

pyCSH: Automated atomic-level structure generation of bulk C-S-H and investigation of their intrinsic properties

Ziga Casar^{a,f}, Jon López-Zorrilla^b, Hegoi Manzano^b, Eduardo Duque-Redondo^{c,*}, Aslam Kunhi Mohamed^{d,e,**}, Karen Scrivener^a, Paul Bowen^a

^a Laboratory of Construction Materials, Institut des Matériaux, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

^b Physics Department, University of the Basque Country (UPV/EHU), 48940 Leioa, Basque Country, Spain

^c Physical Chemistry Department, University of the Basque Country (UPV/EHU), 48940 Leioa, Basque Country, Spain

^d Institute for Building Materials, Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, CH-8093 Zürich, Switzerland

^e Department of Civil Engineering, IIT Madras, Chennai 60036, Tamil Nadu, India

^f Department of Chemistry, Durham University, South Road, Durham DH1 3LE, United Kingdom

ARTICLE INFO

Keywords:

Calcium silicate hydrate
Brick model
Automated structure generation
pyCSH code

ABSTRACT

A new Python code for the automated generation of realistic bulk calcium silicate hydrate (C-S-H) structures is introduced. The code was used to generate 400 structures with Ca/Si of 1.3, 1.5, 1.7, and 1.9. The generated structures are in excellent agreement with experimentally measured C-S-H properties (Ca/Si, 2H/Si, MCL, Si-OH/Si, and Ca-OH/Ca). Molecular dynamics was used to simulate the structures, which were then investigated for their structural features and energetic stability. The results indicate very similar short-range ordering and energetic stability between all 400 structures. Finally, it is shown how computational C-S-H models can be used to understand the experimentally measured pair distribution functions. The code, named pyCSH, is available as open source under GPL-2 license at GitHub <https://github.com/hegoimanzano/pyCSH.git> and <https://github.com/akmiitm/pyCSH.git>.

1. Introduction

The cement and concrete industry is one of the biggest emitters of CO₂ [1]. The partial substitution of the clinker by supplementary cementitious materials (SCMs) can lower the CO₂ footprint of cement [2]. However, to understand the impact of these on properties a fundamental understanding of the unhydrated and hydrated phases in Portland cement is needed [3,4]. The focus of this work is on the main hydration product and primary binder in hydrated cement paste, calcium silicate hydrate (C-S-H).

C-S-H forms around 50 to 60 % of hardened cement paste and binds together other crystalline phases and aggregates [5]. The nucleation and growth of C-S-H dictates the early hydration behavior of Portland cement, from a few hours to around one day, [6]. The hydration results in a continuous nanoporous C-S-H network, which consists of nanofoils with lengths around 100 nm and below 5 nm in thickness [7–10].

Over the past decade significant progress has been made in

understanding the atomic structure of C-S-H. C-S-H can be represented as a highly defective tobermorite 14 Å [11] with variable chemical composition and structure [5,10,12–15]. In tobermorite a calcium oxide sheet has silicate chains on either side, which follow the dreierketten silicate structure. The dreierketten structure of the silicate chain has a repetition of a silicate dimer (Q^{2p}-Q^{2p}) followed by a bridging silicate (Q^{2b}) which is oriented into the interlayer (Fig. 1). The interlayer separates the calcium-silicate sheets of tobermorite and contains one Ca²⁺ ion and seven water molecules. This atomic arrangement in tobermorite 14 Å results in a Ca/Si ratio of 0.83 and a mean chain length (MCL) of infinity, since the dreierketten structure of the silicate chain is not broken. In C-S-H the Ca/Si ratio varies between 1.2 and 2.1, with an average value of 1.7 in plain Portland cement. The increase in the Ca/Si ratio is partly due to missing Q^{2b} silicates [13]. This is evident from ²⁹Si nuclear magnetic resonance (NMR) where a lower intensity of the Q^{2p} and Q^{2b} signals is observed, while the Q¹ (end of chain silicate) signal appears and dominates at high Ca/Si ratios [10,16,17]. From the

* Corresponding author.

** Corresponding author at: Institute for Building Materials, Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, CH-8093 Zürich, Switzerland.

E-mail addresses: eduardo.duque@ehu.eus (E. Duque-Redondo), akm@civil.iitm.ac.in (A.K. Mohamed).

<https://doi.org/10.1016/j.cemconres.2024.107593>

Received 5 February 2024; Received in revised form 23 May 2024; Accepted 20 June 2024

Available online 24 June 2024

0008-8846/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

proportion of those three Q species the MCL can be calculated [12,16]. With an increase in the Ca/Si ratio a decrease in the MCL is observed [5]. The MCL for Ca/Si = 1.7 is around 3.0, which means that only 25 % of the Q^{2b} silicates are present in the silicate chains. The missing Q^{2b} sites are then occupied by Ca^{2+} ions, which is a defining feature of the C-S-H structure [10]. Further increase of the Ca/Si ratio comes from additional Ca^{2+} in the water dominated interlayer. The charge neutrality is achieved by the presence of OH^- ions as indicated by the rising Ca-OH/Ca ratio with increasing Ca/Si ratio [5].

The highly defective atomic level structure of C-S-H is a nontrivial landscape to navigate. Extensive knowledge of the atomic level structure of calcium silicate hydrates is needed to introduce defects into tobermorite 14 Å which will satisfy the experimentally measured C-S-H properties and simultaneously fulfill the charge neutrality of the system over a wide range of Ca/Si compositions. Typical parameters from experimental measurements that are used to construct C-S-H models are Ca/Si ratio, MCL, relative concentration of silanol (Si-OH/Si) and hydroxide groups (Ca-OH/Ca). Due to the poorly ordered nanocrystalline nature of C-S-H the models need to be of sufficient size, nanometers, to accommodate the different types of defects and yield representative structures. The process of constructing such models can be difficult, tiresome and repetitive. Therefore, an automated tool for representative C-S-H structure generation would be of great service to the wider modeling and experimental community.

The first atomistic models of C-S-H appeared in 1996 [18]. Since then, many computational models have been developed [5,12,13,15,19,20]. In 2018 Kunhi Mohamed et al. [12] introduced the brick model, which allows for a reproducible approach for constructing models of bulk C-S-H and surfaces that are able to reproduce experimentally measured properties [5]. The advantage of the approach of Kunhi Mohamed et al. [12] lies in the exact inclusion of defects into the tobermorite 14 Å structure. These defects are missing Q^{2b} silicates, calcium replacement of the Q^{2b} silicate, protonation and deprotonation of Q species silanol groups, missing Q species in the silicate dimers, variable water content and addition of Ca^{2+} and OH^- ions in the interlayer. With the precise control over the introduced defects, it is possible to conduct studies in which a given defect can be studied in detail. The brick model in conjunction with experiments was successfully used in elucidating the bulk structure of high Ca/Si ratio C-S-H [10], as well as incorporation of aluminium [17] and zinc [16] into the bulk C-S-H structure.

The brick model [12] was also successfully used to construct representative models of C-S-H nanofoils with the basal (001) surface which are in excellent agreement with experimentally measured properties [21]. It was concluded that nanofoils which are 3 layers, 2 interlayers thick are in best agreement with the experimentally reported thickness

(< 5 nm) and specific surface area (200–300 m²/g). This thickness underlies the importance of surfaces when discussing the atomic-level structure of C-S-H, since one third of the calcium-silicate layers are associated with the surface. Furthermore, co-adsorption of calcium and hydroxide ions was proposed, leading to the formation of a calcium and hydroxide network on the surface, which can then explain the experimentally observed high Ca/Si ratios.

To fulfill the need for representative C-S-H structure generation, we present a new open source Python code for bulk C-S-H structure generation. The code uses the collected experimental data from Duque-Redondo et al. [5] and the brick model methodology from Kunhi Mohamed et al. [12] to generate automatically thousands of different representative bulk C-S-H structures with the desired chemical composition and very simple input parameters. In this work, we used the automatic code, named pyCSH, to generate 100 structures for each series of Ca/Si ratios (1.3, 1.5, 1.7, and 1.9). The resulting structures were compared with experimentally measured properties (MCL, Si-OH/Si, and Ca-OH/Ca). For a given Ca/Si ratio, the layer spacing, density, and pair distribution function (PDF) for each structure were calculated and thereafter analyzed. Finally, we compare the energetical stability of different defects.

2. Methods

2.1. pyCSH Python code

The pyCSH code works with the methodology of the brick model [12] which was written in C++. Defects are introduced into the unit cell of tobermorite 14 Å [11] by the arrangement of different atom groups. The atom groups considered are shown in Fig. 2. The silicate chains are formed from a Q^{2p} silicate, which can then be followed by a Q^{2b} silicate in the bridging site. The Q^{2b} silicate can either have both non-sharing oxygens deprotonated, or one can be protonated. If the dimer ends with a Q^1 silicate, the non-sharing oxygen can be either protonated or deprotonated. Following the Q^1 silicate, the bridging site can be either unoccupied or occupied with a Ca^{2+} ion. The interlayer is filled with Ca^{2+} and OH^- ions and water molecules. Altogether in the interlayer 5 different positions for Ca^{2+} are possible, 10 for OH^- and 17 for water. While Kunhi Mohamed et al. [12] introduced the defect of the missing dimer silicate, it is not considered in the pyCSH code [10,16]. The absence of a silicate from the Q^1 - Q^1 dimer would result in a new Q^0 silicate, which has a distinguished ²⁹Si NMR chemical shift. ²⁹Si NMR measurements of pure phase C-S-H do not show such a chemical shift [10,16]. An example of a brick which is formed from different atom groups is shown in Fig. 3a.

The automated generation of bulk C-S-H structures is outlined with

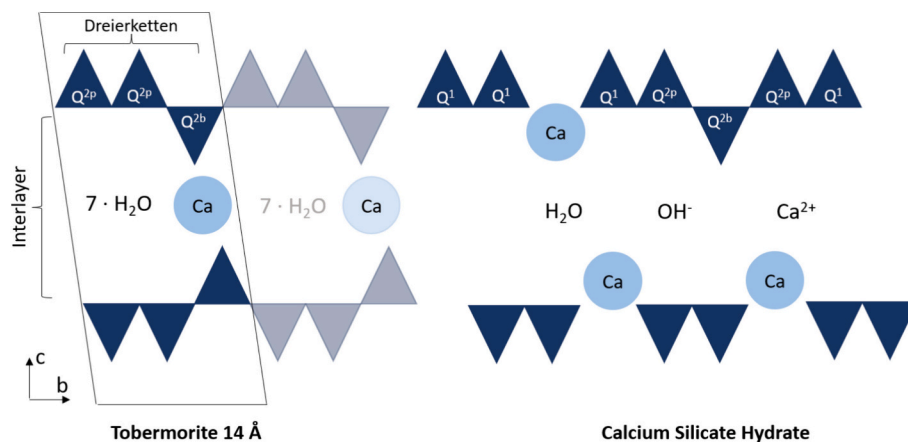


Fig. 1. Schematic representation of the unit cell of tobermorite 14 Å and calcium silicate hydrate. The calcium oxide layers of the calcium-silicate sheets are not shown.

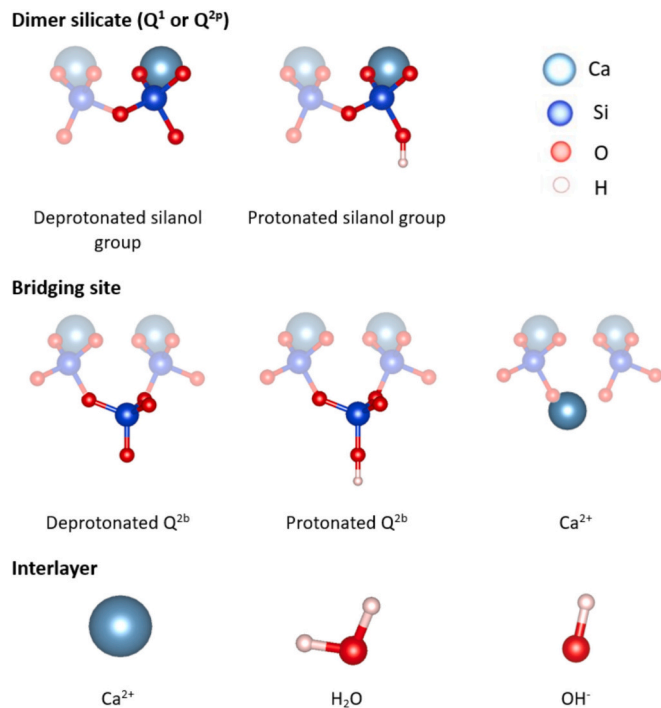


Fig. 2. Considered atom groups (opaque) to generate the representative C-S-H structures. The semitransparent atom groups show the connectivity of the atom groups in the silicate chains.

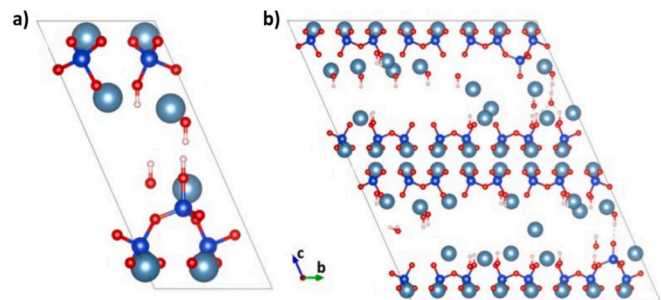


Fig. 3. a) Example of a constructed brick from different atomic groups with a missing Q^{2b} silicate, the site is occupied by Ca^{2+} , a protonated Q^{2b} , and 2 additional Ca^{2+} and 2 OH^{-} in the interlayer. b) A $1 \times 4 \times 2$ C-S-H super cell with $Ca/Si = 1.7$. Water is hidden for clarification. These are unrelaxed structures. The colour legend is the same as in Fig. 2.

the flow chart in Fig. 4. The code needs 9 input parameters. The structural ones are the size of the system (number of defective tobermorite 14 Å unit cells (bricks) in the a, b and c crystallographic directions), the targeted Ca/Si and $2H/Si$ ratios, as well as the number of independent structures which should be generated for the given Ca/Si and $2H/Si$ ratios. The numerical input parameters are the random seed number and the tolerances of the Ca/Si , $Si-OH/Si$ and $Ca-OH/Ca$ ratios, which determine if the generated structure is accepted or not. From the targeted Ca/Si ratio the corresponding $Si-OH/Si$, $Ca-OH/Ca$ and MCL values are taken from the data collected in Duque-Redondo et al. [5].

The C-S-H supercell is formed by a random combination of bricks, each of them composed of a random combination of predetermined atomic groups (Fig. 2). Therefore, the code first generates all possible bricks, although certain limitations are imposed in this process. If a dimer silicate with a silanol group is placed into the silicate chain then a Q^{2b} silicate is prevented from being placed into the bridging site, which can then be only unoccupied or occupied with a Ca^{2+} . This is due to the

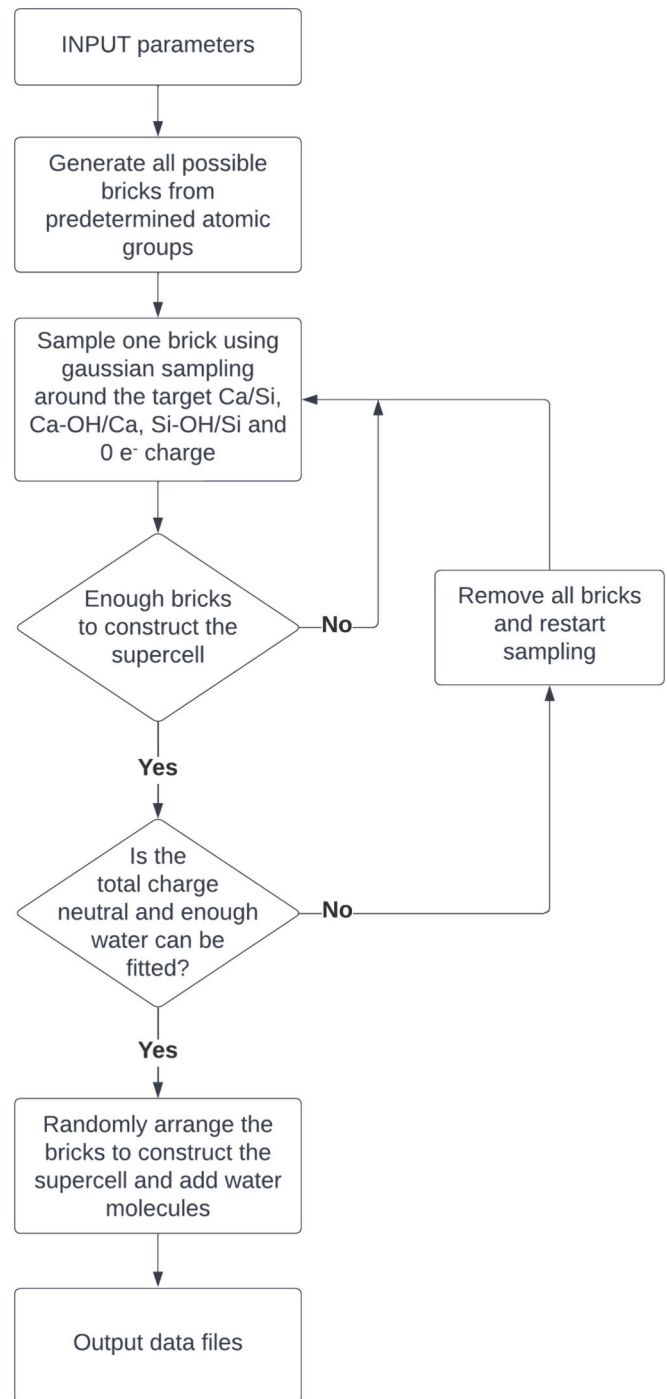


Fig. 4. Flow chart of the pyCSH code.

fact that silicates interlink through oxygen and not hydroxyl groups [11,13]. The overall charge of each brick is calculated. If the charge is less than $-2e$ or more than $+2e$ the brick is discarded and a new one is constructed. Once all bricks are generated, the required number of bricks to construct the supercell is sampled to match the target Ca/Si ratio, as well as the experimental $Si-OH/Si$ and $Ca-OH/Ca$ ratios and the overall charge neutrality of the supercell. This is achieved by a gaussian sampling process, in which the widths of the gaussians used for the sampling are controlled by the above-mentioned input parameters (Ca/Si , $2H/Si$). Once a neutral supercell is obtained, water insertion follows. The number of water molecules that a model can accommodate depends on its constituent bricks; so if a supercell is not able to include enough

water, it is discarded, and the sampling process is restarted. Finally, hydrogen atoms of hydroxyl groups and water molecules are iteratively added trying to minimize the overlap among them. The amount of water needed is calculated from the experimentally measured 2H/Si ratio [5] after the subtraction of silanol groups and hydroxide ions. ^1H NMR is used to experimentally determine the amount of water in C-S-H, but it is difficult to distinguish between hydrogens of water molecules and other hydrogens [22]. Therefore, the reported 2H/Si equals $(\text{H}_2\text{O} + 0.5\text{OH}^-)/\text{Si}$. The water molecules are uniformly distributed through the entire supercell structure. The geometry of the inserted water molecules is from the TIP4P/2005 water model [23], and therefore suitable for force fields that use this water model without additional adjustments. An example of a generated C-S-H model with $\text{Ca}/\text{Si} = 1.7$ measuring 4 bricks in the b-direction and 2 bricks in the c-direction is shown in Fig. 3b.

In addition to the automatic generation of bulk C-S-H structures the program allows for manual creation of structures and surfaces. It is possible to take an automatically generated structure and further change it manually. The details of the alphanumeric notation that is used by the program are given in detail and explained in the Supplementary Information (SI).

2.2. Molecular dynamics

Since the atoms are placed in a predetermined way with fixed coordinates a suitable equilibration protocol was devised. First, three separate 40,000 step runs in the canonical (NVT) ensemble were carried out, where first only the water molecules are allowed to move, followed by water plus hydroxyl groups and hydroxide ions, and finally including all calcium atoms. Next, energy minimization of the simulation box was carried out. With this a more favorable (lower energy) atom arrangement in the interlayer is achieved. Due to the use of the adiabatic core-shell model for the polarizability of silicate oxygens an additional step followed to prevent the drift of the total energy during the simulation. In this step, the motion of cores and shells is decoupled. For details we refer the reader to the paper of Valavi et al. [24], the original paper of the adiabatic core-shell model [25] and the LAMMPS documentation (8.5.5. Adiabatic core/shell model) [26].

A so-called hard minimization followed, where the structure is subjected to cyclic heating. A 100 ps run in the isobaric-isothermal ensemble (NPT) ensemble at 300 K and 1 atm was carried out to relax the simulation box. This was followed by a heating and cooling step in the NVT ensemble, whereby the temperature was raised over 50 ps to 700 K, where the system was kept for 200 ps before a 50 ps cooling to 300 K. This cycle was repeated one more time before the system was equilibrated in the NPT ensemble at 298.15 K and 1 atm for 800 ps.

The production run was then carried out over 20 ns in the NPT ensemble at 298.15 K and 1 atm. The trajectories were recorded every 50 ps. The temperature, pressure, energy, volume and unit cell parameters were averaged and recorded every 20 ps. The average potential energy of core-shell pairs was also averaged and recorded every 20 ps. Due to the polarizable force field and water holding interlayers the bulk C-S-H structure needs roughly 10 ns in the production run to reach steady conditions. Therefore, the reported average values were taken from the last 6.7 ns (see SI section 3).

All simulations were carried out with LAMMPS [26]. The force field used was the Erica FF2 [24], with a time step of 0.2 fs. The only modification to the force field was the adjustment of the mass of the core-shell oxygen pair. To prevent the drift of total energy the mass of the shell was set to 0.4 g/mol, while the core mass was 15.59 g/mol. The K-space solver was the particle-particle particle-mesh solver (pppm) [27] with an accuracy of 10^{-4} . The Nosé-Hoover thermostat and barostat [28,29] were used with Tdamp equal to 100 times the time step and Pdamp equal to 1000 times the time step.

The program TRAVIS - Trajectory Analyzer and Visualizer [30,31] was used to postprocess the trajectory files. The partial pair distribution

functions (PDFs), $G_{\alpha\beta}(r)$, were calculated as the average of the entire 20 ns production run. From the partial PDFs the total pair distribution function, $G(r)$, for each structure was calculated as:

$$G(r) = \sum_{\alpha\beta} \frac{c_\alpha c_\beta b_\alpha b_\beta}{\langle b \rangle^2} G_{\alpha\beta}(r) \quad (1)$$

where b values are the atomic scattering lengths of the atoms, c is the weight factor representing the fraction of atomic species and the sums are over every atom in the sample (except hydrogens, which were excluded from the analysis) [32]. Fixed scattering lengths at zero degrees were used.

For the energetical comparison of structures with the same stoichiometry the enthalpies of the observed C-S-H structures are compared the enthalpies of quartz (SiO_2), lime (CaO) and water (H_2O) via Eq. (2) [12]. To obtain the average enthalpies of quartz, lime and water, the systems were simulated for 20 ns, where the average enthalpy over the last 10 ns was taken into consideration. The simulation protocol was the same as for C-S-H, as stated above.

$$n_1 \bullet \text{CaO} + n_2 \bullet \text{SiO}_2 + n_3 \bullet \text{H}_2\text{O} \rightarrow \left[(\text{CaO})_{n_1} (\text{SiO}_2)_{n_2} (\text{H}_2\text{O})_{n_3} \right]_{\text{CSH}} \quad (2)$$

where the normalization comes by fixing $n_1 + n_2 + n_3 = 1$. A correction for water splitting needs to be considered ($\text{H}_2\text{O} + \text{O} \rightarrow 2 \bullet \text{OH}^-$) and is equal to 2.9 eV per dissociated molecule [24]. Further, the enthalpy of the C-S-H structure is adjusted for the missing Morse bond energy ($N_{\text{OH}} \bullet D$) and potential energy of the core-shell pairs (PE_{CS}):

$$H_{\text{CSH,corr}} = H_{\text{CSH}} + N_{\text{OH}} \bullet D + PE_{\text{CS}} \quad (3)$$

All LAMMPS and TRAVIS input files are provided as part of the SI section.

3. Results

In this results section first the generated structures are analyzed and compared to experimentally obtained C-S-H properties. Afterwards the simulated structures are analyzed, and a comparison of the defect energies is presented.

3.1. Initial structures

Fig. 5 compares the four C-S-H properties (MCL, 2H/Si, Si-OH/Si and Ca-OH/Ca) for the given Ca/Si ratio between the experimentally measured systems and the pyCSH generated models. It is clear from Fig. 5 that the MCL decreases with increasing Ca/Si ratio. However, for Ca/Si ratios above 1.2 the experimental MCL varies significantly, with values between 2.3 and 4.0. The pyCSH values are listed in Table 1. The range of predicted MCL values is in the range of experimentally determined values. For Ca/Si = 1.7 the predicted MCL (2.30 ± 0.11) is at the lower end of experimental data, while for Ca/Si = 1.9 the MCL is in the upper range (3.457 ± 0.083). Two separate ^{29}Si NMR studies on synthetic phase pure C-S-H with Ca/Si = 1.75 underline the variation in MCL as calculated from the proportion of Q species. While both studies used the same precipitation protocol [33], Kumar et al. [10] reported an MCL value of 2.55, while Morales et al. [16] reported 2.89. The targeted 2H/Si ratios are predetermined and are in the range of experimentally measured values.

The generated mean Si-OH/Si ratios for Ca/Si of 1.3 and 1.5 are slightly higher than the experimentally reported values. While the Si-OH/Si for Ca/Si 1.7 and 1.9 are closer to the linear extrapolation of the collected experimental data points. There are no experimental values reported above Ca/Si = 1.6. The experimentally determined Ca-OH/Ca ratios show a higher discrepancy above Ca/Si = 1.5, which may indicate that the trend is not linear for the entire range of Ca/Si ratios. The pyCSH code predicts structures which are in good agreement with the experimental linear fit of Ca-OH/Ca for Ca/Si 1.3, 1.5 and 1.9. The

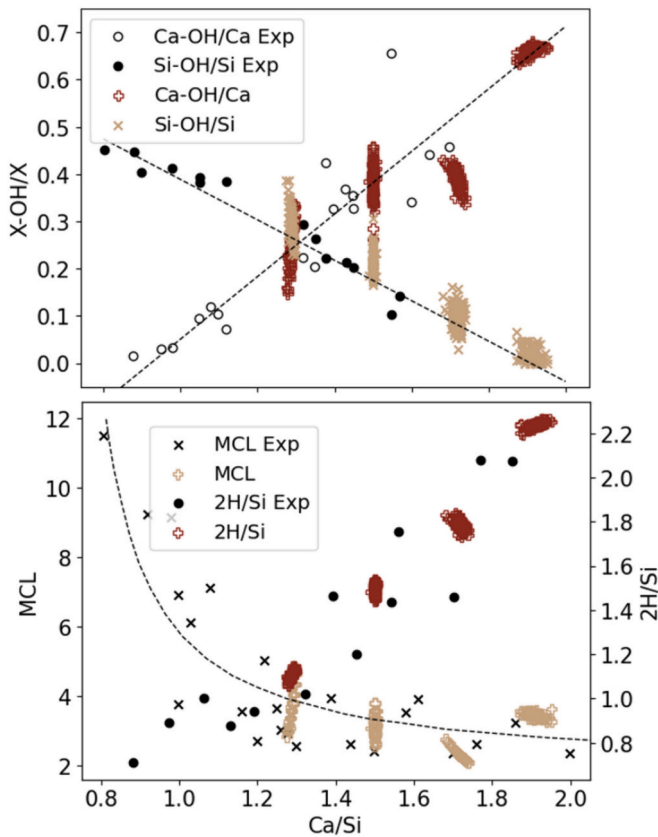


Fig. 5. Characteristic properties (Si-OH/Si, Ca-OH/Ca, MCL and 2H/Si) of bulk C-S-H generated structures for the studied Ca/Si ratios in comparison with experimental data from Duque-Redondo et al. [5].

Table 1

Mean value and standard deviation (std) of properties at a given Ca/Si of bulk C-S-H models as predicted by pyCSH.

Ca/Si	Property	Mean	Std
1.3	MCL	3.668	0.407
	Si-OH/Si	0.294	0.034
	Ca-OH/Ca	0.251	0.043
	2H/Si	1.109	0.020
	MCL	3.161	0.268
1.5	Si-OH/Si	0.213	0.024
	Ca-OH/Ca	0.384	0.036
	2H/Si	1.495	0.022
	MCL	2.299	0.117
	Si-OH/Si	0.097	0.025
1.7	Ca-OH/Ca	0.387	0.021
	2H/Si	1.780	0.022
	MCL	3.457	0.083
	Si-OH/Si	0.018	0.015
	Ca-OH/Ca	0.660	0.010
1.9	2H/Si	2.238	0.014

predicted values for Ca/Si = 1.7 (0.387 ± 0.021) are at the lower end of experimental values.

It is seen that the automated pyCSH code generates bulk C-S-H structures which are in good agreement with experimentally measured properties. The model structures are an assembly of different defective tobermorite bricks, therefore the resulting structures have a random assembly of known C-S-H defects. The samples on which experimental measurements are carried out are often prepared with slightly different approaches (initial materials, water-to-solid ratios, time of hydration), as well as different storage conditions (stored in solution, freeze-dried) [5]. This could partially explain the relatively large scatter of the

experimental data, particularly the MCL and Ca-OH/Ca. Further, the experimental measurements do not only measure the bulk properties, but also include surfaces. It is suggested that the C-S-H nanofoils measure <5 nm in thickness [21], so the surface silicate chains represent one third of the overall silicate chains in C-S-H. While the surfaces should not affect the MCL, they might have a great effect on the Ca-OH/Ca ratio [21]. The solution composition dictates the deprotonation of surface silanol groups. High pH solutions (above pH 12) give high degrees of deprotonation of surface silanol groups [34]. Deprotonation degrees of 80 % and higher would greatly affect the Si-OH/Si ratio of the bulk. At such surface deprotonation degrees, the Si-OH/Si of the bulk would need to be higher than the currently assumed overall value. Simultaneously a higher Si-OH/Si would result in higher Ca-OH/Ca due to the necessity for charge neutrality of the structure. In a C-S-H structure with a given amount of calcium and silicates, for each additional protonation of one silicate oxygen an additional hydroxide ion in the interlayer is needed to maintain the exact same charge of the system. Simultaneously a water molecule must be removed to maintain the 2H/Si ratio.

3.2. Structural analysis

The structural analysis was made after molecular dynamics relaxation of the initial bulk C-S-H structures. Roughly 5 % of simulations crashed. The reported quantities are the average values of the 20 ns production run.

The layer spacing of all calculated structures, including the mean values and standard deviations for each Ca/Si ratio is shown in Fig. 6. The predicted layer spacing increases from Ca/Si ratio of 1.3 (12.18 Å) to 1.5 (13.11 Å) where it stays nearly constant until Ca/Si = 1.7 (13.05 Å). The layer spacing for Ca/Si = 1.9 increases significantly thereafter (16.01 Å). The average Ca/Si = 1.5 structure holds 8 % less water molecules and hydroxide ions than the Ca/Si = 1.7 structure (see $(\text{H}_2\text{O} + \text{OH}^-)/\text{Ca}$ in Table 2). The similar mean values of the layer spacing for the 1.5 and 1.7 structures can be explained by the mean coordination number of oxygen atoms to calcium atoms (Table 2). The coordination number (cutoff at 3 Å) for Ca/Si = 1.5 equals 6.22 ± 0.06 , while for Ca/Si = 1.7 it equals 6.78 ± 0.06 (8 % difference). A higher coordination number results in a denser packing of the bulk and therefore yields a higher density for the same layer spacing. Indeed, a slightly higher mean density (Fig. 7) is observed for Ca/Si = 1.7 (2.40 ± 0.013 g/cm³) in comparison to Ca/Si = 1.5 (2.37 ± 0.017 g/cm³).

It appears that the interlayer spacing linearly increases with increasing $\text{Ca}^{2+}/\text{dimer}$ (see Fig. 6c). Further, the calculated layer spacing is comparable at the same $\text{Ca}^{2+}/\text{dimer}$ as well as Ca-OH/Ca (Figs. 6c and S1). This correlation can be explained by the following observation. The MCL is a measure of how many bridging sites of the silicate chain are occupied by Q^{2b} silicates. Assuming that each missing Q^{2b} is replaced by a Ca^{2+} , which we will consider as part of the silicate chain instead of the interlayer [10,12], one can calculate the number of remaining Ca^{2+} which are associated with the interlayer per each C-S-H brick (defective unit cell of tobermorite). Since the C-S-H brick always consists of the main layer Ca-Si sheet, the Ca^{2+} content can be normalized by the silicate dimer ($\text{Q}^1\text{-Q}^1$, $\text{Q}^{2p}\text{-Q}^{2p}$ or $\text{Q}^1\text{-Q}^{2p}$) of each silicate chain (top and bottom of the interlayer). The property $\text{Ca}^{2+}/\text{dimer}$ is listed in Table 2. The mean value of those interlayer Ca^{2+} at Ca/Si = 1.5 equals $0.69 \text{ Ca}^{2+}/\text{dimer}$, while for Ca/Si = 1.7 equals 0.68. The roughly same amount of Ca^{2+} in the interlayers explains the comparable $\text{Ca}^{2+}/\text{dimer}$ ratios and the similar layer spacings. Indeed, similar layer spacings are also observed for Ca/Si = 1.3 at the same $\text{Ca}^{2+}/\text{dimer}$ ratios (Fig. S1).

The pyCSH predicted MCL for Ca/Si = 1.9 (MCL = 3.5) is higher than the one for Ca/Si = 1.7 (MCL = 2.3) which results in less calcium in the silicate chains and therefore high $\text{Ca}^{2+}/\text{dimer}$ (1.77 versus 0.68). Therefore, care needs to be taken when interpreting the Ca/Si = 1.9 structures. The results suggest that there is an interplay of experimentally measured properties (Ca/Si, MCL, Si-OH/Si, Ca-OH/Ca), which

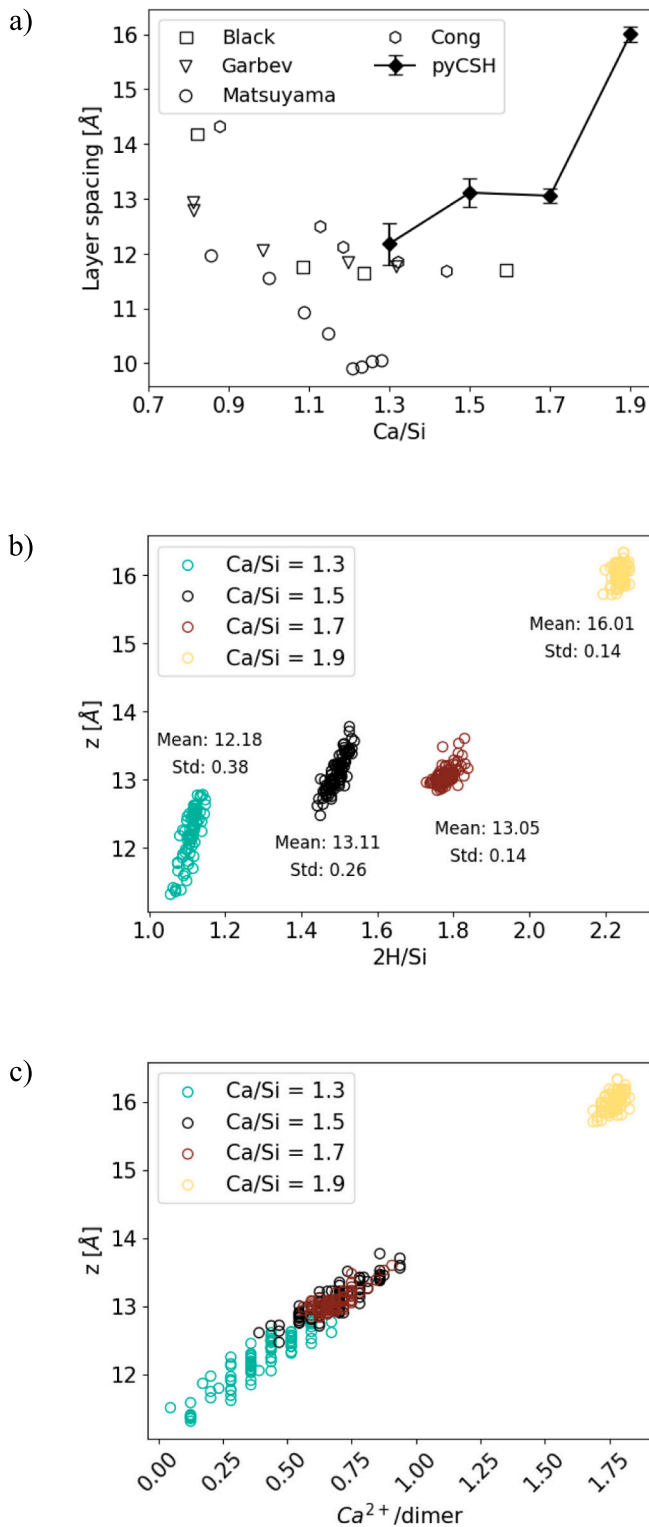


Fig. 6. Layer spacings (z) as a function of Ca/Si ratio (a), $2H/Si$ (b) and $Ca^{2+}/dimer$ (c). Experimental data in (a) are from Black et al. [35], Garbev et al. [36], Matsuyama and Young [37] and Cong and Kirkpatrick [38].

cannot be captured by considering the C-S-H bulk structure only. While the $Ca/Si = 1.7$ structure contains approximately the same amount of hydroxide ions as $Ca/Si = 1.5$, it contains more water, and its silicate chains incorporate more Ca^{2+} ions. The higher Ca^{2+} incorporation results in a higher amount of strongly bound water next to the silicate chains (due to Ca^{2+} incorporation into the Si-chains which is known for

Table 2

Mean value and standard deviation for calcium coordination number (Ca—O CN), H_2O/Ca , $(H_2O + OH^-)/Ca$ and $Ca^{2+}/dimer$ for the given Ca/Si ratios. $Ca^{2+}/dimer$ accounts only for calcium ions in the interlayer, excluding those which are predicted to be in the bridging site.

Ca/Si	Ca—O CN	H_2O/Ca	$(H_2O + OH^-)/Ca$	$Ca^{2+}/dimer$
1.3	6.11 ± 0.05	0.622 ± 0.003	0.747 ± 0.020	0.38 ± 0.14
1.5	6.22 ± 0.06	0.736 ± 0.002	0.926 ± 0.018	0.69 ± 0.12
1.7	6.78 ± 0.06	0.815 ± 0.005	1.008 ± 0.015	0.68 ± 0.06
1.9	6.35 ± 0.04	0.837 ± 0.008	1.167 ± 0.006	1.77 ± 0.03

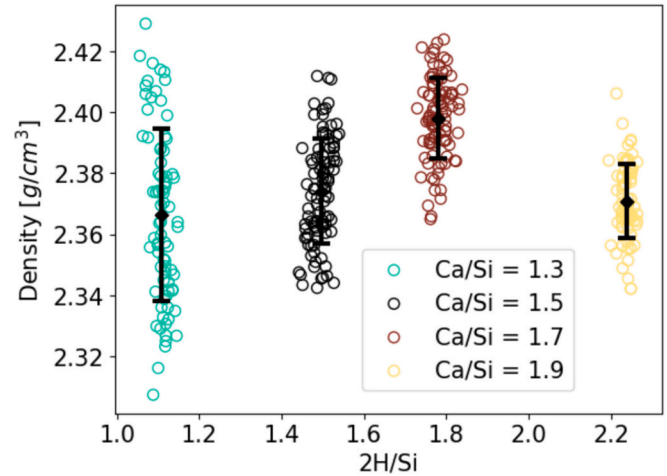


Fig. 7. Density. The lines determine the standard deviation around the mean value.

having a tight hydration shell [39]). This would potentially result in a comparable amount of water in the mid-interlayer (water which is not coordinated to the main layer chains).

Experimentalists estimate the layer spacing from the (001) basal reflection in the X-ray diffraction (XRD) pattern [13,40]. The general trend observed in those studies is a decrease in the layer spacing with increasing Ca/Si ratios (Fig. 6a). The layer spacing seems to be unchanged above $Ca/Si = 1.3$ [13]. Some authors even report constant layer spacings starting at $Ca/Si = 1.1$ [41]. This trend is not observed in the calculated structures with Ca/Si of 1.3 and 1.9. While the MD predicted layer spacings is around ~ 13 Å ($Ca/Si = 1.5$ and 1.7), the measured one for $Ca/Si = 1.5$ and above are typically between 11 and 12 Å [13,40]. However, the experimental readings are from samples dried prior to the measurement [36], possibly underestimating the layer spacing. The predicted $Ca^{2+}/dimer$ for $Ca/Si = 1.3$ is lower than the one for $Ca/Si = 1.5$ (Fig. 6c). The tobermorite 14 Å has $Ca^{2+}/dimer = 0.5$ and is a stable crystal structure. The $Ca^{2+}/dimer = 0.38$ at $Ca/Si = 1.3$ seems in comparison to tobermorite 14 Å ($Ca/Si = 0.86$) very low and could therefore explain the MD predicted layer spacing, which is lower than the $Ca/Si = 1.5$ layer spacing. The large spread of the density at $Ca/Si = 1.3$ (Fig. 7) might originate from the distribution of silicates in bridging sites (Q^{2b}). Bridging sites in opposite positions implies steric repulsions which increases the basal space and consequently reduces the density.

The pair distribution functions (PDF) in Fig. 8 are obtained by taking the mean value over all PDFs of the same Ca/Si ratio. A comparison between individual PDFs with the mean PDF is given in SI section 5 (Fig. S2) and shows a negligible deviation, suggesting that all structures at a given Ca/Si relax into very similar atomic arrangements. By analyzing the individual radial distribution functions, it is possible to assign the individual peaks to the respective atom pairs. The recent models of C-S-H nanofoils predict <5 nm thick structure, which measures 3 layers, 2 interlayers in the z -axis direction [21]. Potential-mean-force calculations of the interaction of C-S-H nanofoils predict an

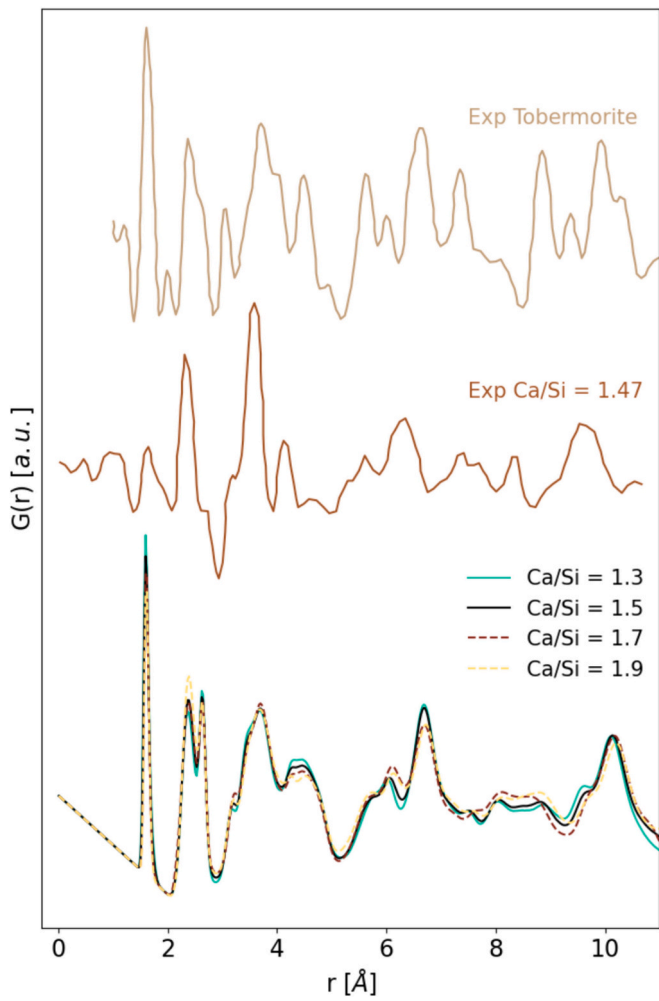


Fig. 8. Mean pair distribution functions (PDF) of bulk C-S-H structures at the given Ca/Si ratio. The experimental (Exp) PDFs are taken from Grangeon et al. [43]. For a direct comparison of experimental and calculated PDFs see SI section 2.

equilibrium distance between nanofoils at 5 nm [42]. Therefore, the experimental PDFs below 20 Å can be safely used to explore the structure of a single C-S-H nanofoil. Since the silicate chain structure of the bulk and surface seems to be the same [21] bulk C-S-H models can be used to help resolve the experimental PDFs (Fig. 8).

There is no significant difference between the MD predicted PDFs, as indicated by Fig. 8. The first three peaks (Si—O at 1.6 Å, Ca—O at 2.39 Å and O—O at 2.65 Å) correspond to the first distances inside the calcium-silicate sheets. This suggests that although the chains at different Ca/Si ratios have a different number of defects (missing Q^{2b} silicates and protonation of silanol groups) the local atomic environment is not

changed. However, with the increase in Ca/Si the intensity of the MD predicted Si—O peak decreases relative to the Ca—O peak, which was previously experimentally observed [43]. This decrease in intensity can be attributed to a lower relative number of silicate groups. Again, the measured PDF did not show differences in structural features below 20 Å, suggesting that short range ordering of C-S-H does not change with additional hydration time.

Grangeon et al. [43] reported the partial PDFs of freeze-dried C-S-H hydrated from $Ca(OH)_2$ and amorphous SiO_2 . While they analyzed samples with Ca/Si from 0.57 to 1.47, the results suggest similar short range structural features for the different samples. Table 3 shows the comparison of structural features (peaks) between the experimentally reported PDFs of Grangeon et al. and our PDFs from simulated pyCSH-generated structures (for Ca/Si ratio of 1.3 to 1.7). While the reported experimental PDF is for the in-house synthesized tobermorite [43], it shows very similar short-range ordering to the C-S-H samples. The partial experimental PDFs show more peaks than are shown in Table 3, which could be attributed to the presence of surfaces and other phases than C-S-H. As seen in Table 1 up to 10 Å the results are in very good agreement, giving confidence to the pyCSH-generated structures and the short range similarity between the tobermorite and C-S-H structure. Due to simulation box size, the PDF calculation was limited to ~10 Å.

It is interesting to analyze the Si—Si distances, which are reported in Table 3 and SI section 2. The first Si—Si distance (3.16 Å) corresponds to the separation of Si atoms of neighboring silicates (d_1 in Fig. 9). 4.39 Å is the distance (d_2) between silicates which are separated by a bridging site (Q^{2b} or Ca^{2+}). 5.54 Å is the approximate nearest distance between silicates of neighboring chains in the a-axis direction (d_a) as well as in the c-axis direction (silicate chain on the opposite side of the calcium-oxide layer, d_c), while 6.69 Å is the separation of a silicate with the second closest silicate of the neighboring chain in a-axis direction. A similar distance corresponds to the closest Q^{2b} silicates across the interlayer. While there are not many Q^{2b} silicates at Ca/Si = 1.5 and above, the 6.69 Å peak should gain intensity with decreasing Ca/Si ratios, which is indeed the case (Fig. 8). The 7.19 Å peak corresponds again to the separation of in-chain silicates (d_{i-c}). The peaks above 7.19 Å are harder to resolve, since distances of around 9.8 Å correspond to silicates of neighboring chains in a-axis direction, as well as to dimer silicates across the interlayer. The comparison of the Si—Si RDF (SI section 2) for all calculated Ca/Si shows indeed an excellent agreement up to 8.5 Å (in-chain silicates) after which the Ca/Si = 1.9 peaks are shifted to slightly higher distances, indicating the increased layer spacing. Distances of 10.7 Å correspond to in-chain silicates (approximate length of a pentamer) and to across the interlayer silicates.

The calcium-silicate main layer structural features are predicted with good accuracy. Therefore, it should be possible to resolve the interlayer spacing by carefully analyzing the short-range Si—Si distances. Due to different defects along the silicate chains, one should expect a slight variation in the Si—Si distance across interlayer silicate chains. Therefore, the PDF peak for the Si—Si across interlayer separation should have a lower intensity, be wider, and located below the XRD reported basal spacings. Fig. 9 shows the partial Si—Si PDF from Grangeon et al. [43],

Table 3

Comparison of individual peaks between MD calculated and experimental PDFs (EXP) of freeze dried tobermorite synthesized from $Ca(OH)_2$ and amorphous SiO_2 [43]. All reported values are in Å.

Si-O		Ca-O		Si-Si		Ca-Si		Ca-Ca		O-O	
MD	EXP	MD	EXP	MD	EXP	MD	EXP	MD	EXP	MD	EXP
1.60	1.60	2.39	2.42	3.16	3.07	3.47	3.59	3.79	3.85	2.65	2.64
3.83	3.63	4.56	4.5	4.39	4.3	4.77	4.62	5.33–5.72	5.64	3.11	3.12
4.28	4.27	5.91	5.71–6.11	5.54	5.58	5.26	5.13	6.66	6.65	4.74	4.89
5.96	5.9	7.13	7.36	6.69	6.72	6.08	6.00	7.57	7.37	5.59	5.77
6.87	6.58	8.16	7.75	7.19	7.36	6.71	6.52	8.76	8.86	6.66	6.72
8.09	8.12	9.01	8.96	9.86	9.71	8.51	8.21	9.28	9.20		
				10.7	10.25	9.70	9.52				

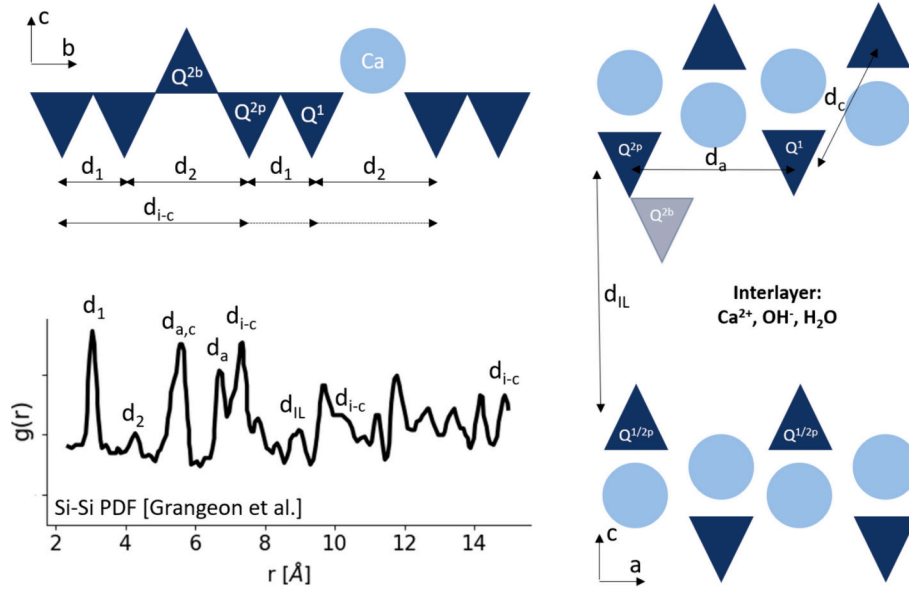


Fig. 9. Si-Si PDF from Grangeon et al. [43] with annotated Si-Si distances as predicted from MD.

where the short-range Si-Si peaks (d_1 , d_2 and d_{i-c} , d_a , d_c) from MD are projected onto the experimental PDF. The predicted interlayer separated dimer Si-Si experimental peak (dried samples) is marked as d_{IL} and is centered around 9 Å. This experimental value is slightly below our MD predicted one (9.8 Å), and therefore in agreement with the layer spacings reported from XRD [13,40].

3.3. Energetic analysis

Kunhi Mohamed et al. [12] studied the energetic stability of silanol groups in bulk C-S-H which were composed of only one type of brick. Their findings suggest an energetic preference for the absence of silanol groups. In the current study the energetic stability of defects from a random defect configuration with representative C-S-H structures was analyzed. The formation enthalpies are calculated through Eq. (2),

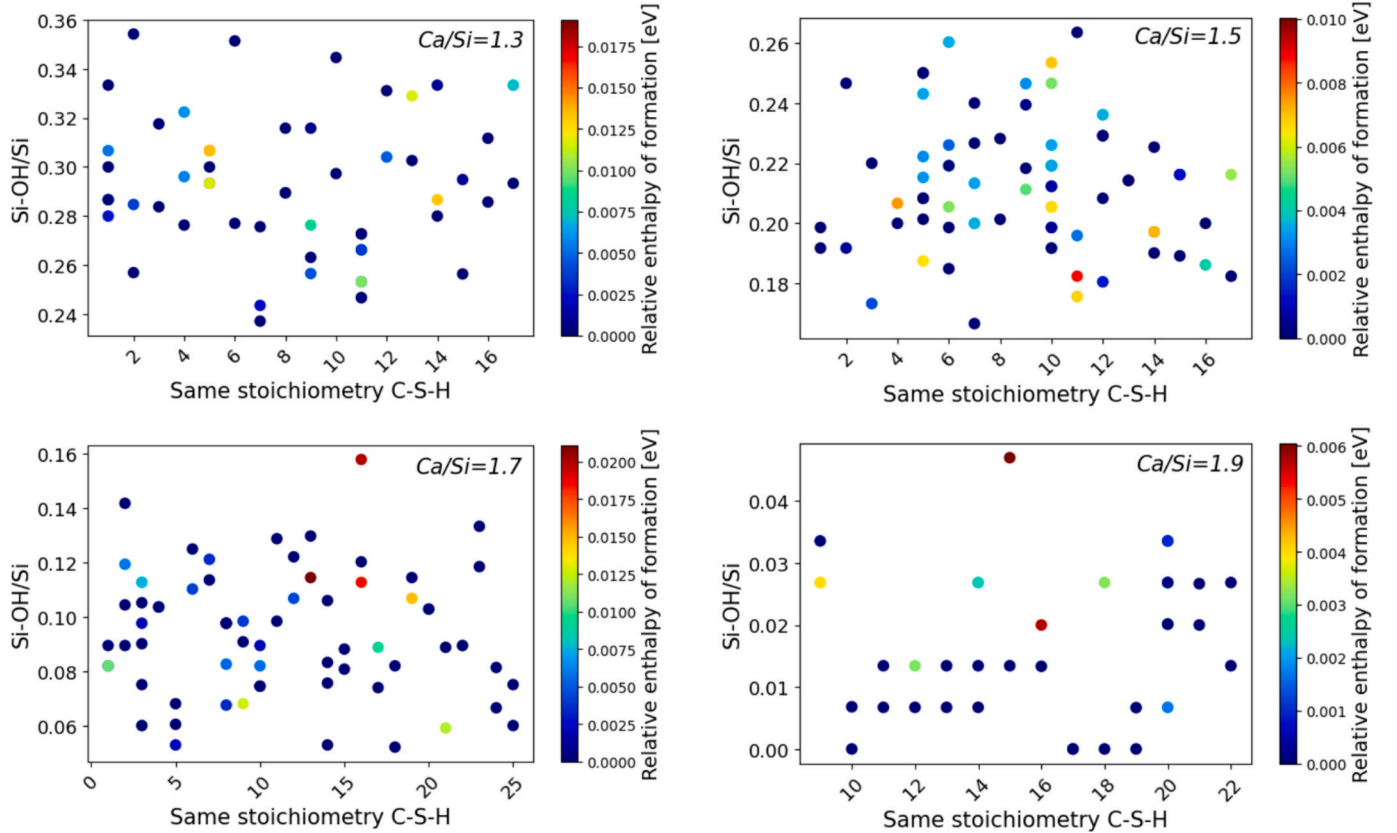


Fig. 10. Relative enthalpy of formation as a function of Si-OH/Si for C-S-H structures with the same stoichiometry. Structures with the same stoichiometry have the same x-axis value. The enthalpy of formation is normalized to the structure with the lowest defect energy for a given stoichiometry.

therefore the comparison can only be done for the C-S-H structures with the same stoichiometry ($C_{n1}S_{n2}H_{n3}$). Structures with the same stoichiometry have the same x-axis value in Fig. 10. The enthalpy of formation of structures with the same stoichiometry is normalized with respect to the lowest enthalpy of formation for a given stoichiometry. Since pyCSH randomly generates structures some of them are unique in their stoichiometry and therefore cannot be compared. Altogether 55 structures with 17 unique stoichiometries for $Ca/Si = 1.3$, 71 structures with 17 unique stoichiometries for $Ca/Si = 1.5$, 68 structures with 25 unique stoichiometries for $Ca/Si = 1.7$ and 34 structures with 22 unique stoichiometries for $Ca/Si = 1.9$ are compared.

While the Si-OH/Si ratio varies by up to 0.1 the variation of enthalpies of formation is minimal. The maximum difference between two structures of the same stoichiometry is 0.02 eV per simulation cell at $Ca/Si = 1.3$, 0.01 eV at $Ca/Si = 1.5$, 0.02 eV at $Ca/Si = 1.7$ and 0.006 eV at $Ca/Si = 1.9$. These very low differences suggest that there is no significant effect of the number of silanol groups on the energetical stability of C-S-H with randomly introduced defects. We remind the reader that $k_B T$ at the simulated temperature (298.15 K) equals 0.0257 eV. Furthermore, no correlation between the number of silanol groups and enthalpies of formation can be found. This implies that there are many different defect assemblies which all show very similar structural features as well as energies. These similarities, along with the extreme thinness (<5 nm) of the particles, give rise to the poorly ordered nanocrystalline properties of C-S-H.

However, one should keep in mind that the comparison is a result of classical molecular dynamics, which does not allow for bond breaking and creation. In the real world we might expect spontaneous deprotonation (and protonation) in C-S-H, particularly of silanol groups on the end of chain Q^1 silicates, where calcium is predicted to play a key role [10,20].

4. Discussion

C-S-H is a nanocrystalline material with foil morphology. Due to the foil thickness of <5 nm, which results in 3 main layer, 2 interlayer thick nanofoils [21] surface effects will significantly influence the C-S-H structure, as discussed below.

In this study bulk C-S-H structures were simulated to fit the experimentally measured parameters: MCL, Si-OH/Si, Ca-OH/Ca, and 2H/Si, for Ca/Si ratios ranging from 1.3 to 1.9 (refer to Fig. 5). However, the bulk C-S-H models for Ca/Si ratios of 1.3 and 1.9 diverge from the experimentally observed trend of layer spacing with the Ca/Si ratio (Fig. 6a). In the experimental data (as discussed in Section 3.2 and by Richardson et al. [13]), a reduction in layer spacing is noted until the Ca/Si ratio reaches 1.3, after which it remains constant. The relation of the layer spacing in the model structures to the Ca/Si ratio is found to be influenced by the packing of Ca^{2+} ions within the interlayer (see Figs. 6c and S1). And it was found that simulations for Ca/Si ratios 1.3 to 1.7 with similar Ca^{2+} /dimer values within the C-S-H interlayer have similar layer spacings.

For this automated generation of C-S-H bulk structures there is an inevitable interplay among various properties characterizing the C-S-H bulk structure: Ca/Si , 2H/Si, MCL, Si-OH/Si, and Ca-OH/Ca. Specifically, when Ca/Si , MCL, and 2H/Si ratios remain constant, an increase in Si-OH/Si will lead to an increase in Ca-OH/Ca, consistent with the requirements of charge neutrality. The introduction of an additional silanol group involves transferring a hydrogen atom from a water molecule to a silicate oxygen, resulting in the formation of a new hydroxide ion.

However, experimental data shows that an increase in the Ca/Si ratio corresponds to a decrease in MCL [5], which comes from an increase in unoccupied bridging sites within the silicate chains. If only the bulk is considered (as in the simulations), this reduced abundance of silanol groups leads to a decrease in Si-OH/Si, which should subsequently result in a decrease in OH^- ions and lower Ca-OH/Ca ratios (at constant Ca^{2+} /

dimer). But experimentally [5] it is seen that there is, an inverse correlation between Si-OH/Si and Ca-OH/Ca ratios. These findings suggest that the observed C-S-H characteristics cannot be fully attributed to the bulk C-S-H structure. Instead, they reflect a combination of influences from both bulk and surface.

In Fig. 11 (left), a bulk C-S-H structure is shown with an assumed MCL of 2.5 and a Ca^{2+} /dimer ratio of 0.7. There are two different limiting assumptions for the distribution of silanol groups: either no silanol groups exist within the bulk structure, or all bridging silicates (Q^{2b}) are fully protonated. The first assumption leads to the lowest feasible Ca-OH/Ca ratio, while the second reflects the highest feasible Ca-OH/Ca ratio. This arrangement results in a bulk structure with a Ca/Si ratio of 1.66 and a Ca-OH/Ca ratio of either 0.31 or 0.39.

Next, the C-S-H nanofoil is considered, consisting of three layers with two interlayers of defined thickness [21], while maintaining the bulk structure as previously described. The surface silicate chains in Fig. 11 (right) a-c also possess an MCL of 2.5. On surface (a), no Ca^{2+} ions are adsorbed, resulting in a nanofoil with a Ca/Si ratio of 1.42 and a Ca-OH/Ca ratio of either 0.24 or 0.31. These Ca-OH/Ca values are similar to the experimentally determined ones as seen in Fig. 5 for Ca/Si at 1.4 (Ca-OH/Ca $\approx$ 0.3).

When Ca^{2+} ions occupy the vacant bridging sites on the surface [10], an increase in Ca/Si ratio is observed at the expense of the Ca-OH/Ca ratio (surface b). Surface (c) further considers the co-adsorption of Ca^{2+} and OH^- , as predicted by Casar et al. [21]. This results in an increase in the Ca/Si ratio of the nanofoil to 1.75 [10], corresponding to an increase in Ca-OH/Ca to either 0.42 or 0.47. The experimentally expected Ca-OH/Ca is >0.4 (Fig. 5). So different surfaces can result in Ca/Si ratios from 1.4 to 1.75 with the same bulk structure.

Given the strong impact of the surface it may even be possible that in real structures with $Ca/Si > 1.75$ all of the extra calcium ions are in fact adsorbed on the surface. This would explain why the bulk simulations for C-S-H with Ca/Si or 1.9 give anomalous results when all the extra calcium is assumed to be in the bulk.

The above observations can also explain the anomalous results for the simulated bulk structure with $Ca/Si = 1.3$. While it is predicted that a bulk structure with a MCL = 3.67 and Ca^{2+} /dimer = 0.38 fits the experimentally measured Ca/Si , Ca-OH/Ca, Si-OH/Si and 2H/Si, it underestimates the layer spacing, when compared to bulk structures with higher Ca/Si ratios (Fig. 6a). However, if the surface structure from Fig. 11a is considered, higher Ca^{2+} /dimer values would be required to model a nanofoil with $Ca/Si = 1.3$. Higher Ca^{2+} /dimer values would increase the layer spacing to values similar to the ones observed for bulk structures with Ca/Si 1.5 and 1.7.

In summary, the integration of atomistic modeling with experimentally measured parameters, including C-S-H layer spacings, Ca/Si ratios, MCL, Si-OH/Si, Ca-OH/Ca, and 2H/Si, leads to two key conclusions. First, the defining factor for the bulk C-S-H structure is foremost the MCL. In a bulk C-S-H structure where Ca^{2+} /dimer is a constant and Ca^{2+} ions occupy the vacant bridging sites of the silicate chains [10], the MCL is the defining variable for the Ca/Si ratio of the bulk C-S-H. Second, surface related phenomena play a crucial role in explaining the observed discrepancies in experimental data (Fig. 5), such as different Ca/Si and Ca-OH/Ca ratios for similar MCL values, as shown in Fig. 11.

5. Conclusion

The present paper introduces a new Python code for the automated generation of representative bulk C-S-H structures. The structures are generated with the methodology of the brick model [12], which introduces defects into the tobermorite 14 Å crystal structure. The code carries out an iterative process in which it generates bulk structures that match the measured C-S-H properties (Ca/Si , 2H/Si, mean chain length, Si-OH/Si and Ca-OH/Ca) which were collected by Duque-Redondo et al. [5].

100 structures were generated and simulated for each of the

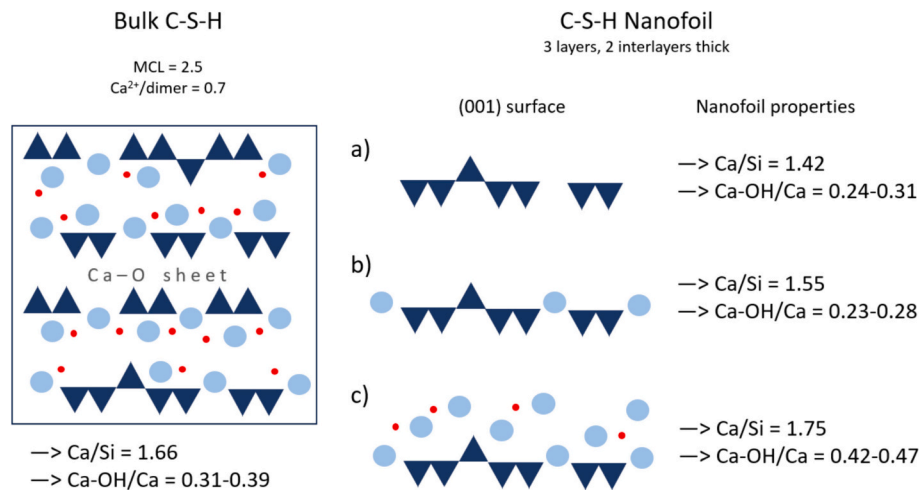


Fig. 11. Impact of the (001) surface on overall properties of the 3 layers, 2 interlayers thick C-S-H nanofoil with a bulk MCL = 2.5 and Ca²⁺/dimer = 0.7. Surfaces (a)-(c) have a MCL = 2.5 with an increasing amount of adsorbed Ca²⁺. Surface (b) considers only Ca²⁺ in the vacant bridging sites, whereby (c) considers co-adsorption of Ca²⁺ and OH⁻ as predicted by Casar et al. [21]. Colour code: dark blue – Si tetrahedra, light blue – Ca²⁺, red – OH⁻. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

investigated Ca/Si ratios (1.3, 1.5, 1.7 and 1.9). The structures differed from each other in the number and distribution of defects. It was found that all structures were in good agreement with the experimentally measured C-S-H characteristics (Ca/Si, MCL, Ca-OH/Ca, Si-OH/Si, 2H/Si). All structures have very similar structural features (pair distribution functions) as well as being energetical very similar, which along with C-S-H's thin nanofoil structure, explains its nanocrystallinity. However, only structures with Ca/Si = 1.5 and 1.7 agree with the experimentally observed trends of the layer spacings with the changing Ca/Si. Due to this we conclude that there is a precise calcium packing criterium of the interlayer (in terms of Ca²⁺/silicate dimer). Therefore, the Ca/Si and Ca-OH/Ca of the C-S-H nanofoil are related to the surface structure as well as mean chain length.

It was shown that with computational insights it is possible to interpret the experimentally obtained data with a higher accuracy. The relatively easy to use automated pyCSH package can produce many structures with equivalent properties at any desired Ca/Si and 2H/Si ratio. This will allow experimentalists to create structures for particular compositions and get a better understanding of the atomistic arrangements. In conjunction with the production of surfaces from Casar et al. [21], which is currently being added to the pyCSH package, a better understanding of additive interactions (e.g. superplasticizers, accelerators) with C-S-H will boost our capacity to better design sustainable cements and concretes in the near future.

The pyCSH code is available for the community at <https://github.com/hegoimanzano/pyCSH.git> and <https://github.com/akmiitm/pyCSH>.

CRediT authorship contribution statement

Ziga Casar: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Jon López-Zorrilla:** Writing – original draft, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation. **Hegoi Manzano:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Eduardo Duque-Redondo:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation. **Aslam Kunhi Mohamed:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Karen Scrivener:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Paul Bowen:** Writing –

review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The code pyCSH is available as open source under GPL-2 license at GitHub <https://github.com/hegoimanzano/pyCSH.git> and <https://github.com/akmiitm/pyCSH.git>.

Acknowledgements

The Swiss National Foundation is acknowledged for financing the contribution of Z.C. (grant no. 179076). This work was in part supported by the “Departamento de Educación, Política Lingüística y Cultura del Gobierno Vasco” (Grant No. IT1458-22), the “Ministerio de Ciencia e Innovación” (Grant No. TED2021-130860B-I00), and the Laboratory for Trans-border Cooperation “Aquitaine-Euskadi Network in Green Concrete and Cement-based Materials” (LTC-Green Concrete). The authors thank for the technical and human support provided by SGiker (UPV/EHU/ ERDF, EU). E.D-R. and J.L.-Z. acknowledge the financial support from the Government of the Basque Country (POS_2020-1-0029 and PRE_2019_1_0025). Dr. Franco Zunino is acknowledged for the fruitful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cemconres.2024.107593>.

References

- [1] K.L. Scrivener, V.M. John, E.M. Gartner, Eco-efficient cements: potential, economically viable solutions for a low-CO₂, cement-based materials industry, *Cem. Concr. Res.* 114 (2018) 2–26, <https://doi.org/10.1016/j.cemconres.2018.03.015>.
- [2] B. Lothenbach, K. Scrivener, R.D. Hooton, Supplementary cementitious materials, *Cem. Concr. Res.* 41 (2011) 1244–1256, <https://doi.org/10.1016/j.cemconres.2010.12.001>.

- [3] O. Heinz, H. Heinz, Cement interfaces: current understanding, challenges, and opportunities, *Langmuir* 37 (2021) 6347–6356, <https://doi.org/10.1021/acs.langmuir.1c00617>.
- [4] A. Kunhi Mohamed, A. Bouibes, M. Bauchy, Z. Casar, Molecular modelling of cementitious materials: current progress and benefits, *RILEM Technical Letters* 7 (2022) 209–219, <https://doi.org/10.21809/rilemtechlett.2022.175>.
- [5] E. Duque-Redondo, P.A. Bonnaud, H. Manzano, A comprehensive review of C-S-H empirical and computational models, their applications, and practical aspects, *Cem. Concr. Res.* 156 (2022) 106784, <https://doi.org/10.1016/j.cemconres.2022.106784>.
- [6] A. Ouzia, K. Scrivener, The needle model: a new model for the main hydration peak of alite, *Cem. Concr. Res.* 115 (2019) 339–360, <https://doi.org/10.1016/j.cemconres.2018.08.005>.
- [7] M.R. Andalibi, A. Kumar, B. Srinivasan, P. Bowen, K. Scrivener, C. Ludwig, A. Testino, On the mesoscale mechanism of synthetic calcium-silicate-hydrate precipitation: a population balance modeling approach, *J. Mater. Chem. A* 6 (2018) 363–373, <https://doi.org/10.1039/c7ta08784e>.
- [8] X. Zhu, I.G. Richardson, Morphology-structural change of C-A-S-H gel in blended cements, *Cem. Concr. Res.* 168 (2023) 107156, <https://doi.org/10.1016/j.cemconres.2023.107156>.
- [9] H. Viallis-Terrisse, A. Nonat, J.C. Petit, Zeta-potential study of calcium silicate hydrates interacting with alkaline cations, *J. Colloid Interface Sci.* 244 (2001) 58–65, <https://doi.org/10.1006/jcis.2001.7897>.
- [10] A. Kumar, B.J. Walder, A. Kunhi Mohamed, A. Hofstetter, B. Srinivasan, A. J. Rossini, K. Scrivener, L. Emsley, P. Bowen, The atomic-level structure of cementitious calcium silicate hydrate, *J. Phys. Chem. C* 121 (2017) 17188–17196, <https://doi.org/10.1021/acs.jpcc.7b02439>.
- [11] E. Bonaccorsi, S. Merlino, A.R. Kampf, The crystal structure of tobermorite 14 Å (plombierite), a C-S-H phase, *J. Am. Ceram. Soc.* 88 (2005) 505–512, <https://doi.org/10.1111/j.1551-2916.2005.00116.x>.
- [12] A. Kunhi Mohamed, S.C. Parker, P. Bowen, S. Galmarini, An atomistic building block description of C-S-H - towards a realistic C-S-H model, *Cem. Concr. Res.* 107 (2018) 221–235, <https://doi.org/10.1016/j.cemconres.2018.01.007>.
- [13] I.G. Richardson, Model structures for C-(A)-S-H(I), *Acta Crystallogr. B Struct. Sci. Cryst. Eng. Mater.* 70 (2014) 903–923, <https://doi.org/10.1107/S2052520614021982>.
- [14] R.J.M. Pellenq, A. Kushima, R. Shahsavari, K.J. Van Vliet, M.J. Buehler, S. Yip, F. J. Ulm, A realistic molecular model of cement hydrates, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 16102–16107, <https://doi.org/10.1073/pnas.0902180106>.
- [15] G. Kovacević, B. Persson, L. Nicoleau, A. Nonat, V. Veryazov, Atomistic modeling of crystal structure of Ca_{1.67}SiHx, *Cem. Concr. Res.* 67 (2015) 197–203, <https://doi.org/10.1016/j.cemconres.2014.09.003>.
- [16] A. Morales-Melgares, Z. Casar, P. Moutzouri, A. Venkatesh, M. Cordova, A. Kunhi Mohamed, K.L. Scrivener, P. Bowen, L. Emsley, Atomic-level structure of zinc-modified cementitious calcium silicate hydrate, *J. Am. Chem. Soc.* 144 (2022) 22915–22924, <https://doi.org/10.1021/jacs.2c06749>.
- [17] A. Kunhi Mohamed, P. Moutzouri, P. Berruyer, B.J. Walder, J. Siramanont, M. Harris, M. Negroni, S.C. Galmarini, S.C. Parker, K.L. Scrivener, L. Emsley, P. Bowen, The atomic-level structure of cementitious calcium aluminate silicate hydrate, *J. Am. Chem. Soc.* 142 (2020) 11060–11071, <https://doi.org/10.1021/jacs.0c02988>.
- [18] P. Faucon, J.M. Delaye, J. Virlet, Molecular dynamics simulation of the structure of calcium silicate hydrates I. Ca₄xSi₆O₁₄2x(OH)₄2x(H₂O)₂(Ox)₁, *J. Solid State Chem.* 127 (1996) 92–97, <https://doi.org/10.1006/jssc.1996.0361>.
- [19] M.J. Abdolhosseini Qomi, K.J. Krakowiak, M. Bauchy, K.L. Stewart, R. Shahsavari, D. Jagannathan, D.B. Brommer, A. Baronnet, M.J. Buehler, S. Yip, F.J. Ulm, K. J. Van Vliet, R.J.M. Pellenq, Combinatorial molecular optimization of cement hydrates, *Nat. Commun.* 5 (2014) 1–10, <https://doi.org/10.1038/ncomms5960>.
- [20] H. Manzano, S. Moeini, F. Marinelli, A.C.T. Van Duin, F.J. Ulm, R.J.M. Pellenq, Confined water dissociation in microporous defective silicates: mechanism, dipole distribution, and impact on substrate properties, *J. Am. Chem. Soc.* 134 (2012) 2208–2215, <https://doi.org/10.1021/ja209152n>.
- [21] Z. Casar, A. Kunhi Mohamed, P. Bowen, K. Scrivener, Atomic-level and surface structure of calcium silicate hydrate nanofoil, *J. Phys. Chem. C* 127 (2023) 18652–18661, <https://doi.org/10.1021/acs.jpcc.3c03350>.
- [22] A.C.A. Muller, K.L. Scrivener, A.M. Gajewicz, P.J. McDonald, Densification of C-S-H measured by ¹H NMR relaxometry, *J. Phys. Chem. C* 117 (2013) 403–412, <https://doi.org/10.1021/jp3102964>.
- [23] J.L.F. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, *J. Chem. Phys.* 123 (2005), <https://doi.org/10.1063/1.2121687>.
- [24] M. Valavi, Z. Casar, A. Kunhi, P. Bowen, S. Galmarini, Molecular dynamic simulations of cementitious systems using a newly developed force field suite ERICA FF, *Cem. Concr. Res.* 154 (2022) 106712, <https://doi.org/10.1016/j.cemconres.2022.106712>.
- [25] P.J. Mitchell, D. Fincham, Shell model simulations by adiabatic dynamics, *J. Phys. Condens. Matter* 5 (1993) 1031–1038, <https://doi.org/10.1088/0953-8984/5/8/006>.
- [26] A.P. Thompson, H.M. Aktulga, R. Berger, D.S. Bolintineanu, W.M. Brown, P. S. Crozier, P.J. In't Veld, A. Kohlmeyer, S.G. Moore, T.D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, S.J. Plimpton, LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Comput. Phys. Commun.* 271 (2022) 108171, <https://doi.org/10.1016/j.cpc.2021.108171>.
- [27] R.W. Hockney, J.W. Eastwood, *Computer Simulation Using Particles*, CRC Press, 2021, <https://doi.org/10.1201/9780367806934>.
- [28] W.G. Hoover, Constant-pressure equations of motion, *Phys. Rev. A (Coll Park)* 34 (1986) 2499–2500, <https://doi.org/10.1103/PhysRevA.34.2499>.
- [29] D.J. Evans, B.L. Holian, The Nose-Hoover thermostat, *J. Chem. Phys.* 83 (1985) 4069–4074, <https://doi.org/10.1063/1.449071>.
- [30] M. Brehm, M. Thomas, S. Gehrke, B. Kirchner, TRAVIS—A free analyzer for trajectories from molecular simulation, *J. Chem. Phys.* 152 (2020), <https://doi.org/10.1063/5.0005078>.
- [31] M. Brehm, B. Kirchner, TRAVIS - a free analyzer and visualizer for Monte Carlo and molecular dynamics trajectories, *J. Chem. Inf. Model.* 51 (2011) 2007–2023, <https://doi.org/10.1021/ci200217w>.
- [32] T. Egami, S.J.L. Billinge, *Underneath the Bragg Peaks Structural Analysis of Complex Materials*, Elsevier, Oxford, 2003.
- [33] M. Harris, G. Simpson, K. Scrivener, P. Bowen, A method for the reliable and reproducible precipitation of phase pure high Ca/Si ratio (>1.5) synthetic calcium silicate hydrates (C-S-H), *Cem. Concr. Res.* 151 (2022) 106623, <https://doi.org/10.1016/j.cemconres.2021.106623>.
- [34] C. Labbez, B. Jönsson, I. Pochard, A. Nonat, B. Cabane, Surface charge density and electrokinetic potential of highly charged minerals: experiments and Monte Carlo simulations on calcium silicate hydrate, *J. Phys. Chem. B* 110 (2006) 9219–9230, <https://doi.org/10.1021/jp057096+>.
- [35] L. Black, K. Garbev, P. Stemmermann, K.R. Hallam, G.C. Allen, Characterisation of crystalline C-S-H phases by X-ray photoelectron spectroscopy, *Cem. Concr. Res.* 33 (2003) 899–911, [https://doi.org/10.1016/S0008-8846\(02\)01089-X](https://doi.org/10.1016/S0008-8846(02)01089-X).
- [36] K. Garbev, G. Beuchle, M. Bornfeld, L. Black, P. Stemmermann, Cell dimensions and composition of nanocrystalline calcium silicate hydrate solid solutions. Part 1: synchrotron-based X-ray diffraction, *J. Am. Ceram. Soc.* 91 (2008) 3005–3014, <https://doi.org/10.1111/j.1551-2916.2008.02484.x>.
- [37] H. Matsuyama, J.F. Young, Synthesis of calcium silicate hydrate/polymer complexes: part I. Anionic and nonionic polymers, *J. Mater. Res.* 14 (1999) 3379–3388, <https://doi.org/10.1557/JMR.1999.0458>.
- [38] X. Cong, R. James Kirkpatrick, 29Si MAS NMR study of the structure of calcium silicate hydrate, *Adv. Cem. Bas. Mat.* 3 (1996) 144–156, [https://doi.org/10.1016/S1065-7355\(96\)90046-2](https://doi.org/10.1016/S1065-7355(96)90046-2).
- [39] M.F. Döpke, J. Lützenkirchen, O.A. Moulto, B. Siboulet, J.-F. Dufreche, J. T. Padding, R. Hartkamp, Preferential adsorption in mixed electrolytes confined by charged amorphous silica, *J. Phys. Chem. C* 123 (2019) 16711–16720, <https://doi.org/10.1021/acs.jpcc.9b02975>.
- [40] S. Grangeon, F. Claret, Y. Linard, C. Chiaberge, X-ray diffraction: a powerful tool to probe and understand the structure of nanocrystalline calcium silicate hydrates, *Acta Crystallogr. B Struct. Sci. Cryst. Eng. Mater.* 69 (2013) 465–473, <https://doi.org/10.1107/S2052519213021155>.
- [41] G. Renaudin, J. Russias, F. Leroux, F. Frizon, C. Cau-dit-Coumes, Structural characterization of C-S-H and C-A-S-H samples-part I: long-range order investigated by Rietveld analyses, *J. Solid State Chem.* 182 (2009) 3312–3319, <https://doi.org/10.1016/j.jssc.2009.09.026>.
- [42] S. Masoumi, D. Ebrahimi, H. Valipour, M.J. Abdolhosseini Qomi, Nanolayered attributes of calcium-silicate-hydrate gels, *J. Am. Ceram. Soc.* 103 (2020) 541–557, <https://doi.org/10.1111/jace.16750>.
- [43] S. Grangeon, A. Fernandez-Martinez, A. Baronnet, N. Marty, A. Poulain, E. Elkaim, C. Roos, S. Gaboreau, P. Henocq, F. Claret, Quantitative X-ray pair distribution function analysis of nanocrystalline calcium silicate hydrates: a contribution to the understanding of cement chemistry, *J. Appl. Cryst.* 50 (2017) 14–21, <https://doi.org/10.1107/S1600576716017404>.