [39] (Figure 9). Furthermore, there are three polymorphs of COM, which may be obtained via dehydration or manipulating the temperature, but they will not be subject to studies here [37].

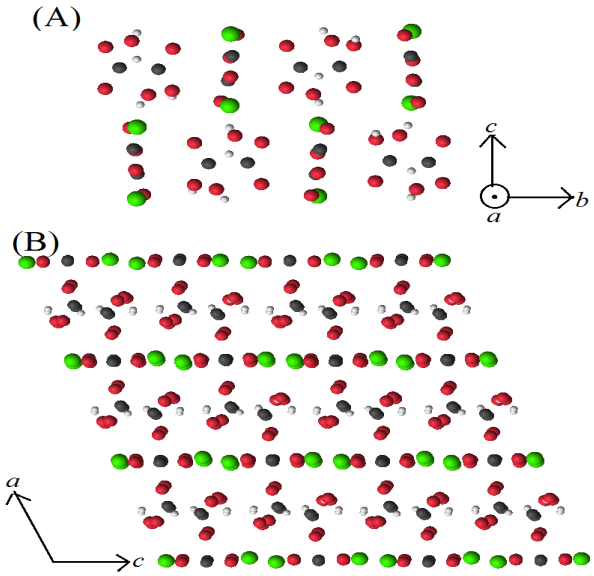


Figure 9: (A) displays the unit cell of COM, and (B) shows the stacking of Ca and oxalate groups along the a-axis [37].

1.7 Outline of the study

As mentioned earlier, this study seeks to model the anisotropic growth of the C-S-H nanoplatelets. For this purpose, the MM method with the $\text{ReaxFF}_{\text{SiO}}$ force field and semi-empirical calculations with the PM6 method will both be used. For a review of the theory behind those methods, please refer to Appendix 5.1-5.3. The computing softwares to be used are LAMMPS and MOPAC, respectively. The performances of the two methods will be compared, but it will only be possible to test smaller system in MOPAC, due to high computational costs for semi-empirical methods. In both cases, the analysis is centred around plotting the total interaction energy as a function of the crystal size. However, the CaO system is first analysed to test the reliability of the $\text{ReaxFF}_{\text{SiO}}$ and PM6 calculations. Furthermore, the analysis of this simple structure will also serve as a reference point for the study of the more complex C-S-H phase.

This study of the C-S-H phase is related to the analysis conducted by Rodríguez MA, where the $\text{ReaxFF}_{\text{SiO}}$ and PM6 methods were used to analyse the growth of C-S-H nanoparticles [18]. Therefore, in this study, as was done in his study, the particles will be modelled as cylinders, and it will be monitored how the energy varies with the system size when only the radius (growth in the xy-plane) or the height (growth in the z-direction) is increased. This study will therefore focus on the energetic driving force. Rodríguez MA concluded that the cylinders displayed a larger gain in energy when only the radius was increased [18]. This was in accordance with the known platelet structure of C-S-H nanoparticles [1]. This study seeks to expand the work by Rodriguez MA by performing the same calculations for larger cylinders. However, it must be mentioned that the cylinder shapes will only be approximative. The reason is that in order to avoid breaking covalent bonds, since the cylinders will be cut out of C-S-H supercells, the geometrical cylinder shape will have to be compromised. This will be explained in detail later.

As the C-S-H cylinder radius and therefore the size increases, so does the significance of size-dependent entropic terms. If these entropic terms could be estimated, it would be possible to combine these with the calculated energies and approximate the optimal sizes of the cylinders. The optimal sizes would correspond to the most stable cylinders from a free energy perspective. However, these entropic terms are difficult to estimate, so this will not be done in this study. Furthermore, it will not be computationally feasible to calculate the energies for system sizes corresponding to real concrete particles, as they have dimensions in the order of tens of nm [1].

It has to be mentioned that this study disregards how the cylinders have been created. Growing crystals in practice as well as modelling it theoretically are both very involved procedures with many different parameters to consider. Therefore, only snapshots of C-S-H cylinders are investigated without any regards to how they have been created. It is simply assumed that the cylinders of various sizes and shapes are stable. This means that the radius/height ratios of the modelled C-S-H cylinders can deviate considerably from the quite flat geometry of real C-S-H nanoplatelets ($\sim$ 50 nm $\times$ 30 nm $\times$ 5 nm) [1]. Furthermore, effects from a surrounding water medium on the cylinder energy as well as dissolution of atoms are not modelled in the study, since the cylinders are considered to be surrounded by vacuum. Entropic effects are also excluded from the calculations, meaning that degradation effects on the C-S-H cylinders are not simulated in this study. Furthermore, the calculations only consider the thermodynamic picture. Kinetic effects are not included, meaning that for instance dendritic growth effects, where the growth shape is dictated by the kinetics [9], cannot be simulated. These restrictions apply to both the force field and semi-empirical calculations in the study.

Fortunately, the neglect of water effects on the energy is not a significant problem. Rodríguez MA did PM6 calculations on C-S-H cylinders in water, where the water influence was simulated via an external potential called COSMO. While the inclusion of COSMO resulted in lower en-

ergies, this energy reduction was very small. Therefore, the calculated C-S-H cylinder energies in this study should not be affected significantly by the calculations being done in vacuum. [18]

In reality, the C-S-H platelets in the hydrated cement paste have high surface charge densities, which is due to titrating silanol groups [40]. But these charges will not be included in this study, because the C-S-H cylinders are considered to be surrounded by vacuum in our calculations, and the real-life process through which concrete is made is not modelled. Furthermore, it will be simpler to perform the calculations on electro-neutral C-S-H cylinders. Furthermore, this study also seeks to analyse the observed anisotropy in the growth of needle-shaped COM crystals, the raphides and styloids. The energy as a function of crystal size will be compared for COM crystals undergoing 3D growth and COM crystals only expanding along the c-axis (1D growth). Therefore, only the energetic contribution to crystal growth will be investigated here too. If elongated shapes are more favourable for COM crystals from an energetic point of view, the 1D growth should result in a more pronounced decrease in energy as the crystal size increases. Also, the earlier explained constraints, such as the neglect of water and kinetic effects, apply to COM crystals as well.

In theory, the crystal systems represented by CaO, C-S-H and COM are infinite. However, in this theoretical study, only clusters of limited size will be analysed. Through the results for these clusters, growth trends for the infinite systems will be investigated.

2 Creation of clusters and computational setup

2.1 CaO system

2.1.1 Creation of supercells

The first part of the study was to create supercells of CaO, based on the unit cell of CaO. The unit cell consisted of a cubic rock salt structure of four atoms each of Ca and O, with the atoms located on each corner of a cube with edge 2.404 Å. A script was written which took the coordinates of the eight atoms of the unit cell and translated them into a supercell using the lattice vectors (4.808 Å, 0, 0), (0, 4.808 Å, 0) and (0, 0, 4.808 Å).

Using this technique, 3D cubic supercells were created for studying the 3D growth of CaO. This was done by simply increasing the number of unit cells in each of the x-,y- and z-axes by one for every step. Therefore, a series of CaO cubes with increasing sides were obtained.

As for studying the 2D growth, only the number of unit cells along the x- and y-axes was increased in steps of 1, while keeping the thickness in the z-axis at 1 unit cell. Therefore, a series of quadratic plates with expanding areas in the xy-plane were studied when analysing the 2D growth. However, due to the very low thickness of the 2D plates, 2.4 Å, all the atoms were at the surface. This introduced considerable surface effects to all the atoms, but the low thickness was nevertheless chosen in order to reduce the number of atoms for the computationally demanding semi-empirical calculations. Furthermore, a comprised thermodynamic stability of the CaO system was not deemed important; CaO was only intended to be a simple modelling system, primarily for testing the accuracies of $\text{ReaxFF}_{\text{SiO}}$ and PM6.

As for the 1D growth, a script was created which allowed a cylinder to be carved out of a 3D supercell, with both the radius and height as adjustable parameters. Using the cylinder shape instead of just a single row of unit cells provided the crystals with more thermodynamic stability; the surface effects were reduced. Then, while keeping the radius at 5 Å, cylinders with various heights were created in order to analyse the 1D growth of CaO crystals. A radius of 5

Å was deemed large enough to resemble some approximate cylinder shape while still keeping the number of atoms down for the semi-empirical calculations.

When limiting the sizes of the cube or plate supercells or cutting the cylinders, no covalent bonds were cut as CaO is a purely ionic system. Also, the cubes and platelets were both electro-neutral, due to the unit cell having no net charge. Furthermore, the 1D cylinders could be considered to be made up of plates stacked along the z-direction. And at a radius of 5 Å, these plates were electro-neutral in themselves, meaning that the supercell cylinders were also electro-neutral.

2.1.2 ReaxFF$_{\mathbf{SiO}}$ calculations (force field)

All the structures were run with ReaxFF$_{\mathrm{SiO}}$ as the force field. Fixed, non-periodic faces were used for the boundary box, since it was only of interest to study samples of limited size, not bulk conditions. Furthermore, the boundary box was made very large, extending from -300 Å to 300 Å in the x-,y- and z-axes, in order to ensure that no atoms would be lost beyond the boundary faces. Since the CaO structure was already known to be a simple and ordered structure, the energy minimization (geometry optimization) was run without any MD, as any relaxation of the structure was deemed unnecessary. As for detailed information about LAMMPS, please refer to the online documentation [41].

The ReaxFF$_{\mathrm{SiO}}$ calculations were very quick, and a table is provided to highlight how short the computation times were.

Table 3: Wall-clock times for some of the ReaxFF$_{\mathrm{SiO}}$ calculations on CaO, using the LAMMPS software.

Growth mode	Dimensions	Number of atoms	Wall-clock time
1D	r = 5 Å, h = 108 Å	552	6 s
1D	r = 5 Å, h = 120 Å	612	10 s
1D	r = 5 Å, h = 132 Å	672	12 s
2D	Side = 64.8 Å, Thickness = 2.4 Å	1568	1 s
2D	Side = 69.6 Å, Thickness = 2.4 Å	1800	4 s
2D	Side = 74.4 Å, Thickness = 2.4 Å	2048	6 s
3D	Side = 55.2 Å	13824	51 s
3D	Side = 64.8 Å	21952	2 min 24 s
3D	Side = 74.4 Å	32768	3 min 12 s

As can be seen, compounds with thousands of atoms could be calculated within seconds, and for tens of thousands of atoms the time was still within minutes (Table 3). Fast computation times are typical for force field methods [42, 43].

2.1.3 PM6 calculations (semi-empirical)

Due to the computational demands of semi-empirical methods, only small structures with at most hundreds of atoms were run with PM6 [42, 43]. All atomic positions were optimized. For more information about MOPAC2016, the used version of MOPAC, please consult the online documentation [44].

The following keywords were used during the PM6 calculations:

PM6 in order to use PM6 as the method.

XYZ to use Cartesian coordinates.

PULAY to speed up the convergence.

PL to monitor the convergence of the density matrix. [44]

The calculations were done a second time, but with the keyword NOOPT added to prevent geometry optimization of all atoms. The energies obtained with and without geometry optimization were then compared.

Table 4: Wall-clock times for some of the PM6 calculations on CaO, as implemented in MOPAC. Geometry optimization was used.

Growth mode	Dimensions	Number of atoms	Wall-clock time
1D	r = 5 Å, h = 12 Å	72	1 min 34 s
1D	r = 5 Å, h = 36 Å	192	23 min 31 s
1D	r = 5 Å, h = 60 Å	312	1 h 59 min
2D	Side = 16.8 Å, Thickness = 2.4 Å	128	5 min
2D	Side = 26.4 Å, Thickness = 2.4 Å	288	1 h 58 min
2D	Side = 31.2 Å, Thickness = 2.4 Å	392	10 h 30 min
3D	Side = 7.2 Å	64	6 s
3D	Side = 12 Å	216	1 h 26 min
3D	Side = 16.8 Å	512	1 d 3 h

The numbers in Table 4 show how quickly the time taken by PM6 computations increased with the number of atoms, a trend which was expected from theory. For a specific size order, the PM6 computation times were also significantly longer than $\text{ReaxFF}_{\text{SiO}}$ computation times, which was also in accordance with theory. Thus, it was not feasible to perform PM6 calculations on compounds with more than hundreds of atoms, even for a system as simple as CaO. This can be compared with the time taken by the $\text{ReaxFF}_{\text{SiO}}$ calculations on CaO, where even compounds with tens of thousands of atoms could be optimized within a few minutes (Table 3).

It was also of interest to extract some computed charges of Ca atoms from the PM6 calculations without geometry optimization. This was done for Ca atoms both in the middle (6 nearest O neighbours) and at the surface (5 nearest O neighbours) of the cubes representing 3D growth, as well as the middle (6 nearest O neighbours) and the end surfaces (5 nearest O neighbours) of the cylinders representing 1D growth. The purpose was to see how the structural environment influenced the charges.

2.2 C-S-H phase

2.2.1 Creation of cylinders

As explained in the introduction, it was not possible to define a single unit cell for the C-S-H phase, because the structure lacks periodicity. A database of C-S-H 2x2x1 monoclinic unit cells with 9 different compositions, optimized by ReaxFF and PM6, was already available at the start of this project, all of them based on the template provided by model 1, as explained in section 3.7 [36]. Furthermore, each composition had hundreds of samples with different internal structures of the atoms represented, a heterogeneity which reflected the disordered nature of the C-S-H phase. Puck, a program written by Veryazov V and Rodríguez MA, was then used to translate these unit cells into supercells and cut cylinders out of the supercells [18]. The

radius, height and cylinder origin were all adjustable parameters, but the cylinder origin was kept fixed at (1 Å, 0, 0.25 Å). Since the translation vectors varied between the optimized unit cells, the translation vectors cannot be listed in full. Furthermore, when later on producing the energy curves, $\text{ReaxFF}_{\text{SiO}}$ would be run on C-S-H cylinders created from unit cells that had been optimized with ReaxFF, and PM6 would be run on cylinders created from unit cells that had been optimized with PM6.

When cutting out cylinders in Puck, covalent bonds were inevitably cut (Figure 10), so a set of connectivity rules were applied in the program in order to remedy this. Also, charge rules were also applied in order to compensate for any net charges obtained during the cutting process and create electro-neutral cylinders. The detailed working mechanisms of these rules are explained in Appendix 5.4, and a detailed review of the original version of Puck, released in 2016, is provided by its creators [18].

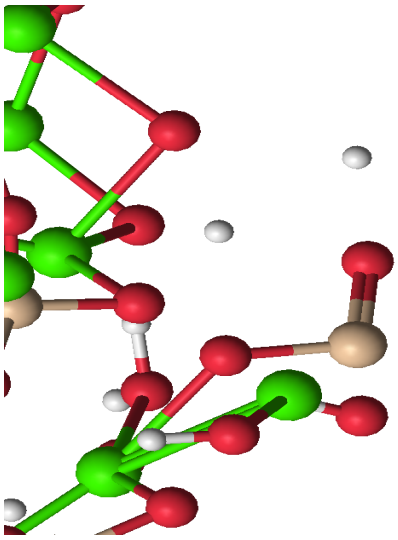


Figure 10: One example of a C-S-H cylinder with broken connectivity is shown; a H atom without any O neighbours may be seen, along with a Si atom with fewer than four O neighbours.

By using a script, Puck was run on the database of C-S-H unit cells, and C-S-H cylinders spanning a wide range of radii and heights were cut out of the supercells, However, Puck, despite the connectivity rules, produced some unwanted cylinders which had Si atoms that were not tetrahedrally coordinated to O, and/or single H atoms. Therefore, a script was used to loop over all the Si atoms in each cylinder, and if any Si atom did not have four O atoms within a sphere of radius 2.0 Å, that cylinder was rejected. Furthermore, the H atoms were also looped over, and if any H atom did not have an O atom within a sphere of radius 1.1 Å, that cylinder was rejected. The limits of these spheres were based on the rules used in Puck when applying the connectivity rules [18].

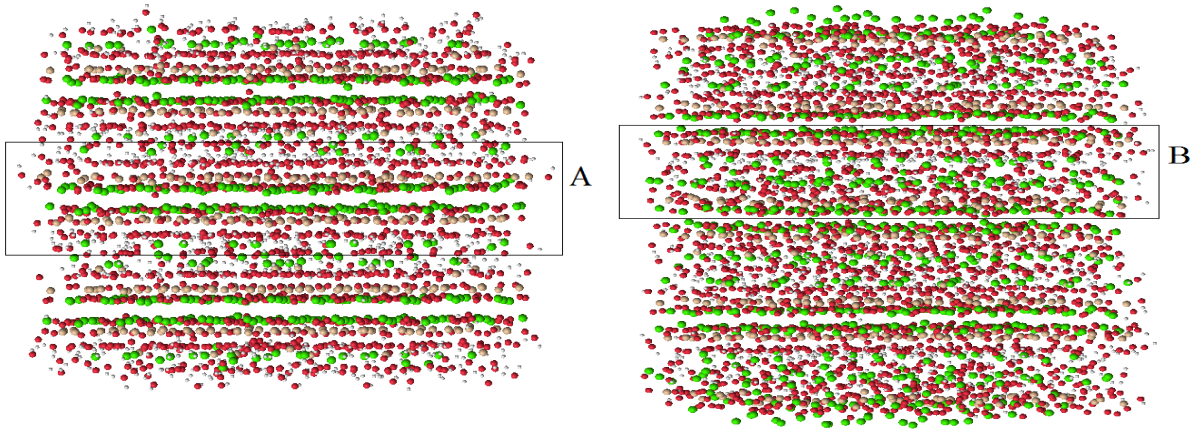
Though it was possible to cut out relatively smooth and geometrical cylinders with defined radii and heights in Puck, atoms were inevitably removed when the connectivity and charge rules were applied. This made it practically impossible to create cylinders that were both geometrical and in agreement with the charge and creativity rules in Puck. Therefore, since the removal of atoms introduced considerable and unpredictable irregularities into the geometries, the increments in the radii and heights of the cylinders to be used in the calculations were not very small.

Furthermore, due to the ordering of the cylinder layers in the z-direction, care had to be taken

when comparing the energies of cylinders with the same radii but different heights. It had to be ensured that the cylinders being compared had the same ordering of layers in the z-direction. As can be seen, the cylinder in Figure 11a had roughly half a interlayer region extending to both the top and bottom surfaces, while the cylinder in Figure 11b had full-length interlayer regions extending to the top and bottom surfaces [33]. Therefore, the energies of these cylinders, which represented the only two types of cylinders that could be obtained without breaking any covalent bonds, could not be compared with each other.

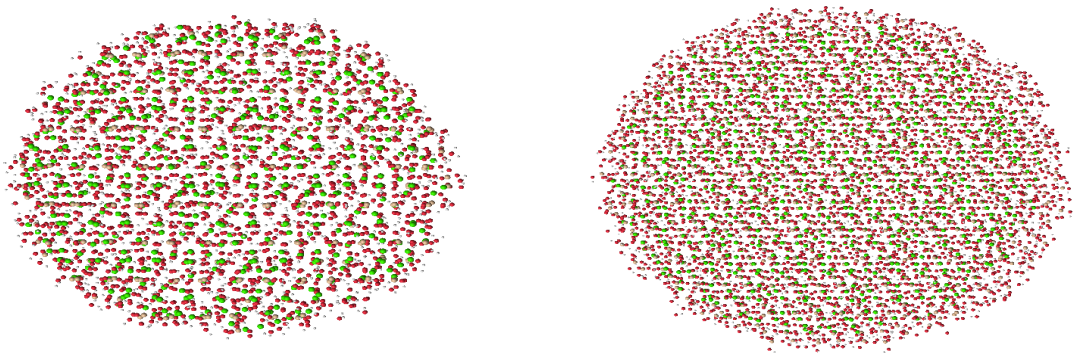
The figures also shows the two different cell types, A and B, where only the former had empty spaces in the middle (Figure 11). Disregarding the atomic disorder of the C-S-H phase, the two cell type were structurally symmetric around the midplane in the z-direction. Cylinders of the type in Figure 11a were built by periodically repeating cell type A along the z-axis, while cylinders of the type in Figure 11b were likewise constructed from cell type B. As can be gathered, a cell type had a height of roughly 12 Å (Figure 11). Therefore, it seemed appropriate to use a step size of 12 Åwhen increasing the height. While this step size deviated from the 11 Å interlayer spacing for the tobermorite model in this study [35, 36], the height of the intralayer region was not well defined. 12 Å was deemed enough to include a complete inter- and intralayer region in every step.

On the other hand, for any value of the radius, the ordering of the structure in the xy-direction was the same (Figure 12). Therefore, the energies of any cylinders with the same height but different radii could be compared freely.



(a) (b)

Figure 11: Comparisons of cylinders with the same radii but different heights, in order to show the differences in ordering of the layered structure in the z-direction. The composition of cylinder (a) is $Si_{404}Ca_{501}O_{1876}H_{1134}$, with dimensions $r = 20$ Å, $h = 36$ Å, whereas the cylinder (b) has composition $Si_{502}Ca_{793}O_{2512}H_{1430}$, with dimensions $r = 20$ Å, $h = 48$ Å. The cell types A and B, from which the cylinders can be considered to be made up of, have also been marked.



(a) (b)

Figure 12: Top view of two C-S-H cylinder structures in the xy-dimension, the radial plane. The composition of cylinder (a) is $Si_{381}Ca_{595}O_{1983}H_{1252}$, with dimensions $r = 25$ Å, $h = 24$ Å, whereas cylinder (b) has composition $Si_{801}Ca_{974}O_{3707}H_{2262}$, with dimensions $r = 35$ Å, $h = 24$ Å. The ordering of the atomic structure as viewed along this direction does not differ much between the cylinders.

2.2.2 ReaxFF$_{SiO}$ calculations (force field)

All the created C-S-H cylinders were run in LAMMPS by using ReaxFF$_{SiO}$ as a force field. The boundary box had the same dimensions as for CaO, extending from -300 Å to 300 Å in all dimensions, and fixed, non-periodic faces were used. Just as for CaO, the study only aimed to study samples of limited sizes, which is why the box was made non-periodic. Furthermore, a very large box would ensure that no atoms would be lost beyond the box faces during the calculations, even for large cylinders. For an example of a LAMMPS script, please refer to Appendix 5.5. It must be stated that since the C-S-H system is a solid, the scheme was designed with the intention of preserving this solid system.

The optimization scheme was as follows:

1. Energy minimization of the entire system except for SiO_2, which were frozen.

2. MD of all atoms for 2.5 ps, except for SiO_2, which were frozen. The canonical ensemble was used for the time integration, and the external temperature was raised from 0.1 K to 500 K. Nose-Hoover thermostat was used.

3. Energy minimization with no frozen atoms.

4. Low-temperature MD of all atoms for 2.5 ps. The canonical ensemble was used for the time integration, and the external temperature was kept constant at 0.1 K. Nose-Hoover thermostat was used.

5. Energy minimization with no frozen atoms.

The purpose of the MD in between the energy minimization steps was to relax the structure towards an energy minimum and therefore aid the energy minimization [43]. During the energy minimization in step 1 and MD in step 2, the SiO_2 were frozen in order to preserve the

main structural elements of the intralayer regions. The unfrozen intralayer Ca atoms should not have been affected too much by the high-temperature MD in step 2, due to the very high melting point of CaO, over 2800 K [25]. However, the interlayer regions were more disordered than the intralayer regions (Figure 11). Therefore, it was considered necessary to explore many different interlayer geometries in order to surpass energy barriers from the energy minimization in step 1. This is why the temperature of the MD in step 2 was set to as high as 500 K; a high-temperature MD was used to significantly alter the interlayer geometry into one more suitable for further energy minimization. Furthermore, the MD in step 4, where no atoms were frozen, was done at a low temperature, since any further pronounced changes of the interlayer geometry at that point were regarded as unnecessary. The low-temperature MD in step 4 also preserved the main structural elements of the more ordered intralayer geometry.

Table 5: Wall-clock time for some of the $\text{ReaxFF}_{\text{SiO}}$ calculations on the C-S-H cylinders created in Puck, using the LAMMPS software.

Dimensions	Number of atoms	Composition	Wall-clock time
$r = 5$ Å, $h = 12$ Å	91	$Si_8Ca_{11}O_{42}H_{30}$	27 s
$r = 5$ Å, $h = 24$ Å	177	$Si_{17}Ca_{24}O_{84}H_{52}$	1 min 7 s
$r = 10$ Å, $h = 12$ Å	368	$Si_{32}Ca_{37}O_{167}H_{132}$	3 min 22 s
$r = 10$ Å, $h = 48$ Å	1325	$Si_{124}Ca_{199}O_{632}H_{370}$	15 min 36 s
$r = 15$ Å, $h = 36$ Å	2210	$Si_{213}Ca_{316}O_{1055}H_{626}$	29 min 12 s
$r = 15$ Å, $h = 60$ Å	4081	$Si_{364}Ca_{443}O_{1872}H_{1402}$	59 min 38 s
$r = 20$ Å, $h = 60$ Å	6554	$Si_{684}Ca_{808}O_{3138}H_{1924}$	1 h 39 min
$r = 25$ Å, $h = 48$ Å	8026	$Si_{812}Ca_{1274}O_{3912}H_{2028}$	2 h 13 min
$r = 30$ Å, $h = 48$ Å	12066	$Si_{1128}Ca_{1590}O_{5680}H_{3668}$	3 h 33 min
$r = 30$ Å, $h = 60$ Å	15493	$Si_{1458}Ca_{2129}O_{7332}H_{4574}$	4 h 33 min
$r = 35$ Å, $h = 60$ Å	19785	$Si_{2056}Ca_{2493}O_{9482}H_{5754}$	5 h 36 min
$r = 40$ Å, $h = 60$ Å	26460	$Si_{2719}Ca_{3369}O_{12662}H_{7710}$	9 h 1 min

The results revealed that $\text{ReaxFF}_{\text{SiO}}$ computation times were longer for C-S-H samples than for CaO samples of the same size order (Table 3, 5). This is due to the C-S-H phase being much more complex than the CaO system, and to more steps being used in the $\text{ReaxFF}_{\text{SiO}}$ routine for C-S-H, where several energy optimization steps were alternated with MD. Nevertheless, samples with hundreds of atoms were calculated within minutes, and samples with thousand of atoms within hours (Table 5). The program displayed a very large amount of converged solutions, even for very large cylinders, and it is clear that the computation time did not increase very quickly with the number of atoms (Table 5). Therefore, LAMMPS was robust enough to be used for large-scale calculations for small to very large C-S-H cylinders.

However, it can clearly be seen here than even very large cylinders were well below the dimensions (tens of nm) of real C-S-H particles [1]. Since real particles therefore probably have hundreds of thousands of atoms, it would require computationally expensive, parallellized MD to perform calculations for theoretical systems correspond to that size. Those calculations were beyond the scope of this study.

2.2.3 PM6 calculations (semi-empirical)

Only smaller crystals were run with PM6, due to considerable computational costs. The same keywords as for the CaO calculations were used here as well. All atomic positions were optimized.

Table 6: Wall-clock time for some of the PM6 calculations on the C-S-H cylinders created in Puck, as implemented in MOPAC. Geometry optimization was used.

Dimensions	Number of atoms	Composition	Wall-clock time
$r = 5$ Å, $h = 12$ Å	100	$Si_9Ca_{11}O_{46}H_{34}$	12 min 34 s
$r = 5$ Å, $h = 24$ Å	171	$Si_{17}Ca_{21}O_{81}H_{52}$	1 h 33 min
$r = 5$ Å, $h = 36$ Å	302	$Si_{28}Ca_{37}O_{141}H_{96}$	8 h 18 min
$r = 5$ Å, $h = 48$ Å	370	$Si_{38}Ca_{47}O_{177}H_{108}$	(Convergence failed)
$r = 10$ Å, $h = 12$ Å	355	$Si_{35}Ca_{44}O_{168}H_{108}$	(Convergence failed)
$r = 10$ Å, $h = 12$ Å	387	$Si_{32}Ca_{42}O_{175}H_{138}$	7 h 41 min

As can be seen, the PM6 computation time increased very quickly with the number of atoms (Table 6), which was expected from semi-empirical methods. By contrast, the scaling of LAMMPS computation times (C-S-H phase) with the number of atoms was much lower (Table 5).

The structural complexity of the system was an important factor which increased the computation time. This is why the computation times for the disordered C-S-H cylinders were longer than for CaO samples of the same size order (Table 4). The structural complexity also explains why some C-S-H structures with sizes of a few hundreds of atoms failed to converge (Table 6). MOPAC did not produce any messages stating the specific reason(s) for convergence failure. To conclude, only small C-S-H cylinders were suitable for PM6 calculations. This can be compared with the $\text{ReaxFF}_{\text{SiO}}$ calculations on C-S-H, where even the solutions for samples with several thousands of atoms converged and the number of converged solutions was higher (Table 5).

2.2.4 Creation of energy curves from database

The Puck runs resulted in a large database of around 500 cylinders with 5 Å $\leq r \leq$ 40 Å and 12 Å $\leq h \leq$ 60 Å. The radii were increased in steps of 5 Å, while the heights were increased in steps of 12 Å, referring to the 12 Å height of the structure cell types A and B (Figure 11). The cylinders with the same radii and heights belonged to the same group, and the cylinders within each group displayed different compositions. Within each subgroup of compositions, various internal structures of the C-S-H phase were represented. $\text{ReaxFF}_{\text{SiO}}$ and PM6 calculations were only run on the most common composition within each respective group of cylinders. If there was not a unique most common structure within a group, a random one of the most common structures was selected. The calculations resulted in a spread of calculated energy values within each group due to the different internal structures. Statistical methods were used to reject results in the tails of the energy distribution within each group, assuming a normal distribution. More details about these statistical methods can be found in the results and discussion section.

When analysing the growth, the total system energy was plotted as a function of either increasing radius (growth in the xy-plane) or height (growth in the z-direction), while keeping the other variable constant. It was analysed whether these two kinds of curves displayed different trends. In his thesis, Rodríguez AM had already shown that the reduction in energy should be more pronounced when the radius increases [18]. Since it was not computationally feasible to calculate the energies of real-life sized C-S-H nanoparticles, the results from the smaller systems in this study were extrapolated.

2.3 COM crystals

2.3.1 Creation of supercells

The procedure was similar to the creation of CaO supercells, though the monoclinic unit cell, which may be seen in Figure 9a, and translation vectors were more complex. The fractional atomic coordinates for one of the four asymmetric units were retrieved from the study by Tazzoli V and Domeneghetti C [45] (Table 7).

Table 7: Fractional atomic coordinates for one of the four asymmetric units of COM at 123 K, though standard deviations have been omitted [39]. All the atomic parameters came from structure refinement of atomic coordinates from Tazzoli V and Domeneghetti C [39, 45].

Atom	x	y	z
C(1)	0.9770	0.32008	0.2493
C(2)	0.9983	0.42690	0.2505
C(3)	0.5147	0.12697	0.1804
C(4)	0.4426	0.11592	0.3132
O(1)	0.9665	0.28174	0.1303
O(2)	0.9898	0.46552	0.1375
O(3)	0.9707	0.28100	0.3537
O(4)	0.9898	0.46559	0.3600
O(5)	0.3573	0.14503	0.0684
O(6)	0.7378	0.12132	0.2051
O(7)	0.2793	0.12102	0.3031
O(8)	0.6025	0.10490	0.4269
Ca(1)	0.96197	0.12372	0.05344
Ca(2)	0.99212	0.12283	0.43515
W(1)	0.3873	0.34435	0.1009
W(2)	0.5850	0.3926	0.3890
H(11)	0.479	0.367	0.049
H(12)	0.388	0.2792	0.091
H(21)	0.473	0.378	0.431
H(22)	0.513	0.378	0.291

By utilizing the translation symmetry of the symmetry space group, $P2_{1/c}$, the asymmetric unit with fractional coordinates as in Table 2.3.1 was transformed into the unit cell. These symmetry operations were:

$(x, y, z) \longrightarrow (-x, -y, -z)$

$(x, y, z) \longrightarrow (-x, y + 1/2, -z + 1/2)$

$(x, y, z) \longrightarrow (x, -y + 1/2, z + 1/2)$

Afterwards, in order to make the unit cell more compact, the value 1 was added to every negative fractional coordinate in the unit cell. It was then translated into supercells by means of the unit cell parameters a =6.250 Å, b =14.471 Å, c =10.114 Å and $\beta = 109.978^o$ (at 123(1) K) [39].

In order to study the 3D growth, a series of supercells were created, each with the same number of unit cells along the a-, b- and c-axes. As for the 1D growth, three sets of needle-shaped structures expanding along the c-axis were created. The first one had a thickness of 1 unit cell

along the a- and b-axes, whereas the other two had a thickness of 2 unit cells along either the a- or b-axis. Just as for CaO, the number of unit cells in the growth direction(s) was increased in steps of 1.

Since the unit cell did not contain any broken covalent bonds, (Figure 9) [37], any unchemical connectivity situations were avoided when creating the supercells. Additionally, due to the unit cell being electro-neutral, the supercells were all electro-neutral as well.

2.3.2 ReaxFF$_{SiO}$ calculations (force field)

The ReaxFF$_{SiO}$ force field was optimized for compounds containing Si and SiO_2 [50], and a force-field is usually not suitable for a set of compounds for which it has not been optimized [43]. Some test calculations showed that this force field was unsuitable for COM crystals, and no other suitable force fields for COM crystals were available. Therefore, only PM6 calculations were used for the COM clusters.

2.3.3 PM6 calculations (semi-empirical)

Once again, only smaller crystals were run with PM6, due to considerable computational costs. Furthermore, the same keywords as for the CaO and C-S-H calculations were used here as well. All atomic positions were optimized.

3 Results and discussion

3.1 CaO system

In this section the results of the LAMMPS and MOPAC calculations are displayed and discussed for the CaO system. The CaO plots concern the 3D growth of the cube, 2D growth of the platelets, and the 1D growth of the cylinders.

3.1.1 Results for ReaxFF$_{SiO}$ calculations (force field)

The unit cell consisted of eight atoms (four Ca and four O) with side 2.4 Å. Therefore, the number of atoms in each cube was calculated as $n^3 * 8$ and the side of the cube as $(2 * n - 1) * 2.4$Å (Figure 13). As for the plates, the number of atoms in each structure was calculated as $n^2 * 8$ and the side in the xy-plane as $(2 * n - 1) * 2.4$ Å, with n being the number of unit cells along each side (Figure 14). As explained earlier, the plates had a thickness of 2.4 Å. As for the cylinders, they could be regarded as being made up of plates stacked along the z-axis. One plate with radius 5 Å consisted of 12 atoms (6 of each kind). Therefore, the number of atoms in each cylinder was calculated as $n * 12$ and the height along the z-axis as $(n - 1) * 2.4$Å, with n being the number of stacked layers (Figure 15).

All three curves displayed the same trend, with an increasing crystal size in terms of volume, xy-area or z-height all leading to a decreased energy (Figure 13-15). However, the asymptote for the 3D growth was at roughly -5.36 eV (Figure 13), while the asymptote for the 2D growth corresponded to a higher energy, roughly -5.18 eV (Figure 14). And the asymptote for the 1D growth was even higher up at approximately -5.01 eV (Figure 15). This showed that is most energetically favourable for CaO crystals to grow in three dimensions, which agrees with the isotropic nature of the cubic rock salt structure. Furthermore, the absolute slope of the energy curve, which reflected the energy gain, decreased as the crystals grew larger. This indicated that it is less and less energetically favourable for a crystal to increase its size as its becomes larger - one possible explanation for why crystals do not grow indefinitely. (Figure 13-15)

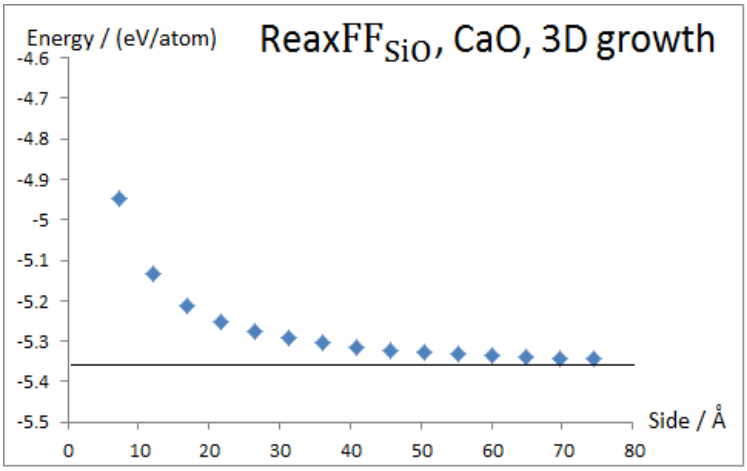


Figure 13: The energies of the CaO cubes representing 3D growth, are displayed as a function of the system size. Asymptote has been drawn.

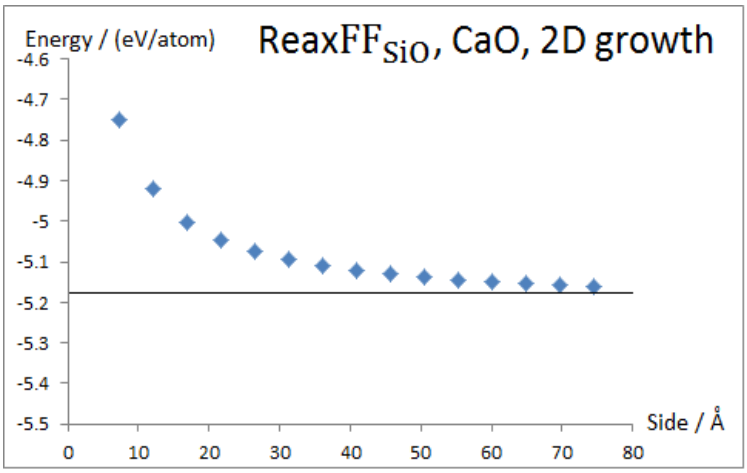


Figure 14: The energies of the CaO plates representing 2D growth are displayed as a function of the system size. Asymptote has been drawn.

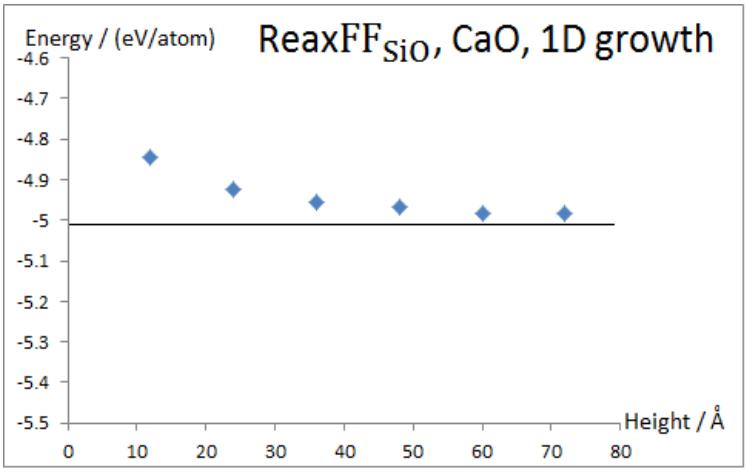


Figure 15: The energies of the CaO cylinders representing 1D growth are displayed as a function of the system size. Asymptote has been drawn.

Even the highest energy obtained in these calculation was remarkably low at around -4.75 eV/atom (Figure 13-15). This would require a temperature of roughly 55000 K, as calculated from the entropic quantity k_bT, in order for the entropy to disintegrate the crystal. Therefore, entropic effects could justifiably be neglected for the CaO crystals, as this temperature was more than nine times the effective temperature of the sun: 5772 K [46]. This suggested that energetic effects in these calculations were far more important than thermodynamic entropic effects for governing the growth of CaO crystals. Taking entropic effects into account would have given a more accurate treatment, but it was beyond the scope of this study and would likely not have yielded any significant differences.

One reason for these very low energies is that the $\text{ReaxFF}_{\text{SiO}}$ scheme in LAMMPS simply minimized the total potential energies of the crystals by iteratively adjusting atom coordinates, without regard to any other factors [41]. Another reason was that the process in which individual atoms are supplied from a surrounding medium to the crystals was not modelled, as the creation process was not part of this study. Therefore, the calculated minimum energies were not compared with some initial stage that the dissociated crystal components would experience in real-life.

3.1.2 Results for PM6 calculations (semi-empirical)

It can be seen in Figures 16-18 that including geometry optimization of the CaO structure resulted in lower energies, which was expected. However, the simple, periodic structure of the CaO system was already perfectly set up by the unit cell parameters used when creating the supercells. So, although geometry optimization allowed further decreases in energy beyond the quantum chemical calculations on the initial structure, it also compromised this ordered structure. This is why the energy curves corresponding to the calculations without geometry optimization were smoother (Figure 16-18).

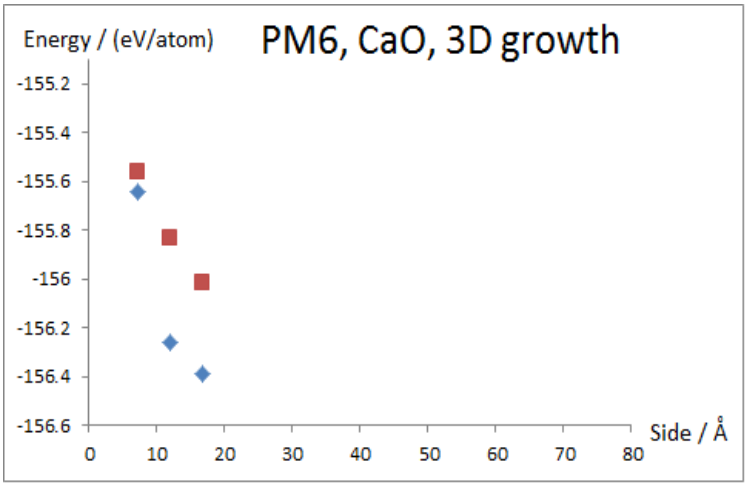


Figure 16: The energies of the CaO cubes representing 3D growth are displayed as a function of the side of the cubes. The *red squares* correspond to the calculations without geometry optimization, whereas the *blue diamonds* correspond to the calculations with geometry optimization. Asymptote has not been drawn, due to a lack of data points.

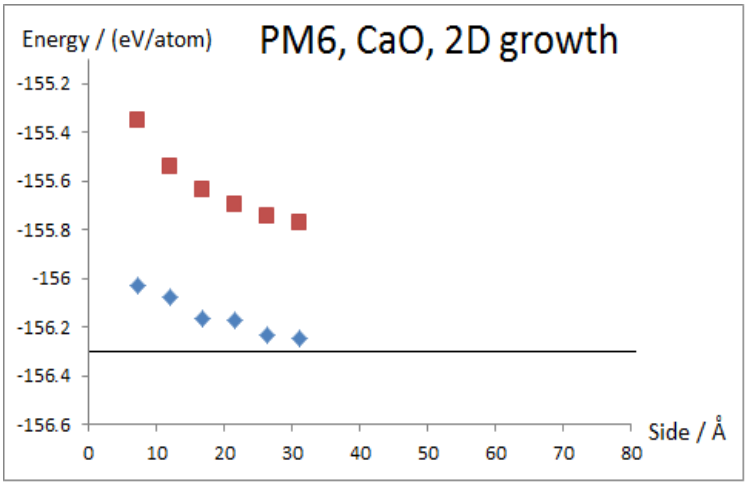


Figure 17: The energies of the CaO plates representing 2D growth are displayed as a function of the side of the plates in the *xy*-plane. The *red squares* correspond to the calculations without geometry optimization, whereas the *blue diamonds* correspond to the calculations with geometry optimization. Asymptote has been drawn for the calculations with geometry optimization.

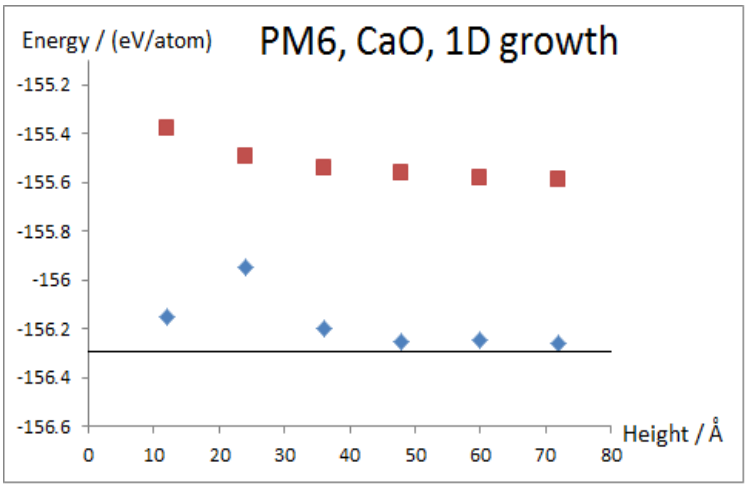


Figure 18: The energies of the CaO cylinders representing 1D growth are displayed as a function of the height of the cylinders along the *z*-axis. The *red squares* correspond to the calculations without geometry optimization, whereas the *blue diamonds* correspond to the calculations with geometry optimization. Asymptote has been drawn for the calculations with geometry optimization.

Due to the limited number of data points for the 3D growth, no asymptote has been drawn for that curve (Figure 16). The asymptotes for the 2D and 1D growth curves (*blue diamonds*; with geometry optimization) both appeared to correspond to energies of around -156.3 eV, it is difficult to determine which asymptote corresponded to the lower energy (Figure 17-18). However, it is clear that the asymptote for the 3D growth curve (*blue diamonds*; with geometry optimization) would correspond to an even lower energy, provided more data points had been obtained (Figure 16). Consequently, the PM6 (semi-empirical) calculations confirmed the observed $\text{ReaxFF}_{\text{SiO}}$ (force field) trend: that it is most energetically favourable for CaO crystals to grow isotropically. Furthermore, the PM6 curves (Figure 13-18), just like the $\text{ReaxFF}_{\text{SiO}}$ results (Figure 13-15), showed that the absolute slopes of the energy curves decreased as the crystals grew larger. Therefore, both methods also showed that the energetic driving force for

crystal growth becomes weaker as the systems size increases.

The absolute values of the energies obtained with the two methods differed significantly, due to different points of zero energy being used in the two methods. But the interesting part in this study here was the trend, not the absolute values. Because the more accurate semi-empirical calculations produced the same trends as the force-field calculations, the methods both appeared reliable. This gave good promise when the methods were later applied on the much more complex C-S-H phase.

The bond in solid CaO is of an ionic nature [25], so it was excepted that the charges of Ca and O should be significantly affected by the structural environment. Analysing atomic charges is a good way to determine if an atom is experiencing almost bulk-like or surface-like conditions, or something in between. Therefore, it was of interest to investigate how the charge of, in this case Ca, varied between atoms found at the surface and in the middle of some of the crystals. Investigating the Ca charge as a function of the crystal size was also of interest.

In figures 19 and 20, plots concerning the comparison of computed Ca charges at the surface and in the middle of the cubes representing 3D growth and cylinders representing 1D growth are shown. The omission of geometry optimization meant that it was feasible to perform the calculations on cubes larger than in the calculations with geometry optimization.

As can be seen in Figures 19 and 20, for both the cubes representing 3D growth and the cylinders representing 1D growth, the charges of the Ca atoms in the middle were higher than those at the surface. This was expected due to the higher number of O neighbours around the Ca atoms in the middle, which consequently increased the electronegative attraction of electrons from the Ca atoms to the O atoms. As the system size increased, the charges also increased for Ca atoms both in the middle and at the surface (Figure 19,20). This was due to the increased number of O neighbours around the Ca atoms.

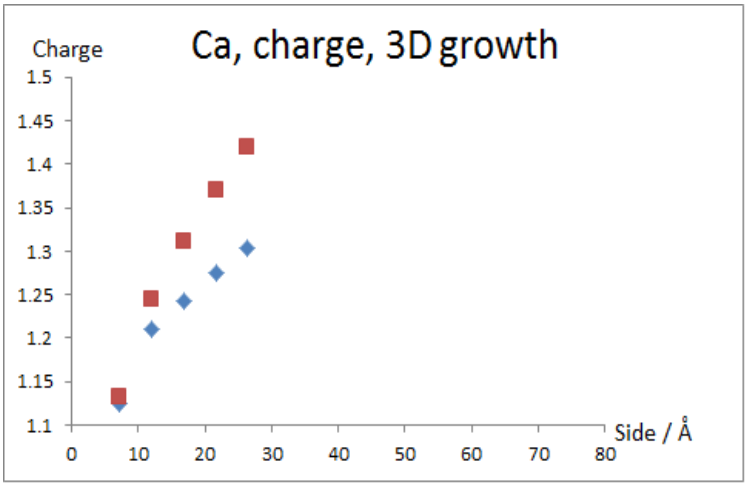


Figure 19: The computed charges of Ca atoms in the cubes representing 3D growth are displayed as a function of the system size. The calculations were done with the PM6 method without any geometry optimization. The *red squares* correspond to Ca atoms in the middle of the cubes (6 nearest O neighbours), while the *blue diamonds* correspond to Ca atoms at the surface of the cubes (5 nearest O neighbours).

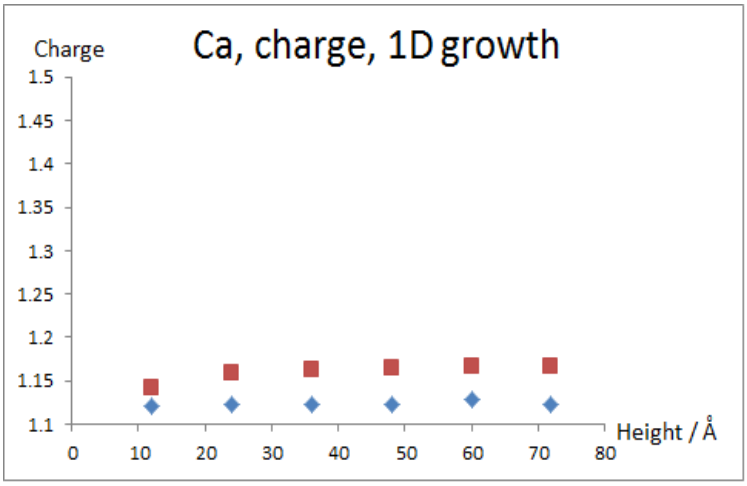


Figure 20: The computed charges of Ca atoms in the cylinders representing 1D growth are displayed as a function of the system size. The calculations were done with the PM6 method without any geometry optimization. The *red squares* correspond to Ca atoms in the middle of the cubes (6 nearest O neighbours), while the *blue diamonds* correspond to Ca atoms at the surface of the cubes (5 nearest O neighbours).

Regarding the cubes, the difference in charges between the Ca atoms in the middle and at the surface increased with the system size (Figure 19). This was explained by that the Ca atoms in the middle were less and less influenced by surface effects as the crystal became larger. As for the cylinders, the differences between the charges of Ca atoms in the middle and at the end surfaces were relatively small (Figure 20). This is because the cylinders were thin, so the atoms in the middle were strongly influenced by surface effects, regardless of the length of the cylinders.

Furthermore, the increase of the Ca charge with crystal size was much more pronounced for the cubes than for the cylinders. This can be explained by an increase in the number of neighbouring O atoms in the cubes, due to expansion in three dimensions, compared to the cylinders, which only expanded in one dimension. Because of computational limitations, the calculations were not performed for even larger cubes, but the shape of the curves suggested that they were staring to level out (Figure 19). This indicated that the number of atoms was eventually so large that the atoms in the middle of the cubes experienced almost pure bulk conditions, while the surface atoms similarly approached pure surface conditions. However, the curves for the cylinder levelled out for relatively small crystal sizes (Figure 20). The explanation was that the cylinders could only approach bulk-like size orders in one dimension, which happened for much smaller geometrical dimensions than for the cubes.

3.2 C-S-H phase

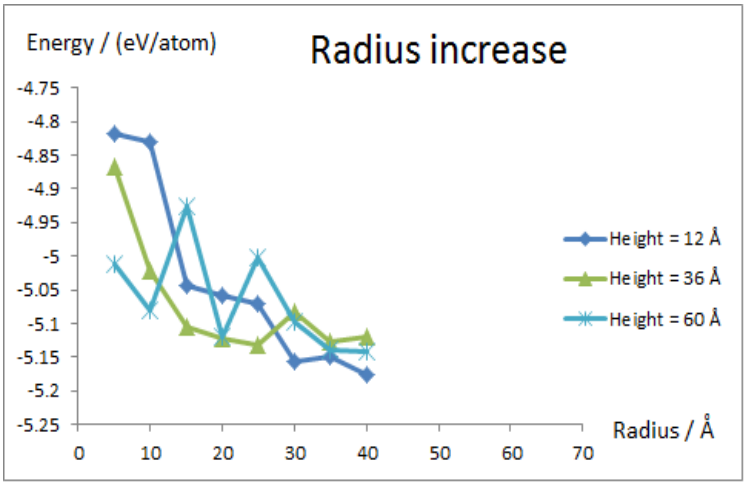
3.2.1 Results for ReaxFF$_{SiO}$ calculations (force field)

Figures 21 and 22 in this section show how the total energy of the C-S-H cylinders changed when only the radius was increased (growth in the xy-plane). This is shown for different values of the cylinder height (growth in the z-direction). It is then followed by figures depicting how the total energy changed when only the height was increased. This is shown for different values of the cylinder radius.

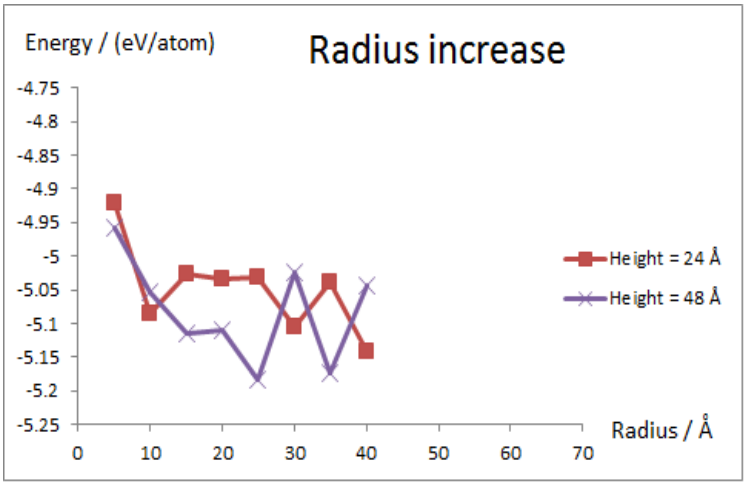
Strictly speaking, the methods used to reject results in the tails of the energy distributions were pseudo-statistical, since only a limited amount of cylinders were found in each group (same radius and height of every C-S-H cylinder within a group). Any results lower than the

first quartile, Q_1, or higher than the third quartile, Q_3, were rejected. Q_1 and Q_3 were computed as the medians of the lower and upper data halves, respectively, with the median of the full data set excluded from both halves, according to method 1 in a paper by Joarder AH and Firozzaman M [47]. For an odd number of elements, the median was the innermost element, and for an even number of elements, the median was the mean of the innermost pair.

Around 100 cylinders were used in the force field calculations. It also has to be mentioned that the amount of C-S-H cylinders within each group varied quite considerably, ranging from around 5-11 for cylinders with a few hundreds of atoms to 1-3 for cylinders with thousands of atoms. This was due to the larger difficulty present in creating large cylinders that satisfied both the charge and connectivity rules. However, since the energy spread for the groups with smaller cylinders was small, this was also considered to be the case for groups with larger cylinders, despite the relative lack of sampling data in the latter groups. An analogy with the tree forest works here: while a few trees are considerably taller than the others, they still intrinsically comply with the same biological and chemical laws as the other trees, if they belong to the same species. Therefore, the limited results obtained for groups with big cylinders were still deemed representative for those groups.



(a)

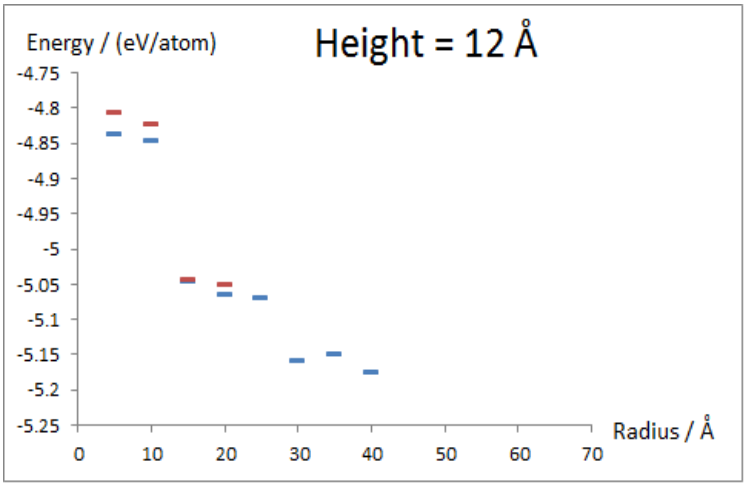


(b)

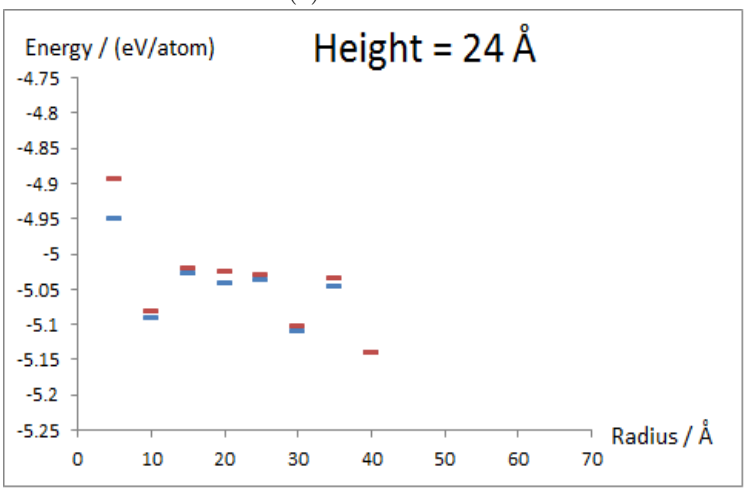
Figure 21: The energies of the C-S-H cylinders are displayed as a function of increasing radius. For each curve, the height is kept constant, and the points represent the median energies. If there is only one energy value within a group, the corresponding point represents that value.

After visualizing some of the cylinders, it was concluded that the 12 Å step length in the z-direction, though deemed suitable in the methods section, was a source of error. Incomplete layers were introduced to the top and bottom surfaces of the C-S-H cylinders with $h = 24$ Å and 48 Å, though they were to some extent limited by the connectivity rules in Puck and energy minimization and MD steps in the ReaxFF$_{\mathrm{SiO}}$ scheme. Nevertheless, these incomplete layers, though relatively small, were to the detriment of the stability of these cylinders, which is why results for $h = 24$ Å and 48 Å are displayed in a separate graph (Figure 21). A more systematic approach when choosing the height increment would have been to investigate how the C-S-H cylinder energy varies with the step length. The optimal step length would then have corresponded to the minimum energy and therefore the most stable cylinder.

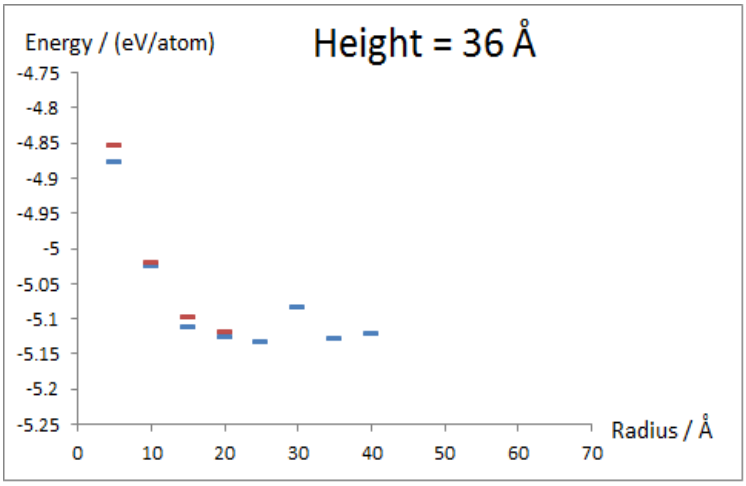
It was not known beforehand how exactly how the cylinder shapes would be affected by the step length. The effects could only be analysed after the time-consuming processes of creating the cylinders in Puck and calculating the energies had been completed. Though the remaining discussion in this section is still valid for cylinders with $h = 24$ Å and 48 Å as well, it must be stressed that the stabilities or these cylinders were decreased by the incomplete layers.



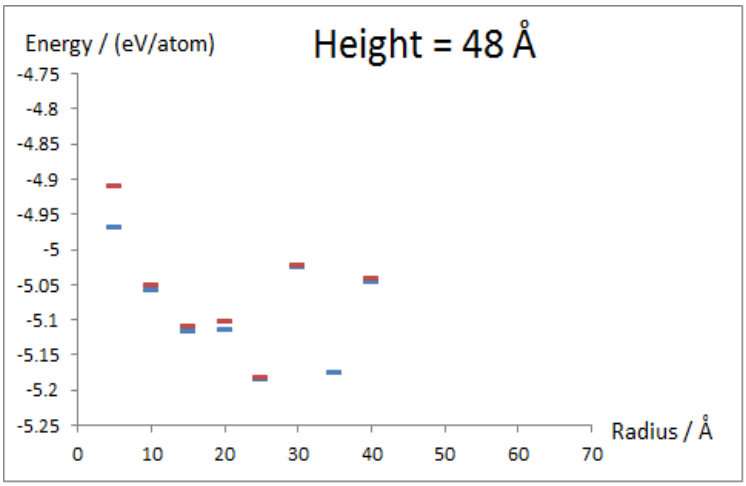
(a)



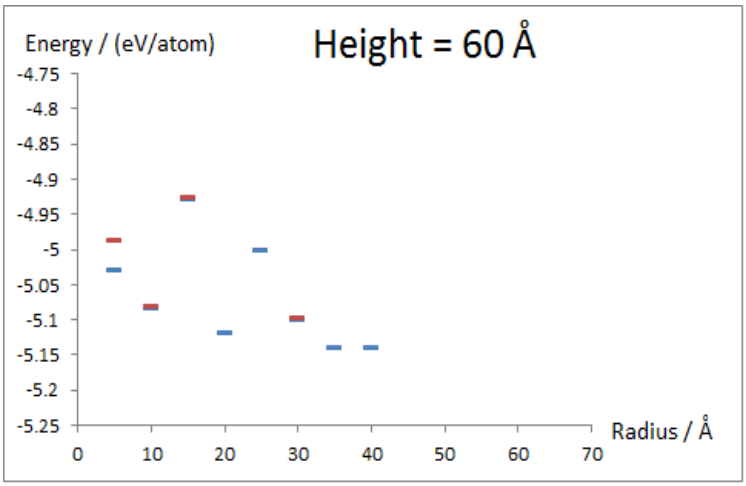
(b)



(c)



(d)



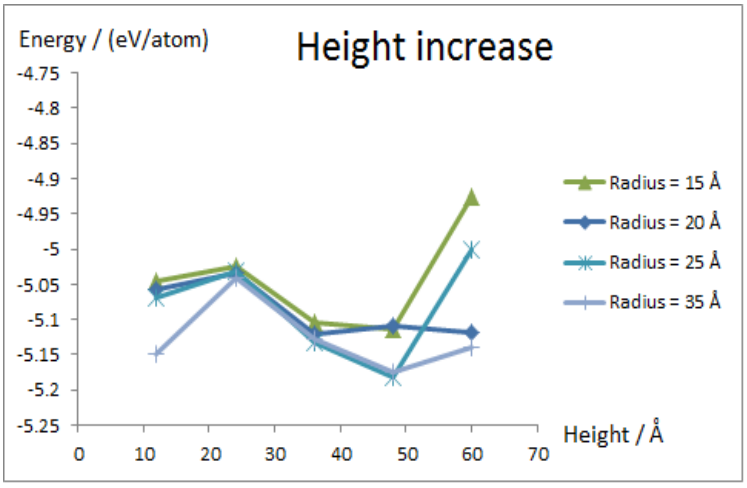
(e)

Figure 22: The energies of the C-S-H cylinders are displayed as a function of increasing radius. Within each figure, the radius is kept constant, and the *red* and *blue* points for each group represent Q_3 and Q_1, respectively. If there is only one energy value within a group, that sole value is shown. The heights are: a) 12 Å, b) 24 Å, c) 36 Å, d) 48 Å and e) 60 Å.

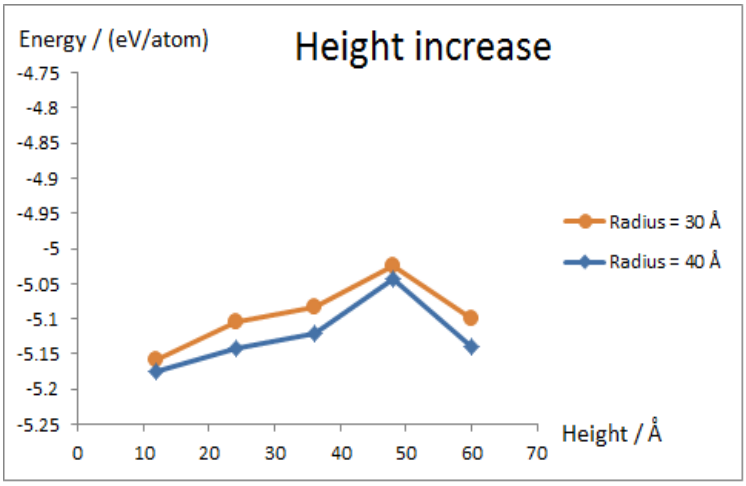
The height values for the figures depicting radial growth (growth in the xy-plane) were almost the same as those used by Rodríguez MA [18]. However, the radii in this study went up to 40 Å for all heights, whereas Rodriguez MA went all the way to 80, 50 and 45 Å for $h = 11.5$, 23 and 34.5 Å, respectively, and only to 35 and 30 Å for $h = 46$ and 57.5 Å, respectively (the step increment of the radii was the same in both studies) [18]. By our own judgement, a radius of 40 Åwas deemed large enough for the atoms in the middle of the cylinders to experience almost bulk-like conditions. Therefore, cylinders up to this size order were considered to represent a large enough size span to display clear trends in the crystal growth.

For all the height values, the overall trend (the energy bumps are discussed towards the end of the ReaxFF$_{\text{SiO}}$ section) was that an increased radius lead to a decease in energy (Figure 24). Furthermore, the slope of the energy curve decreased as the cylinder size increased (Figure 24), which was also observed for the much simpler CaO system (Figure 13-18). Both of these observations were in agreement with the results obtained by Rodriguez MA [18]. However, the ReaxFF$_{\text{SiO}}$ energy curves for the complex C-S-H phase were naturally less smooth than those of CaO. Nevertheless, to conclude, the ReaxFF$_{\text{SiO}}$ results overall showed than an increased radius lead to a significant decrease in energy.

The next step was to analyse the C-S-H cylinder energy as a function of the height. Results for $r = 5$ Å and $r = 10$ Å were omitted, since cylinders that thin experienced more considerable surface effects than cylinders with larger radii. Therefore, the energy curves for $r = 5$ Å and $r = 10$ Å were affected by this, making comparisons with energy curves for cylinders with larger radii complicated. (Figure 23)



(a)



(b)

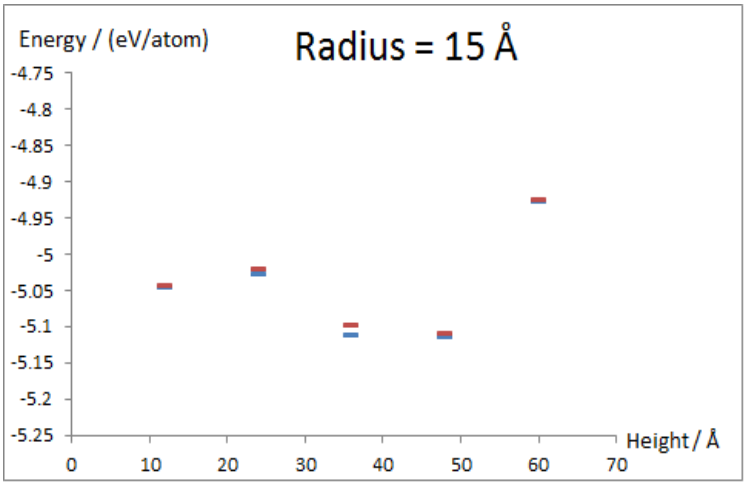
Figure 23: The energies of the C-S-H cylinders are displayed as a function of increasing height. For each graph, the radius is kept constant, and the points represent the median energies. If there is only one energy value within a group, the corresponding point represents that value.

The energy curves have been divided into two groups, based on the overall shapes of the curves, though the correlation between the radius and the curve shape could not be determined (Figure 23). Regardless, no significant energy fluctuations were expected when only increasing the height [18, 1]. Therefore, it was of interest to investigate the C/S ratios of the various cylinder groups to find potential sources of error (Table 8).

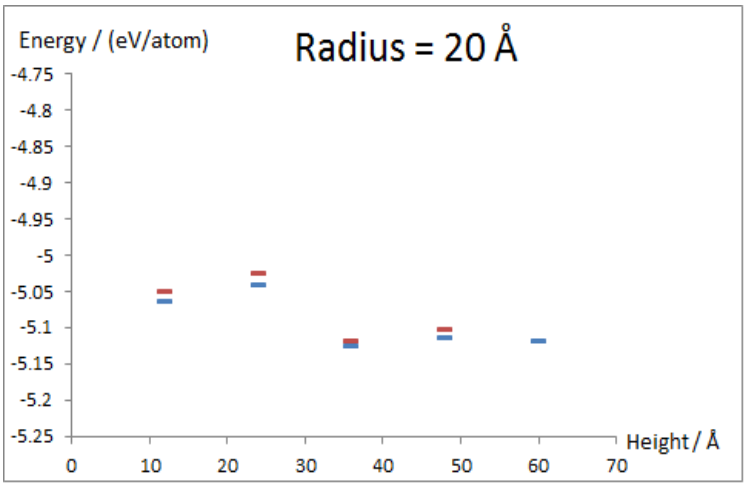
Table 8: Calculated C/S ratios for all the groups of C-S-H cylinders.

r/h	12 Å	24 Å	36 Å	48 Å	60 Å
5 Å	1.38	1.41	1.15	1.41	1.51
10 Å	1.16	1.62	1.41	1.60	1.62
15 Å	1.23	1.18	1.48	1.58	1.22
20 Å	1.31	1.20	1.24	1.58	1.18
25 Å	1.15	1.56	1.21	1.57	1.45
30 Å	1.19	1.20	1.17	1.41	1.46
35 Å	1.19	1.22	1.18	1.39	1.21
40 Å	1.19	1.21	1.20	1.42	1.24

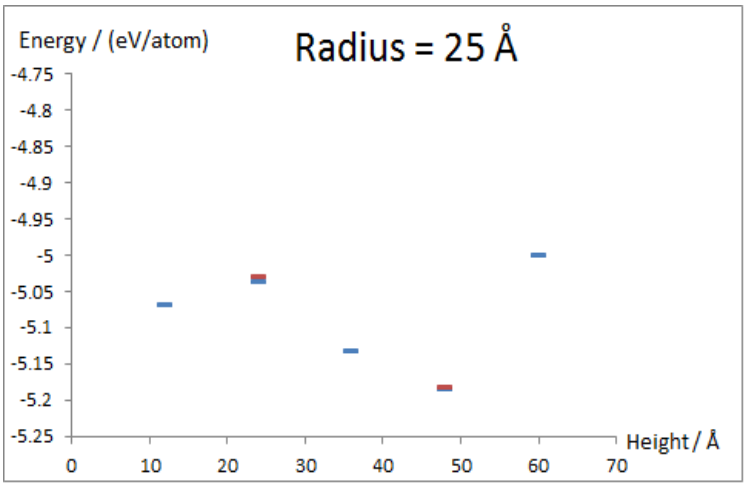
It is evident that the calculated C/S ratios of the C-S-H cylinder groups deviated considerably from the bulk value of 1.67 [36], and they also fluctuated between the groups (Table 8). This was due to that the connectivity rules used by Puck were designed to avoid producing cylinders with broken covalent bonds. Atoms at the cylinder surfaces were removed or added with regards to these connectivity rules only and without any respect to the bulk value of C/S. The fluctuations in the C/S ratio meant that the individual contributions from the atomic species (Si, Ca, O and H) were not equal between the groups, consequently undermining the accuracy of using units of energy/atom. Plotting the energy per atom of one species, for instance Si, might have given smoother curves for not only the energy dependence on the cylinder height (Figure 23), but also the energy dependence on the radius (Figure 21. Furthermore, that the C/S ratios deviated from the bulk value impaired the accuracies of the individual energies within each group. (Table 8)


Energy / (eV/atom)
Radius = 15 Å
-4.75
-4.8
-4.85
-4.9
-4.95
-5
-5.05
-5.1
-5.15
-5.2
-5.25
0
10
20
30
40
50
60
70
Height / Å

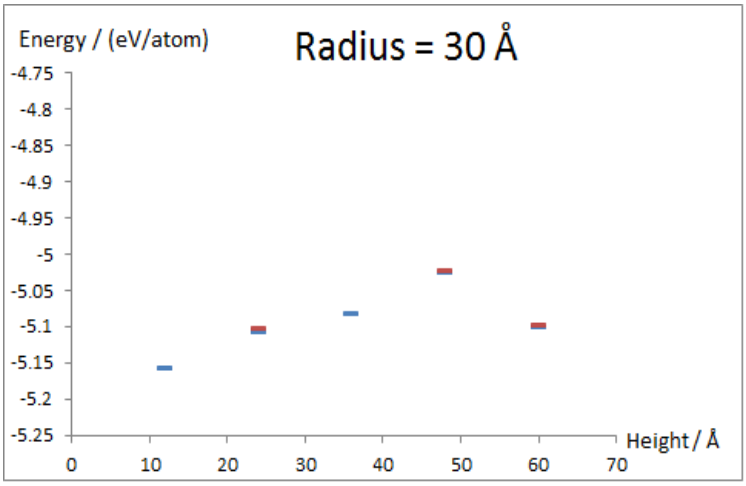
(a)


Energy / (eV/atom)
Radius = 20 Å
-4.75
-4.8
-4.85
-4.9
-4.95
-5
-5.05
-5.1
-5.15
-5.2
-5.25
0
10
20
30
40
50
60
70
Height / Å

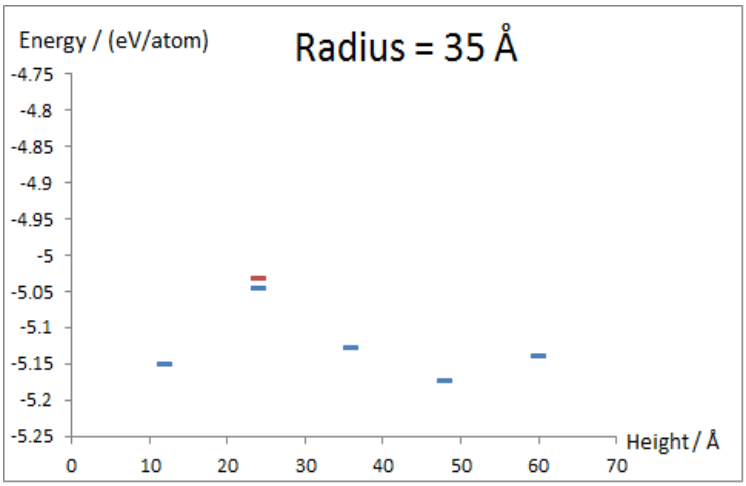
(b)


Energy / (eV/atom)
Radius = 25 Å
-4.75
-4.8
-4.85
-4.9
-4.95
-5
-5.05
-5.1
-5.15
-5.2
-5.25
0
10
20
30
40
50
60
70
Height / Å

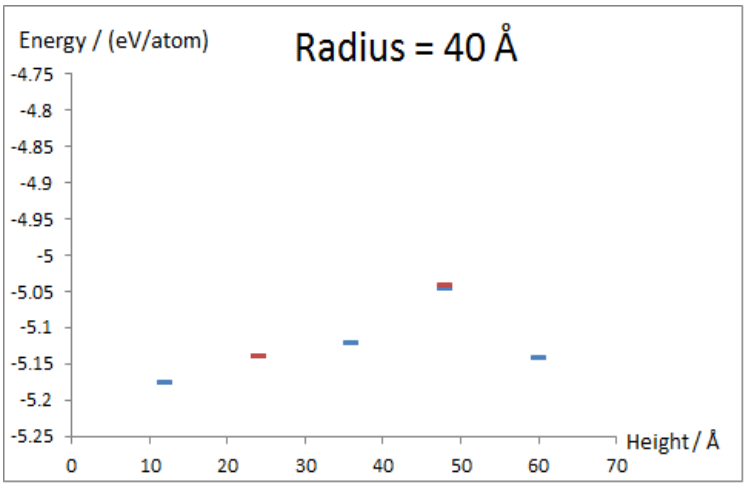
(c)



(d)



(e)



(f)

Figure 24: The energies of the C-S-H cylinders are displayed as a function of increasing height. Within each figure, the radius is kept constant, and the *red* and *blue* points for each group represent Q_3 and Q_1, respectively. If there is only one energy value within a group, that sole value is shown. The radii are: a) 15 Å, c) 20 Å, c) 25 Å, d) 30 Å, e) 35 Å, and f) 40 Å.

As mentioned earlier, the ordering of the layers in the z-direction had to be considered when analysing the change in energy as a function of height (Figure 11). Therefore, for each radii, the cylinders had to be divided into two groups based on the ordering of layers in the z-direction: $h = 12$, 36 and 60 Å, with half a interlayer region extending to both the top and bottom surfaces, and $h = 24$ and 48 Å, with full-length interlayer regions extending to the top and bottom surfaces. Energies could only be compared within these two groups.

After taking the above considerations into account, it can be seen that there was no clear trend depicting a pronounced energy decrease when the height was increased (Figure 24). While there was sometimes an energy decrease when increasing the height, it was not consistent when looking at the different radius values (Figure 24), and often less pronounced than for the curves depicting the energy dependence on the radius (Figure 22).

The $\text{ReaxFF}_{\text{SiO}}$ energy curves for height growth displayed a few bumps indicating local instabilities (Figure 24), which may also be seen in the results from reference [18]. A cylinder with a size corresponding to one of those bumps will collapse to a smaller size equivalent to the closest energy curve valley, which can be prevented if growth is promoted so that the cylinder size surpasses this energy summit. The existence of these bumps for C-S-H but not CaO ($\text{ReaxFF}_{\text{SiO}}$ curves) can be explained by the much more complex nature of the C-S-H phase, which consisted of layers (Figure 11). Where the cut was made in these layers affected the energetic stability. A real-life analogy is found in the construction of tall buildings. During the construction, the building is so unstable that it requires supporting structures in order to not collapse. However, once the construction has progressed far enough, the building is eventually so stable that the support can be removed. Similar reasoning explains why the energy curves for radial growth also contained bumps (Figure 22). Even though the structure in the radial plane (xy-plane) was not layered (Figure 12), the energetic stability of the disordered and complex C-S-H phase could likewise be affected by where the radius cut was made.

Is can be seen from the graphs that the highest energies were around -4.8 eV/atom (Figure 22, 24). In the $\text{ReaxFF}_{\text{SiO}}$ CaO results section, the highest energy, approximately -4.75 eV, was concluded to be much more significant than entropic effects, k_bT. Consequently, the energetic contribution to the growth was the most important one for C-S-H also. Furthermore, the $\text{ReaxFF}_{\text{SiO}}$ results suggested that the C-S-H cylinders have an energetic preference for radial growth. This was also concluded in the previous study [18], and in agreement with the known platelet shape of C-S-H nanoparticles [1]. The next step was to verify this with PM6 calculations.

3.2.2 Results for PM6 calculations (semi-empirical)

Only smaller cylinder with at most a few hundreds of atoms were calculated on, due to the high computational costs for semi-empirical methods (Table 6). Around 30 cylinders were used in the PM6 calculations, but the number of groups was so small that no conclusive trends could be obtained. Therefore, no PM6 results for C-S-H are shown.

3.3 COM crystals

3.3.1 Results for PM6 calculations (semi-empirical)

Again, due to computational demands of semi-empirical methods, the sizes of the studied crystals were limited to hundreds of atoms. These constraints limited the number of of data points for the 3D growth so much that no clear trend could be obtained for that growth mode. However, results for 1D growth are displayed. When it came to needle-shaped structures with a thickness of 2 unit cells along either the *a*- or *b*-axis, only the latter type converged. Therefore, these structures are displayed in the figure depicting 1D growth, together with the needle-like structures with a thickness of 1 unit cell.

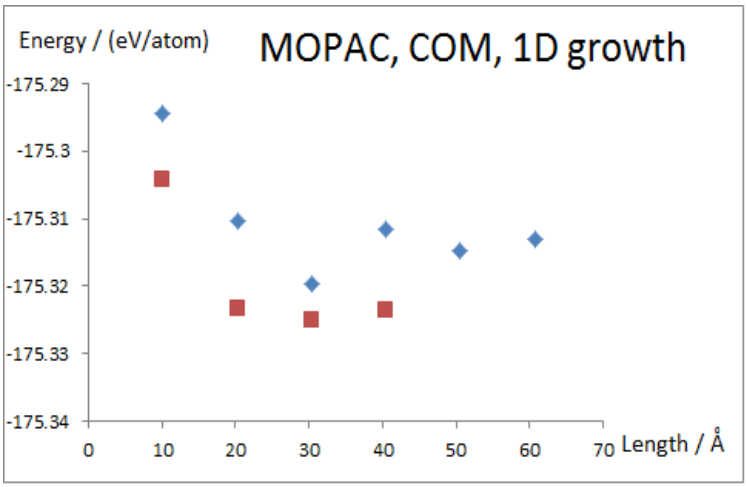


Figure 25: The energies the needle-shaped COM structures depicting 1D growth, as calculated with PM6 without geometry optimization, are displayed as a function of the length along the *c*-axis. The *blue diamonds* correspond to the structures with a thickness of 1 unit cell along both the *a*- and *b*-axes, whereas the *red squares* correspond to the structures with a thickness of 2 unit cells along the *b*-axis.

It can be seen that the 1D growth mode displayed a decrease in energy as the system size increased (Figure 25). Furthermore, increasing the thickness of the needle-like structures lowered the energy, which was expected (Figure 25). As was stated in the introduction, the semi-empirical calculations were done in vacuum, without any regards to water effects. This deviated considerably from reality, as real COM crystals may be found in aqueous environments in the plant cells. Also, kinetic effects were not accounted for; the calculations were thermodynamic in nature. If these effects had been included, the energy curves might have looked different.

Due to the absence of results for 3D growth, it cannot be concluded if growing to isotropic (3D growth) or elongated shapes (1D growth) is more energetically favourable for COM crystals.

4 Conclusions

This project investigated into the driving force for anisotropic growth of crystals, both natural and artificial. The main objective in this project was to confirm the observed anisotropy in the growth of C-S-H nanoplatelets, which make up the bulk phase in concrete. Models of C-S-H clusters with approximate cylinder shapes were created, and the variation of the energy with crystal size when only increasing the radius or height was investigated. The energies were computed using both molecular mechanics, with $\text{ReaxFF}_{\text{SiO}}$ as the force field, and semi-empirical methods, with PM6 as the Hamiltonian.

By using Puck, a database of C-S-H cylinders, with radii and heights ranging from 5 Å to 40 Å and 12 Å to 60 Å, respectively, was created. In total, approximately 100 cylinders were included in the force field calculations, while around 30 were used in the semi-empirical calculations. However, also including cylinders not used in the calculations brings the size of the database up to over 500. This database could potentially be used in the future as a starting point for studies extending to even larger C-S-H cylinders.

The $\text{ReaxFF}_{\text{SiO}}$ calculations showed that the energetic gain was higher when the radius increased (growth in the xy-plane) than when the height increased (growth along the z-axis). However, the computational demands of semi-empirical methods meant that the PM6 calculations could only converge for small C-S-H cylinders. Therefore, too few PM6 data points were produced to confirm the $\text{ReaxFF}_{\text{SiO}}$ trends. Furthermore, due to computational limitations, even the largest C-S-H cylinders in the force field calculations only reached diameters of approximately 8 nm. By contrast, real C-S-H nanoparticles have xy-dimensions of around 50 $\times$ 30 nm.

It was observed that Puck is robust enough for creating C-S-H cylinders with sizes of up to a few tens of thousands of atoms. Unfortunately, since some structures with unsatisfying Si and/or H connectivity were created, a complementary script had to be developed to reject those. So, despite the solid performance of Puck, there is room for improvement in the source code.

The project also briefly analysed the observed anisotropic growth of elongated crystals of COM - raphides and styloids. Since the $\text{ReaxFF}_{\text{SiO}}$ force field was incompatible with COM crystals, only PM6 calculations were used. However, the computational demands of semi-empirical methods meant that too few results for 3D growth could be obtained to conclude if 3D or 1D growth is more energetically favourable for COM crystals.

References

[1] Labbez C, Jönsson B, Woodward C, Nonat A and Delhorme M, Phys Chem Chem Phys, 16, (2014).

[2] Raman V, Horner HT and Khan IA, J Plant Res, 127, (2014).

[3] Cashman KV and Marsh BD, Contributions Mineral Petrol, 99, (1988).

[4] Duval P and Lorius C, Earth Planet Sci Lett, 48, (1980).

[5] Hartl WP, Klapper H, Barbier B, Ensikat HJ, Dronskowski R, Müller P, Ostendorp G, Tye A, Bauer R and Barthlott W, Botany, 85, (2007).

[6] Otálora F and García-Ruiz J, Chem Soc Rev, 43, (2014).

[7] Van Driessche A. Gypsum crystals of the Naica cave [Internet], Copyright (2010) Apr [cited 2019 Jan 29], Link to license: https://creativecommons.org/licenses/by/3.0/legalcode

[8] Rosenberger F and Meehan EJ, J Cryst Growth, 90, (1988).

[9] Jackson KA, Kinetic Processes: crystal growth, diffusion, and phase transformations in materials, John Wiley & Sons; (2006).

[10] Pynn R, Neutron Scattering-A Primer, Los Alamos Science Summer, (1990).

[11] Givand JC, Teja AS and Rousseau RW, AlChe J, 47, (2001).

[12] Sumiya H and Satoh S, Diam Relat Mater, 5, (1996).

[13] Schulz M, Nature, 399, (1999).

[14] Jiao Y, Salce A, Ben W, Jiang F, Ji X, Morey E and Lynch D, JOM, 63, (2011).

[15] Wettlaufer JS, Jackson M and Elbaum M, J Phys A; Math Gen, 27, (1994).

[16] Kotrla M, Comput Phys Commun, 97, (1996).

[17] Glotzer SC and Solomon MJ, Nat Mater, 6, (2007).

[18] Rodríguez MA, Theoretical study about Calcium Silicate Hydrate (C-S-H) anisotropy growth [master's thesis], [Lund]: Lund University; (2016), 47 p.

[19] Harrison JJ. Quartz, Tibet [Internet], Copyright (2009) Feb [cited 2019 Jan 13], Link to license: https://creativecommons.org/licenses/by-sa/4.0/legalcode

[20] Levi AC and Kotrla M, J Phys Condens Matter, 9, (1997).

[21] Paul. Dendritic Silver Crystals - From Batopilas, Chihuahua, Mexico [Internet], Copyright (2011) Feb [cited 2019 Jan 13], Link to license: https://creativecommons.org/licenses/by-sa/4.0/legalcode

[22] A snowflake [Internet], [cited 2019 Jan 28], Link to license: https://creativecommons.org/publicdomain/zero/1.0/legalcode

[23] Miura H, Yokoyama E and Tsukamoto K, Introduction to Phase-Field Model and Its Applications in the Fields of Crystal Growth and Planetary Science, AIP Conf Proc, (2010).

[24] Younsi A and Cartalade A, J Comput Phys, 325, (2016).

[25] Saunderson DH and Peckham GE, J Phys C, 4, (1971).

[26] Roy A and Bhattacharya J, Nanotechnology in Industrial Wastewater Treatment, London: IWA Publishing; 2015.

[27] PCA-The Portland Cement Association-America's Cement Manufacturer, How Concrete is Made, https://www.cement.org/cement-concrete-applications/how-concrete-is-made (Date consulted 2018 Sep 8)

[28] PCA-The Portland Cement Association-America's Cement Manufacturer, How Cement is Made, https://www.cement.org/cement-concrete-applications/how-cement-is-made (Date consulted 2018 Sep 8)

[29] Poppe AM and De Schutter G, Cem Concr Res, 35, (2005).

[30] Thomas J and Jennings H, The Science of Concrete, http://iti.northwestern.edu/ cement/monograph/Monograph3_2.html (Date consulted 2010 Aug 14)

[31] Masoero E, Del Gado E, Pellenq RJM, Ulm FJ and Yip S, Phys Rev Lett, 109, (2012).

[32] Belkowitz JS, The Investigation of Nano Silica in the Cement Hydration Pro [master's thesis], [Denver]: University of Denver; (2009), 42 p.

[33] Kovačević G, Persson B, Nicoleau L, Nonat A and Veryazov V, Cem Concr Res, 67, (2015).

[34] Renaudin G, Russias J, Leroux F and Frizon F, J Solid State Chem, 182, (2009).

[35] Merlino S, Bonaccorsi E and Armbruster T, Eur J Mineral, 13, (2001).

[36] Kovačević G, Nicoleau L, Nonat A and Veryazov V, Z Phys Chem, 230, (2016).

[37] Thomas A, Biomimetic Growth and Morphology Control of Calcium Oxalates [dissertation], [Dresden]: TU Dresden; 2009. 185 p.

[38] Agong1. Raphides from liquidized variegated ivy [Internet], Copyright (2010) July [cited 2019 Jan 13], Link to license: https://creativecommons.org/ publicdomain/zero/1.0/legalcode

[39] Echigo T, Kimata M, Kyono A, Shimuzi M and Hatta T, Mineral Mag, 69, (2005).

[40] Delhorme M, Labbez C, Turesson M, Lesniewska E, Woodward CE and Jönsson B, Langmuir, 32, (2016).

[41] Plimpton S, J Comput Phys, 117, (1995).

[42] Atkins P and Friedman R. Molecular Quantum Mechanics. New York: Oxford University Press Inc; (2011).

[43] Lewars EG, Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics, Dodrecht: Springer; (2016).

[44] Stewart J, MOPAC2016, Version 17.361L, http://openmopac.net (Date Consulted 2018 Sep 22)

[45] Tazzoli V and Domeneghetti C, Am Mineral, 65, (1980).

[46] Williams DR, Sun fact Sheet, https://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html (Date consulted 2018 Feb 23)

[47] Joarder AH and Firozzaman M, Teach Stat, 23, (2008).

[48] Stewart JJP, J Mol Model, 13, (2007).

[49] Thiel W, Wiley Interdiscip Rev Comput Mol, 4, (2013).

[50] Van Duin ACT, Strachan A, Stewman S, Zhang Q, Xu X and Goddard WA, J Phys Chem A, 107, (2003).

[51] Parrinello M and Rahman A, J Appl Phys, 52, (1981).

5 Appendix

5.1 Molecular quantum mechanics

5.1.1 Ab initio methods

One category of molecular quantum mechanics calculations is *ab initio* methods. These are based on solving the non-relativistic, time-independent Schrödiger equation, making use of fundamental constants for a theoretically pure approach. One example of ab initio methods is the *Hartree-Fock self-consistent field* (HF-SCF) method, in which the starting point is the HF equations for closed-shell species with doubly occupied wavefunctions:

$$f_1\psi_m(1) = \epsilon_m\psi_m(1) \tag{2}$$

where ϵ_m are the orbital energy eigenvalues, ψ_m are the spatial wavefunctions and f_1 is the *Fock operator*:

$$f_1 = h_1 + \sum_{m'}\left[2J_{m'}(1) - K_{m'}(1)\right] \tag{3}$$

Here, h_1 is the one-electron Hamiltonian (hydrogenic Hamiltonian for a single electron and a nucleus with charge Ze) and $J_{m'}$ and $K_{m'}$ represent the Coulomb and exchange operators, respectively:

$$J_{m'}(1)\psi_m(1) = j_o\int \psi^*_{m'}(2)\frac{1}{r_{12}}\psi_m(1)\psi_{m'}(2)d\tau_2 \tag{4}$$

$$K_{m'}(1)\psi_m(1) = j_o\int \psi^*_{m'}(2)\frac{1}{r_{12}}\psi_{m'}(1)\psi_m(2)d\tau_2 \tag{5}$$

where $j_o = e^2/4\pi\epsilon_o$. The Coulomb operator represents the Coulomb repulsion between electron 1 and 2, and the exchange operator includes the effects of spin correlation. [42]

One important point is that each electron is considered to be affected by the average field of all the other electrons, so *electron correlation* is neglected. Consequently, instantaneous electron-electron interactions as well as quantum mechanical effects on the electron distribution are disregarded. The *Born-Oppenheimer approximation* is also used, meaning that the electrons are considered to be able to respond immediately to changes in the positions of the nuclei, due to the electron mass being much lower than the mass of the nucleus. Nevertheless, one procedure is then to write each spatial wave unction ψ_m as a linear combination of basis functions χ_o, which for instance be approximated as atomic orbitals centred on each of the N_o atoms:

$$\psi_m = \sum_{o=1}^{N_o} c_{om}\chi_o \tag{6}$$

with c_{om} representing expansion coefficients. After inserting this expansion into the HF equations (2) and performing some operations, it is convenient to introduce the *overlap matrix*, $\mathbf{S}$, and the *Fock matrix*, $\mathbf{F}$:

$$S_{o'o} = \int \chi^*_{o'}(1)\chi_o(1)d\tau_1 \tag{7}$$

$$F_{o'o} = \int \chi^*_{o'}(1)f_1\chi_o(1)d\tau_1 \tag{8}$$

This results in the *Roothaan equations*:

$$\mathbf{Fc} = \mathbf{eSc} \tag{9}$$

where $\mathbf{c}$ and e are $N_o \times N_o$ matrices of the expansion coefficients and orbital energies ϵ_m, respectively, the latter of which is a diagonal matrix. The equation system (9) has a non-trivial solution only if

$$\det|\mathbf{F} - \mathbf{eS}| = 0. \tag{10}$$

However, the Roothaan equations (10) must be solved in a self-consistent field way by first choosing basis functions and trial expansion coefficients and then calculating the overlap and Fock matrix elements, (7) and (8). From (10) a new set of expansion coefficients and orbital energies is then obtained, which are used to compute new matrix elements and obtain another set of expansion coefficients and orbital energies. This cycle of computations is then repeated until the change of c_{om} and ϵ_m from cycle to cycle is not very large, i.e. until convergence has been achieved. [42]

5.1.2 Semi-empirical methods

However, ab initio method, consume a considerable amount of computing power [42]. Therefore, for more than a few dozen atoms, they may be replaced by *semi-empirical* methods [42]. Here, the Hamiltonian is simplified and adjustable parameters taken from experimental data are also introduced, which saves computing power at the expense of accuracy [42]. Furthermore, the computations only involve the valence electrons, each of which is considered to be influenced by atomic cores consisting of the nuclei and other electrons [43].

In the previous section, it was explained that the basis functions can be approximated as, for example, atomic orbitals centred on the atoms. One type of functions fulfilling this is *Slater type orbitals* (STOs), which for an orbital with quantum numbers n ,l and m_l looks like:

$$\Psi_{n,l,m_l}(r,\theta,\varphi) = Nr^{n_{\text{eff}}-1}\mathrm{e}^{-Z_{\text{eff}}\rho/n_{\text{eff}}}Y_{l,m_l}(\theta,\varphi) \tag{11}$$

with N being a normalization constant, $Y_{l,m_l}(\theta,\varphi)$ a spherical harmonic and $\rho = r/a_o$ (a_o is the Bohr radius). Z_{eff} and n_{eff} represent the effective nuclear charges and principal quantum numbers, respectively, and may be retrieved from a table. [42]

In semi-empirical methods, the Roothaan equations are again considered, but the overlap matrix $\mathbf{S}$ is set to the unity matrix [43]. Then, the core Hamiltonian for valence electron i is introduced:

$$h_i^v = -\frac{\hbar^2}{2m_e}\nabla_i^2 + V_i^{v,eff}, \tag{12}$$

where $V_i^{v,eff}$ is the effective potential field experienced by valence electron i from the nuclei and the inner-shell electrons. This gives for the Fock matrix elements (8) of a closed-shell molecule with N_v valence electrons:

$$F_{o'o}^v = h_{o'o}^v + \sum_{q,r} P_{qr}\{(o'q|or) - \frac{1}{2}(o'q|ro)\} \tag{13}$$

where

$$h_{o'o}^v = \int \chi_{o'}^*(1)h_1^v\chi_o(1)d\tau_1 \tag{14}$$

and

$$P_{qr} = 2\sum_m c_{qm}^* c_{rm}. \tag{15}$$

Here, the simplified notation $(ab|cd)$ is used for the two-electron integrals:

$$(ab|cd) = j_o \int \chi_a^*(1)\chi_b^*(2)\frac{1}{r_{12}}\chi_c(1)\chi_d(2)d\tau_1 d\tau_2. \tag{16}$$

[42]

Since the calculation of two-electron integrals (16), especially three- and four-centre ones, consume a considerable amount of computing time, some extent of differential overlap is neglected in order to reduce the number of integrals [43]. In the *neglect of diatomic differential overlap* (NDDO) approximation, $\chi_a^*(1)\chi_c(1)$ is neglected if the basis functions χ_a and χ_c correspond to different atoms, i.e. all one-centre two-electron integrals of the form (16) are retained [42]. Furthermore, NDDO makes use of the earlier introduced STOs as basis functions [43].

One example of a semi-empirical method that can be used in these calculations is *parametric method number 6* (PM6) which is one of the the latest developments in NDDO [48]. This method, which makes use of diatomic properties, offers better accuracy than previous NDDO-types, such as AM1 and PM3, and also allows for the parametrization of up to 70 elements [48]. Also, it treats hydrogen bonds better and yields better heats of formation than AM1 and PM3 [43]. When it comes to semi-empirical methods, there are three sources of error: insufficient or inaccurate reference data, poor assumptions made during approximations or insufficiently optimized parameters [48]. In the case of PM6, data from over 9000 compounds was used for the parametrization [43].

Like ab initio methods, semi-empirical methods make use of several basis functions for each atom [42]. Furthermore, the scaling is normally $O(N^3)$, where N is the number of basis functions [49]. Conventional semi-empirical methods are only recommended for compounds with up to 1000 non-hydrogen atoms [49], and it is not computationally feasible to perform semi-empirical calculations on very large compounds.

5.1.3 Density functional theory

A third alternative, apart from ab intio and semi-empirical methods, is to apply *density functional theory* (DFT). This is fundamentally based on two theorems. The first one is the *Hohenberg-Kohn existence theorem*, which states that the ground-state energy is uniquely determined by the electron density. The second one is the *Hohenberg-Kohn variational theorem*, which states that the energy functional $E_o[\rho']$ cannot be less that the molecular ground-state energy ($\rho'(r)$ is the trial electron density function). In order to minimize the energy functional, the *Kohn-Sham equations* must be solved, where also electron correlation is included. [42, 43].

The inclusion of electron correlation represents one advantage over ab initio and semi-empirical methods [42, 43]. The electron density is also unbound by the permutation symmetry considerations that limit the acceptable choices of wavefunctions [42]. However, the exchange-correlation functional $E_{XC}[\rho]$ (which represents exchange and correlation energy) is not known exactly, and there are many proposed forms for this functional [42, 43]. Furthermore, DFT calculations are much slower than semi-empirical calculations [49].

5.2 Molecular mechanics (force field)

For large systems, it is not computationally feasible to use quantum mechanical calculations. Instead, it is worth considering using *molecular mechanics* (MM) methods, in which each atom with associated nucleus and electrons is treated as a single particle. The interatomic interactions are approximated via a spring-like system, but since electronic effects are disregarded,

bond-breaking and bond-forming, etc. cannot be simulated. Also, the set of parameters and equations used to describe the potential energy is referred to as a *force field* [42]

There are various force fields available, including for instance **AMBER** and **CHARMM**, which are used for proteins and nucleic acids. The force fields are usually reliable only for the specific set of compounds that they are parametrized for, and transferring parameters between force-fields is not recommended [43]. Regardless of the type of force field being used, there are typical expressions for the potential energies. First of all, there is the potential energy due to stretching of the bonds:

$$E_{str} = \sum \frac{1}{2} k_b (r - r_0)^2 \tag{17}$$

where the sum is over all bonds in the molecule and k_b and r_0 are the empirical force constants and equilibrium bond lengths, respectively, for each bond. The zero-order term represents the potential energy at the equilibrium bond length, which may be set to zero for convenience. Furthermore, the first-order term is zero due to that potential energy representing an energy minimum, a stationary point. [42]

Another potential energy takes into account the bending of the bonds:

$$E_{bend} = \sum \frac{1}{2} k_\Theta (\Theta - \Theta_0)^2 \tag{18}$$

where k_Θ and Θ_0 are empirical constants and equilibrium angles, respectively, for every type of angle, with the sum extending over all bonds. A third expression takes into account the potential energy associated with the torsional motion of the atoms:

$$E_{tors} = \sum A[1 + \cos(n\pi - \varphi)] \tag{19}$$

where the sum extends over all types of torsional motion, with A being the amplitude, φ the dihedral displacement and n testifying of the symmetry of the torsional motion. For instance,$n = 2$ indicates a periodicity of π. These constants are assigned to each kind of torsional motion. Lastly, a fourth expression takes into account the non-bonded interactions: the van der Waals, repulsion and Coulomb interactions:

$$E_{nb} = \sum_{i>j} \left[-\frac{C_{ij}}{r_{ij}^6} + \frac{D_{ij}}{r_{ij}^{12}} + \frac{Q_i Q_j j_o}{\epsilon_r r_{ij}} \right] \tag{20}$$

where the sum is over each pair of non-bonded atoms i and j. The empirical coefficients C_{ij} and D_{ij} in (20) are assigned to each such pair, and Q_i and Q_j represent the electric charges of the non-bonded atoms. Since, for each atom, the interactions with non-bonded atoms far outnumber the interactions with neighbouring, bonded atoms, the computation of (20) consumes the most computing power. [42]

The force-field to be used in this study is $\text{ReaxFF}_{\text{SiO}}$, which has been developed for systems containing silicon and silicon oxides [50]. However, the force-field parameters in this study were also optimized for C and Ca, among others [41]. Unlike other force fields, $\text{ReaxFF}_{\text{SiO}}$ does not use fixed connectivity assignments, which allows for bonds to be created and dissociated during computations [50].

5.3 Molecular dynamics

Molecular Dynamics (MD) methods disregard the Schrödinger equations and instead use as a starting point Newton's equations of motion:

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \sum_{j \neq i} \frac{1}{\mathbf{r}_{ij}} \frac{d\phi(r_{ij})}{dr_{ij}} \mathbf{r}_{ij}. \tag{21}$$

Here, the masses m_i and displacement vectors $\mathbf{r}_i$ of each particle i are considered in a system of N particles. The pairwise interaction potentials are represented by $\phi(r_{ij})$, where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. The equations (21) are solved numerically, and the system eventually reaches equilibrium as time translates. The solution to this MD method is a microcanocial ensemble, which means that V, E and N are free variables. [51]

Since MM methods are unreliable if the optimization is done to a stationary point that is not a minimum [43], MD can be used to relax the structure towards a minimum (not necessarily a global minimum).

5.4 Details about Puck

In order to create the super cells from which the cylinders are cut, Puck first takes the highest value of the height or twice the radius, and doubles it. This value is then divided by the length of the longest of the three translation vectors, as counted from the cylinder origin. This value, rounded to the nearest integer, represents the number of unit cells in the positive and negative x-, y- and z- directions.

Afterwards, all atoms that are not located within the cylinder with the specified origin, radius and height are rejected. This is where the cylinder is actually cut out from the supercell. Since covalent bonds are broken after cutting out the cylinders, there are eight separate rules for restoring connectivity. All of the connectivity rules are carried out twice, because the changes done to comply with one rule might violate another rule.

Rule 1: Any H atom which does not have any O atoms inside the cylinder and within a distance of 1.1 Å is removed.

Rule 2: If an O atom has two H atoms within a distance of 1.1 Å, but one of them is outside the cylinder, they are all included in the cylinder.

Rule 3: If an O atom has two H atoms within a distance of 1.1 Å, but both of them are outside the cylinder, they are all included in the cylinder.

Rule 4: Any O atom which does not have any Si atoms inside the cylinder and within a distance of 2.0 Å, while one or several Si atoms are outside the cylinder and within 2.0 Å, is removed.

Rule 5: If a Si atom has four O atoms within a distance of 2.0 Å, but one of them is outside the cylinder, they are all included in the cylinder. If this O atom is also within 2.0 Å from another Si atom inside the cylinder, it is a bridging oxygen. This is a valid case, but if it is not connected to another Si atom, a H atom is added at a distance of 1.0 Å from the O atom, and a SiOH bond angle of 116.7 ^{o} is enforced.

Rule 6: If a Si atom has four O atoms within a distance of 2.0 Å, but two of them are outside the cylinder, they are all included in the cylinder. For both of these O atoms, the procedure for bridging O in rule 5 is then applied

Rule 7: If a Si atom has four O atoms within a distance of 2.0 Å, but three of them are outside the cylinder, they are all included in the cylinder. For all of these three O atoms, the procedure for bridging O in rule 5 is then applied.

Rule 8: If a Si atom has four O atoms within a distance of 2.0 Å, but all of them are out-

side the cylinder, they are all included in the cylinder. For all of these four O atoms, the procedure for bridging O in rule 5 is then applied.

After going through all the above rules twice, a sanity check is applied to review both Si and H connectivity. All Si atoms should be coordinated to four O atoms inside the cylinder and within a distance of 2.0 Å, and each H atom should have one O atom inside the cylinder and within a distance of 1.1 Å. If either connectivity is not satisfied, the program stops and the geometry file is not created. Due to the many different connectivity scenarios for O and Ca, these atoms are not subjected to connectivity rules [18].

As for establishing the charge neutrality, the defined charges are +4 for Si, +2 for Ca, +1 for H and -2 for O. Based on this, the computed charge can be positive, neutral or negative, and the absolute value can be even or odd. If the charge is neutral, the geometry file is created, otherwise, these charge rules are applied:

Rule 1: If the charge is positive and odd, then an OH group is removed to make the charge even. Puck looks for an O atom which has a H atom inside the cylinder and within a distance of 1.1 Å and removes this OH group. If no O and H atoms meet this criterion, Puck is terminated without creating the geometry file.

Rule 2: If the charge is positive and even, Ca atoms are removed until the charge is zero. Only Ca atoms that are located further away from the cylinder origin are removed.

Rule 3: If the charge is negative, Puck stops without creating the geometry file.

5.5 Example input script for ReaxFF$_{SiO}$ calculations on C-S-H

```
boundary f f f
read_data r___10.00_h___12.00_Si______32Ca______37OO_____167HH_____132n______0.lmp
pair_style reax/c NULL
pair_coeff * * ffield.reax Si O Ca O H
neighbor 0.8 bin
thermo 10
thermo_style custom step time temp etotal press lx ly lz xy xz yz
neigh_modify delay 0 check yes
group silica type <> 1 2
group Ca type == 3
group water type <> 4 5

fix charge all qeq/reax 1 0.0 10.0 1.0e-6 reax/c

# 1. minimization of all atoms except silica

print """Starting" minimization of all atoms except "silica"""
min_style cg
fix 1 silica setforce 0.0 0.0 0.0

minimize 1.0e-5 1.0e-7 300 1000
unfix 1


# 2. MD of all atoms except silica 2.5 ps

timestep 0.1
velocity silica zero linear
velocity Ca create 0.1 1881677804
velocity water create 0.1 1514666106
fix 1 silica setforce 0.0 0.0 0.0
fix 2 Ca nvt temp 0.1 500.0 1000.0
fix 3 water nvt temp 0.1 500.0 1000.0
print """running" 0.1 K->500 K "MD"""

run 2500
unfix 1
unfix 2
unfix 3


# 3. E minimization with no frozen atoms

print """Starting" minimization of all atoms in the "cell"""
min_style cg
dump 1 all xyz 1000 r___10.00_h___12.00_Si______32Ca______37OO_____167HH_____132n______0_step3.xyz
dump_modify 1 element Si O Ca O H

minimize 1.0e-5 1.0e-7 300 1000
undump 1


# 4. low temperature MD of the whole system

timestep 0.1
velocity all create 0.1 722092004
print """running" 0.1 K MD for the whole "system"""
fix 1 all nvt temp 0.1 0.1 1000.0

run 2500
unfix 1


# 5. E minimization with no frozen atoms

print """Starting" minimization of all atoms in the "cell"""
min_style cg
dump 1 all xyz 1000 r___10.00_h___12.00_Si______32Ca______37OO_____167HH_____132n______0_step5.xyz
dump_modify 1 element Si O Ca O H

minimize 1.0e-5 1.0e-7 300 1000
undump 1
```