

ARTICLE OPEN



A machine-learned interatomic potential for silica and its relation to empirical models

Linus C. Erhard¹, Jochen Rohrer¹, Karsten Albe¹ and Volker L. Deringer²

Silica (SiO₂) is an abundant material with a wide range of applications. Despite much progress, the atomistic modelling of the different forms of silica has remained a challenge. Here we show that by combining density-functional theory at the SCAN functional level with machine-learning-based interatomic potential fitting, a range of condensed phases of silica can be accurately described. We present a Gaussian approximation potential model that achieves high accuracy for the thermodynamic properties of the crystalline phases, and we compare its performance (and performance–cost trade-off) with that of multiple empirically fitted interatomic potentials for silica. We also include amorphous phases, assessing the ability of the potentials to describe structures of melt-quenched glassy silica, their energetic stability, and the high-pressure structural transition to a mainly sixfold-coordinated phase. We suggest that rather than standing on their own, machine-learned potentials for silica may be used in conjunction with suitable empirical models, each having a distinct role and complementing the other, by combining the advantages of the long simulation times afforded by empirical potentials and the near-quantum-mechanical accuracy of machine-learned potentials. This way, our work is expected to advance atomistic simulations of this key material and to benefit further computational studies in the field.

npj Computational Materials (2022)8:90; <https://doi.org/10.1038/s41524-022-00768-w>

INTRODUCTION

Silica (SiO₂) is among the most common materials on Earth¹, with applications in window glass, oscillators², aerogels³, and catalysts⁴. This wide range of applications is due, in parts, to the existence of various polymorphs with different properties. The ambient form of crystalline silica, α -quartz, is a piezoelectric material⁵; the α -cristobalite polymorph has a negative Poisson ratio indicating unconventional elastic properties⁶. High-pressure phases like stishovite are important for understanding reactions in the Earth's mantle⁷. Low-density zeolite polymorphs are used as catalysts in the petrochemical industry⁸, and amorphous silica finds application in optical fibres⁹.

Atomistic modelling by means of interatomic potentials has a long history for silica, reflected in the sheer amount of models and parameterisations available in the literature. Among the arguably most successful silica potentials is that of van Beest, Kramer, and van Santen (BKS)¹⁰—a two-body energy function fitted for α -quartz that is also applicable to silica glasses^{11–13}. This potential, however, does not correctly predict the phase boundaries of the different polymorphs¹⁴. Re-parameterisations of BKS-type potentials were developed for nanoclusters¹⁵ or for an improved description of amorphous silica¹⁶. There are also more complex potentials implementing polarisation terms¹⁷, charge transfer^{18–20}, and three-body effects^{21–24}. For example, the Tangney–Scandolo potential¹⁷, which includes self-consistently calculated dipoles, improves the description of many properties of tetrahedral silica compared to BKS²⁵. However, these potentials typically describe only a subset of the crystalline and amorphous forms, and often lack a fully accurate description of the (relative) energetic stability of different polymorphs, resulting in inaccurate temperature–pressure phase diagrams.

Machine learning (ML) based approaches have recently been established as an alternative route to interatomic potential

fitting^{26–30}. ML potentials are trained with quantum-mechanical reference data, typically obtained from density-functional theory (DFT), and reproduce the associated potential-energy surface at a substantially lower computational cost—thereby making much longer time scales and larger length scales accessible. Emerging applications of ML potentials range from modelling structural transitions in disordered matter^{31–33} to battery materials³⁴ or molecular crystal-structure prediction³⁵. Oxide materials are becoming amenable to ML potential modelling as well—as demonstrated, for example, in an early study for zinc oxide³⁶, more recently for titania (TiO₂, ref. 37) and hafnia (HfO₂, ref. 38), and for surface studies of IrO₂³⁹. In the case of silica, Novikov and Shapeev reported an ML potential with a view to analyse the role of long-range interactions in the construction of the model; however, this study was focused on the α -quartz polymorph⁴⁰. Two recent studies described and validated ML potential fits for liquid silica^{41,42}.

In the present work, we surpass the limitations of currently available silica potentials by training an ML potential for a range of crystalline polymorphs as well as models for liquid and amorphous phases. The potential shows high thermodynamic accuracy, enabled by the use of the strongly constrained and appropriately normed (SCAN) functional of Sun et al.⁴³ for the computation of reference data. Thus, the ML potential shows superior behaviour in predictions of phase coexistence lines and high-pressure structural transitions compared to several existing, empirically fitted interatomic potentials. The results also allow us to discuss, more generally, the criteria for evaluating the quality of an ML potential—thereby placing it in the context of more established simulation methods in terms of their computational cost, predictive power, and applicability to practical research questions.

¹Institute of Materials Science, Technische Universität Darmstadt, Otto-Berndt-Strasse 3, 64287 Darmstadt, Germany. ²Department of Chemistry, Inorganic Chemistry Laboratory, University of Oxford, Oxford OX1 3QR, UK. ✉email: rohrer@mm.tu-darmstadt.de; albe@mm.tu-darmstadt.de; volker.deringer@chem.ox.ac.uk

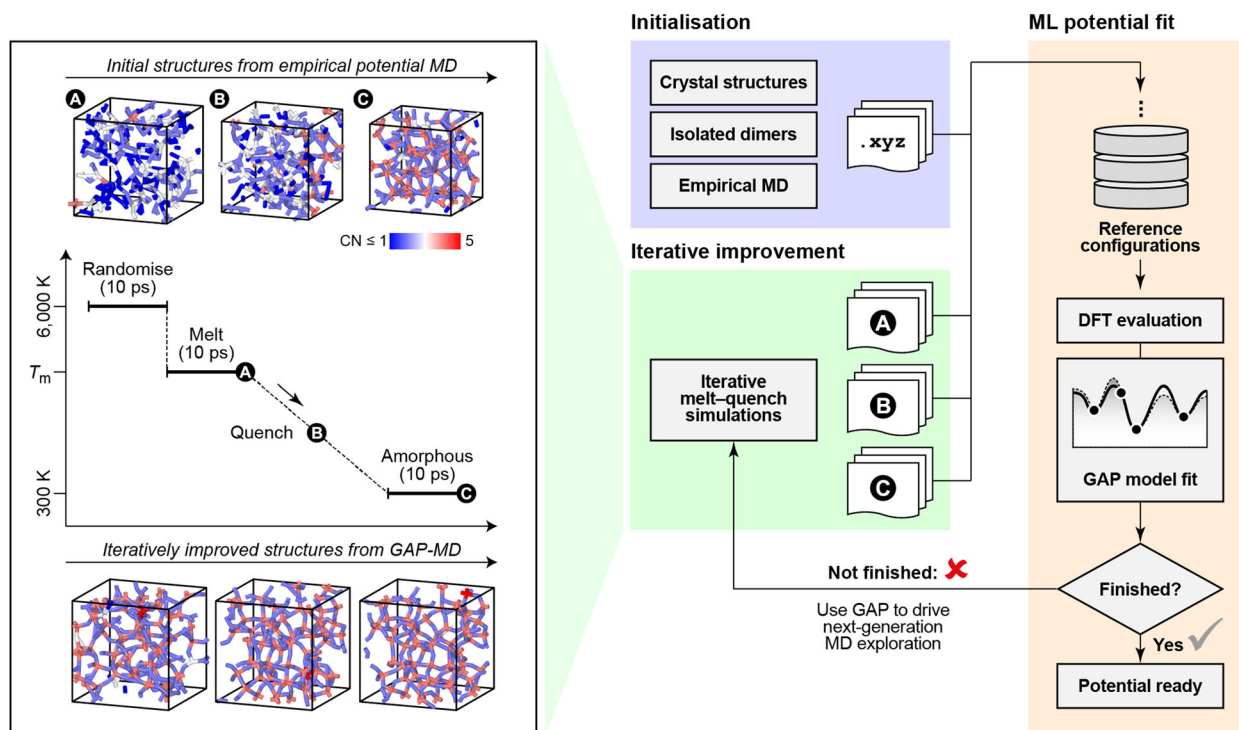


Fig. 1 Iterative exploration and potential fitting. We generated a GAP ML model for silica by fitting it to a gradually extended set of DFT reference data, as is now increasingly done in the field⁴⁴. Here, we initialised the process with simulation cells describing distorted crystalline structures, isolated dimers, and snapshots of molecular-dynamics (MD) simulations with a fast empirical potential (blue), and then alternated single-point DFT evaluation and GAP fitting (orange) with GAP-driven melt-quench simulations (green) until the database was considered complete. The left-hand side of the figure illustrates the melt-quench protocol and includes selected snapshots (at points A, B, and C during the simulations, respectively). In some iterations, we started from β -cristobalite rather than from randomised structures; some used NVT (constant-volume) and others used NPT (constant-pressure) simulations. Full details of the protocol are given in the Supplementary Information. The right-hand side of the figure is drawn in the style of ref.¹⁰⁴, where the same colour-coding is used to identify the different components of an iterative random structure searching (GAP-RSS) protocol; see also references therein.

RESULTS

A machine-learned interatomic potential for silica

We fitted a ML potential for silica using the Gaussian approximation potential (GAP) framework^{27,44}. Initial versions were fitted to reference data obtained with the computationally inexpensive PBEsol functional for DFT; the final potential version was fitted to data obtained using the SCAN functional, which our results show to provide an accurate description of the material (see Supplementary Fig. 1 for details). The use of existing reference structures re-labelled with higher-level data has been shown previously to be successful in GAP fitting (for example, LDA \rightarrow dispersion-corrected DFT for carbon⁴⁵). In the following, we will analyse the effect of changing the computational level more closely.

Figure 1 illustrates the development of the reference database, which is one of the three defining ingredients of an ML potential (in addition to the representation of the structure and the regression task itself)⁴⁶. As in previous work with GAP⁴⁷ and other ML potentials⁴⁸, we used iterative fitting to explore the potential-energy surface. Here, we started the iterative cycles by carrying out molecular-dynamics (MD) simulations with a fast empirical potential, so as to sample a range of structures that require only single-point DFT computations for the generation of reference data. Rather than in active learning, where a decision criterion is implemented within the cycle, we here use 'batch' learning, where sets of structures are fed back into the reference database in multiple iterations until the composition of the database is deemed sufficient. Details of the protocol are given in the Supplementary Material. Figure 1 already indicates (if highly qualitatively) the progressive improvement of the model: initial liquid structures, generated with an empirical potential, contain

many low-coordinated atoms (top left), whereas the structures from the last GAP-MD iteration are largely 'correctly' coordinated (4-fold silicon, 2-fold oxygen; bottom left), even in the liquid phase. The latter may be partly a consequence of constant-volume (NVT) simulations being carried out in this case.

Figure 2a shows the energy-volume curves of several silica polymorphs as computed with DFT (using the SCAN functional) and with the GAP. Accurately reproducing these curves is an important initial quality indicator for the thermodynamic behaviour of the potential: the 0 K energy-volume diagram of the crystalline phases are highly relevant to both pressure- and temperature-induced phase transitions. We find that the GAP (lines in Fig. 2a) reproduces the SCAN-DFT data (symbols) very well; as a consequence, the derived hydrostatic bulk moduli and the ground-state volumes closely match the SCAN reference, with the remaining deviations being indicative of the error in the ML model. In addition, both the SCAN and GAP predictions are in good agreement with earlier experimental observations (Table 1).

In the following, we present further tests of the GAP for elastic constants of various polymorphs, as a more sensitive way of assessing the quality of predicted forces in particular. Here, the focus is on a comparison to experimental data and on contrasting the accuracy of the GAP with that of existing, empirically fitted potentials. In particular, we have chosen the models of Broughton²³, Vashista²¹, Munetoh²⁴, and the BKS¹⁰ and the Carré, Horbach, Ispas, and Kob (CHIK)¹⁶ potentials. In the case of BKS, we used a cut-off of 15 Å; for simulations at high temperatures, we also included a harmonic term for short distances according to Shen et al.⁴⁹. For CHIK, which has been explicitly fitted to data for amorphous and liquid silica, we used a cut-off of 6.5 Å for

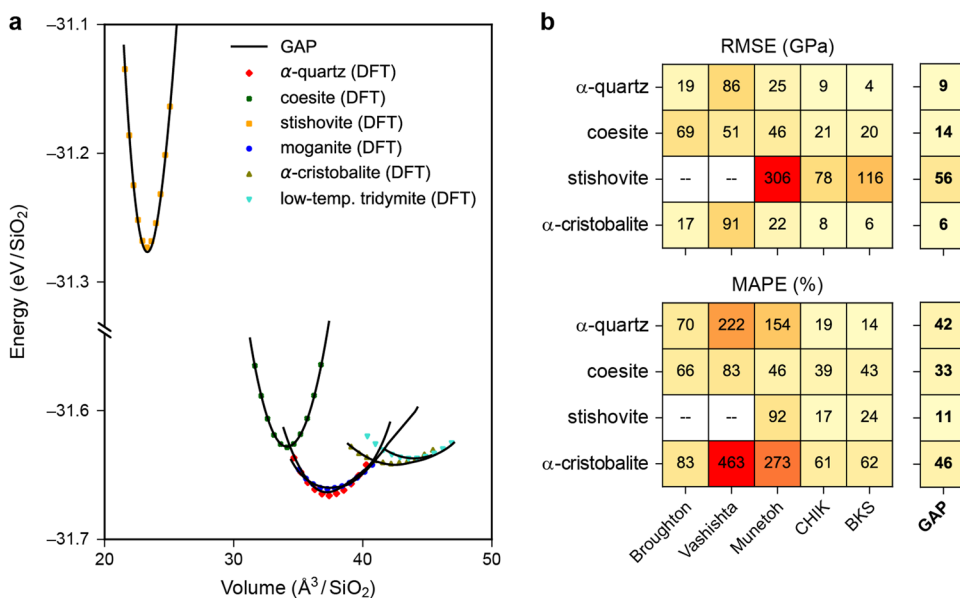


Fig. 2 Elastic properties. **a** Energy–volume curves of different silica polymorphs. Solid black lines correspond to calculations using the GAP model for silica developed in the present work. The GAP agrees well with the SCAN-DFT data (coloured symbols). For the low-temperature tridymite structure, however, the GAP predicts that it becomes unstable at certain strains and data are therefore only shown in the stability region. **b** Root mean square error (RMSE) and mean absolute percentage error (MAPE) of elastic constants as predicted by various potentials against experimental data for α -quartz¹⁰⁵, α -cristobalite⁶, coesite¹⁰⁶, and stishovite¹⁰⁷. The Vashishta and Broughton potentials do not predict a stable stishovite structure (“--”). Note that large deviations from a small reference value can lead to errors exceeding 100%. Colour coding from yellow to red emphasises the magnitude of the respective error.

Table 1. Ground-state volumes (V) and bulk moduli (K) for silica polymorphs^a.

	K (GPa)			V (Å ³ /SiO ₂)		
	Expt.	SCAN	GAP	Expt.	SCAN	GAP
α -quartz	37.7 ⁹⁴	39.7	36.6	37.8 ⁹⁵	37.4	37.2
coesite	94.0 ⁹⁶	106.8	106.4	33.9 ⁹⁷	34.1	34.1
stishovite	295.0 ⁹⁶	300.1	307.5	22.9 ⁹⁸	23.3	23.3
chabazite	—	42.3	50.8	64.9 ⁹⁹	65.4	65.7
α -cristobalite	16.4 ⁶	15.5	13.5	42.8 ¹⁰⁰	42.3	42.5
moganite	32.2 ¹⁰¹	26.8	25.2	38.1 ¹⁰²	37.4	37.5
low-tridymite	—	19.0	18.7	44.0 ¹⁰³	43.8	44.0

^aA Birch–Murnaghan fit was used for hydrostatically deformed cells.

simulation of the amorphous and liquid phases and 10 Å for crystalline phases as suggested in the original work¹⁶.

In Fig. 2b, we report the root mean square error (RMSE) and the mean absolute percentage error (MAPE) of the components of the stiffness tensor as calculated using the various interatomic potentials, with experimental values taken as benchmarks. We find that the GAP model shows the lowest total RMSE and MAPE among all potentials considered here: this is consistent with the fact that ML potentials commonly include many (distorted) copies of crystalline-phase unit cells in the reference database and are therefore expected to accurately describe the associated regions of the potential-energy surface. Moreover, larger discrepancies of the GAP in the RMSE can be associated with large absolute values of elastic constants (particularly for stishovite), and in these cases the corresponding MAPEs are reasonable. Hence, studying both RMSE and MAPE of the elastic constants provides a more detailed view of the quality of the potential than looking at only one or the other.

Phonons and phase diagrams

Phonons are important fundamental properties of materials, and their computational demands (in terms of resources, and also required accuracy) make them an interesting target for study with ML potentials. For example, it was shown that, given a suitably designed reference database, GAP models for silicon can be brought to within 0.1–0.2 THz prediction error for a range of existing and hypothetical crystalline allotropes⁵⁰. Phonons may also be used to assess the quality of an ML potential as compared to empirically fitted ones—as done, for example, in a recent study for carbon⁴⁵. We continue our assessment of silica potentials here by studying phonon dispersion curves as well as the prediction of phase diagrams, which requires accurate phonon data.

Figure 3 shows the phonon dispersion curve of α -quartz predicted by the various potentials, compared to experimental results from ref. ⁵¹. In line with earlier work on other materials^{45,50}, the GAP shows high accuracy and closely matches previous experimental data (Fig. 3a). A small deviation in the highest mode at Γ (experiment: 6.25 ± 0.14 THz; ref. ⁵¹) is the only exception. The situation is more varied for the faster, empirical potentials tested here (Fig. 3b–f). For both the Broughton and the Vashishta potentials, the phonon dispersion seems to be scaled by a near-constant factor compared to experimental results; the Munetoh and BKS potentials differ in aspects of the qualitative description. For example, at the K point, the Munetoh potential predicts that the three lowest modes nearly coincide, at variance with the experiment. Although the CHIK potential gives a reasonable result, the GAP reproduces the experimental data with the highest accuracy among the tested potentials.

Beyond α -quartz, Supplementary Fig. 2 shows GAP-computed phonon dispersion curves for the α -cristobalite and stishovite polymorphs. In both cases, the GAP results agree reasonably well with experiments^{52,53}, although there are some deviations at the M-point for α -cristobalite, and at higher frequencies for stishovite. For high-pressure stishovite, we observe the vanishing of the lowest acoustic phonon at around 65 GPa, and the vanishing of the lowest optical mode at the Γ -point for pressures above 110

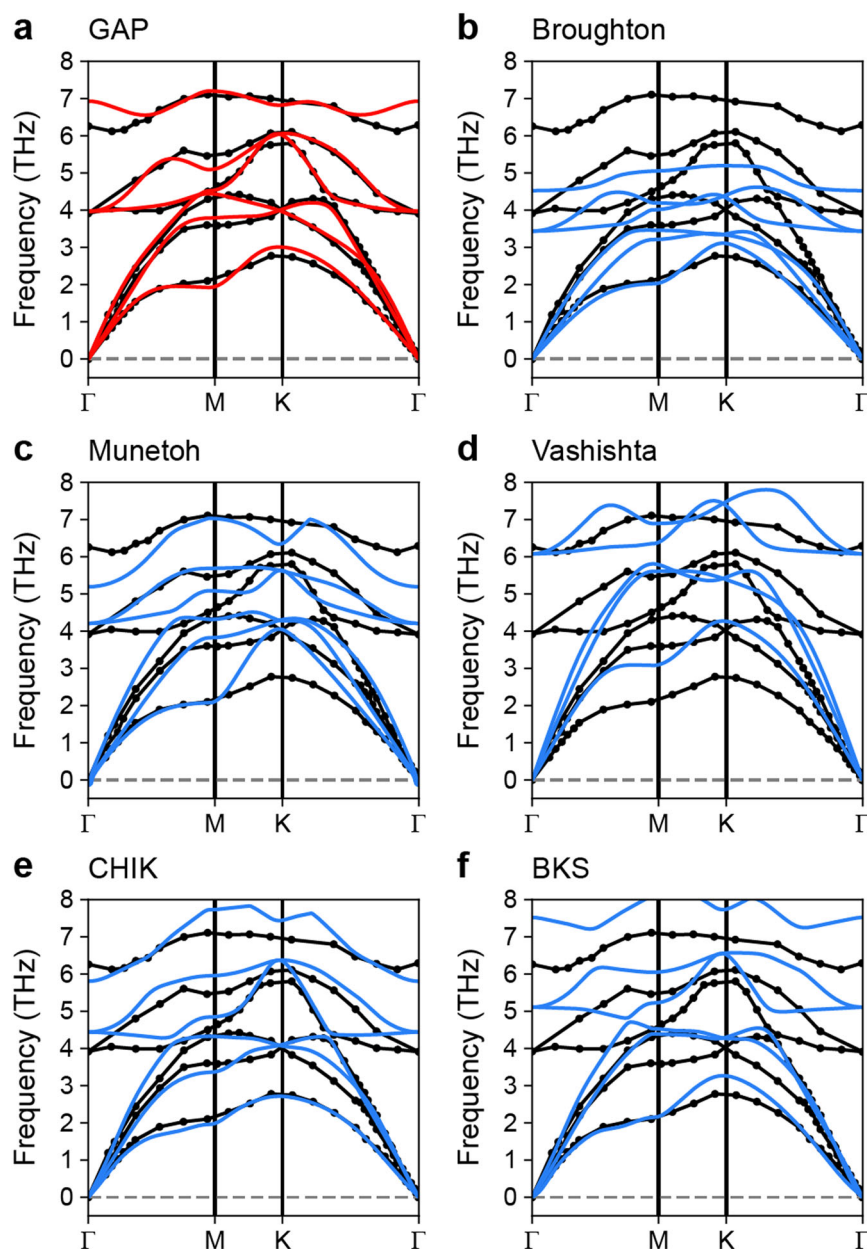


Fig. 3 Phonon dispersions of α -quartz. Phonon dispersion curve of α -quartz, along the $\Gamma \rightarrow M \rightarrow K \rightarrow \Gamma$ pathway through reciprocal space, as computed using the **a** GAP, **b** Broughton, **c** Munetoh, **d** Vashishta, **e** CHIK, and **f** BKS potentials (red and blue lines, respectively), compared to experimental data by Dorner et al. (taken from ref. ⁵¹; black symbols connected by lines to guide the eye).

GPa, in good agreement with the computational study of Tsuchiya et al. (ref. ⁵⁴).

In Fig. 4, we assess the thermodynamic behaviour of various potentials by considering the predicted phase stability of α -quartz, coesite, and stishovite as a function of temperature and pressure. The phase diagrams shown in Fig. 4 were constructed within the quasi-harmonic approximation (Methods section) and are compared to experimental data from refs. ⁵⁵ and ⁵⁶. The Broughton and Vashishta potentials do not yield a stable stishovite structure, which does therefore not appear in the corresponding phase diagram, and neither is the α -quartz–coesite coexistence line predicted correctly with either of the potentials. The Munetoh potential gives a good description of the high-pressure coesite–stishovite boundary, but is rather inaccurate for that of α -quartz–coesite. The CHIK potential predicts the coesite–stishovite coexistence line at too high pressures; moreover, it does not

predict α -quartz to be stable at ambient conditions. This behaviour is similar to that of the BKS potential, as already analysed in earlier studies¹⁴. The BKS potential predicts stishovite to be stable at ambient conditions; instead, coesite and α -quartz become stable at negative pressures. The GAP, on the other hand, is capable of accurately describing both coexistence lines.

We emphasise that we do not include β -quartz and other higher-temperature polymorphs here, due to issues with imaginary modes and the need for a more advanced model to describe the thermodynamics. The α – β quartz coexistence line near ambient pressure is therefore not included in the diagrams shown in Fig. 4.

Structure and energetics of disordered silica

An important task for interatomic potential models is the accurate atomistic simulation of amorphous phases, which is often difficult

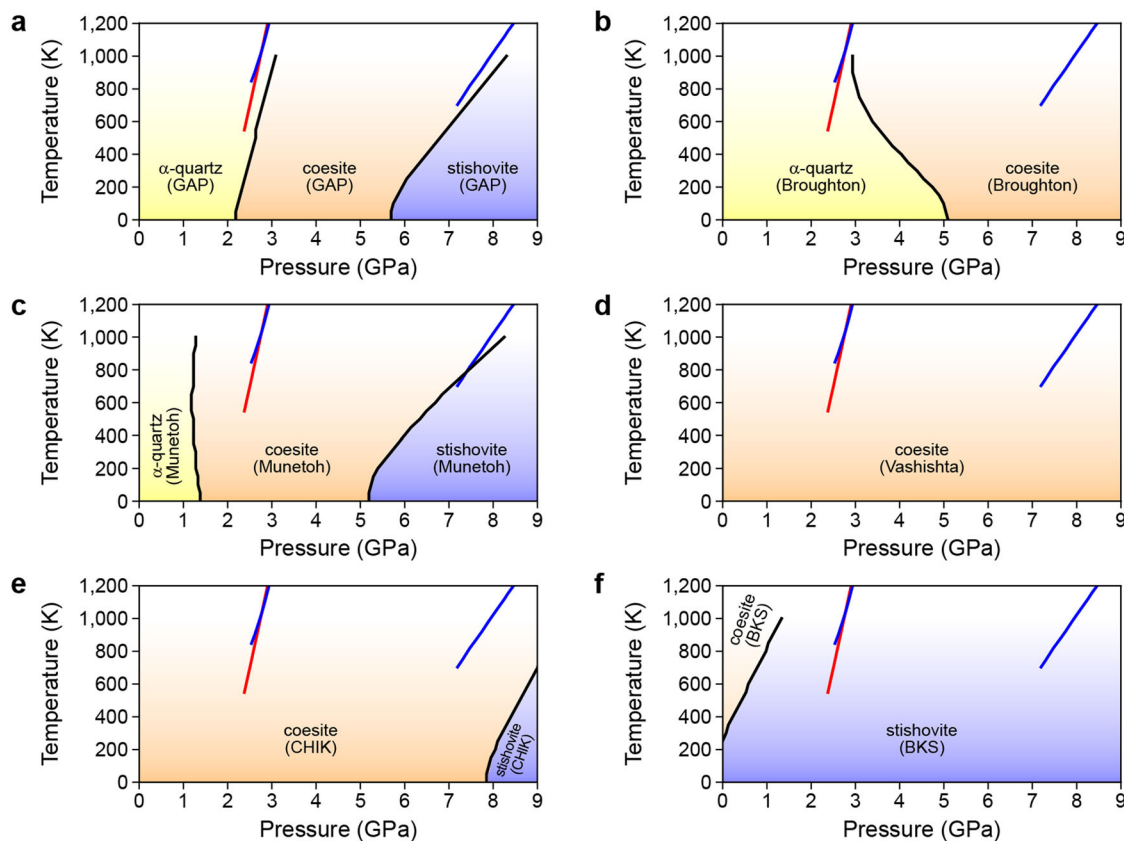


Fig. 4 Phase diagrams. Stability regions of the three silica polymorphs α -quartz, coesite, and stishovite were determined using the quasi-harmonic approximation using the **a** GAP, **b** Broughton, **c** Munetoh, **d** Vashishta, **e** CHIK, **f** BKS potential. The red and blue lines are the experimental transition lines from Bohlen and Boettcher⁵⁵ and Akaogi et al.⁵⁶, respectively. Since we used the quasi-harmonic approximation, the GAP data are only shown up to 1,000 K.

(or even impossible) for DFT-based methods because of their computational cost. In Fig. 5a–c, we assess the ability of various interatomic potentials to describe amorphous silica in rapid quench simulations. There is a natural limitation to such tests because ML potentials are computationally much more demanding than their empirically fitted counterparts—therefore, extremely slow quenches, e.g., as reported in ref.⁵⁷ (with simulation times of up to 1 μ s with BKS), are challenging for GAP or any similar ML potential. This does not, however, preclude the usefulness of faster quenching as a way to assess the predictions of various potential models side-by-side. In this sense, the present tests provide a “slice” through the entire parameter space that controls the properties of amorphous silica.

Figure 5a shows the total structure factor, $S(q)$, of amorphous SiO_2 structural models created by empirical interatomic potentials. Whilst simulations with all potentials qualitatively reproduce the experimental features, there are significant differences in detail. The CHIK potential predicts the position of the first sharp diffraction peak (FSDP; highlighted by arrows) with high accuracy and also the peak height, which is close to experiment. Our GAP model underestimates the height of the FSDP, however, its results coincide well with experiment for higher values of q . The $S(q)$ for structures simulated by the Vashishta and BKS potentials agrees well with experiment; however, in both cases the first peak is shifted to slightly higher q values. Our simulation using the Munetoh potential reproduces the experimental structure factor less accurately: the first peak is significantly underestimated, and the second peak is shifted to lower q . (The GAP also underestimates the first peak, but reproduces the second one accurately.)

Despite the importance of the structure factor as a primary means of validating computed amorphous structures, other quantities need to be considered for a full picture. Figure 5c documents the ability of the GAP to considerably reduce the number of defects in structures obtained from melt–quench simulations. We list the frequency of different coordination numbers of silicon and oxygen atoms in the amorphous structural models. For a realistic amorphous structure, one would expect a low number of wrongly coordinated atoms. This means that nearly all silicon (oxygen) atoms should be 4-fold (2-fold) coordinated, respectively, consistent with the structures of the ambient-pressure crystalline polymorphs. This expectation is not universally met by all potentials tested: in particular, melt-quenching with the Munetoh potential leads to a structure with a high proportion of wrongly coordinated atoms (>10%). The Vashishta, CHIK, BKS, and GAP potentials all yield significantly fewer wrongly coordinated atoms in the amorphous phase. The CHIK potential predicts 50% fewer defects than the Vashishta potential for the quenching protocol we use, and the GAP yields almost no wrongly coordinated atoms.

Further insight into the simulated amorphous phases can be gained by inspecting their relative energetic stability, expressed through the excess energy (the energy of a given structure relative to the more stable crystalline phase). This is not a trivial task, because different interatomic potential models will most likely predict different *local* energy minima, and therefore quantities such as the excess energy will depend on two aspects: (i) the method, typically a specific force field, used to generate the structure, and (ii) the method used to evaluate the energy of the quenched structure. These methods are often the same in practice, but they need not be. Reference data are available in

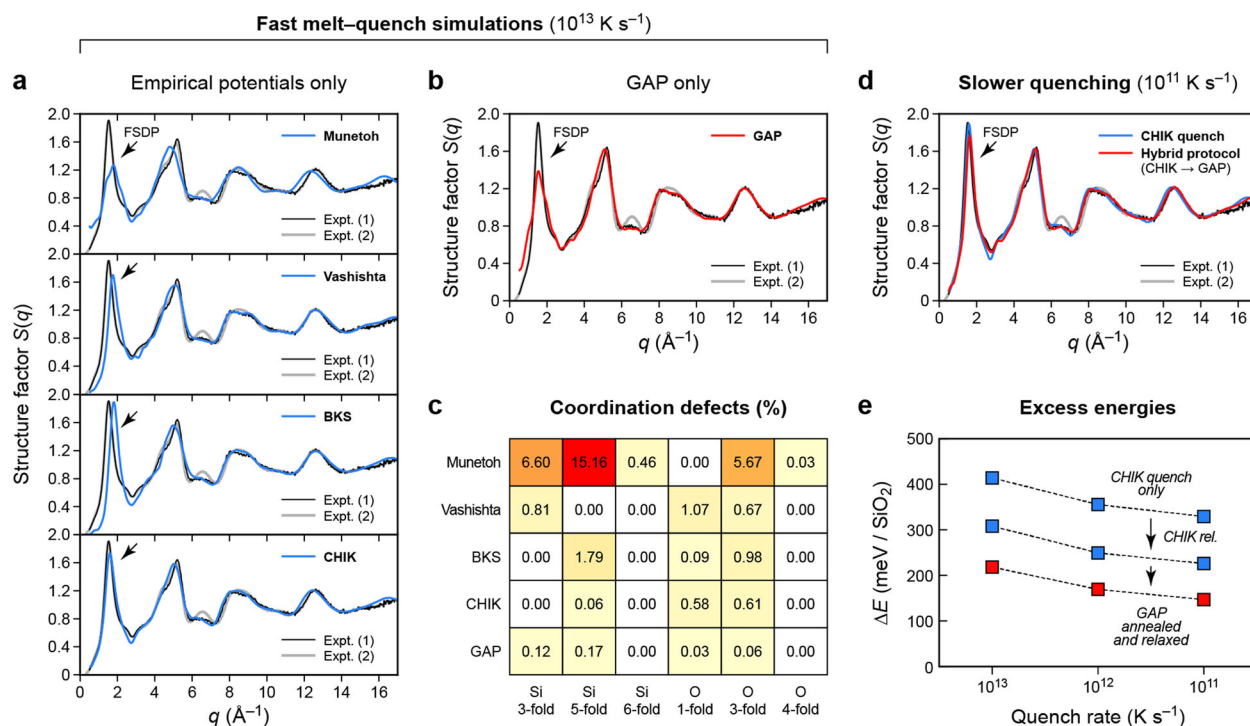


Fig. 5 Amorphous silica. **a** Simulated X-ray structure factor of amorphous silica at 300 K, for melt-quenched structures generated using different empirical interatomic potentials (blue lines) and compared to experimental data (“1”; ref. 72, and “2”; ref. 108). An arrow indicates the location of the first sharp diffraction peak (FSDP). **b** Same for a melt-quenched structure generated by the GAP (red line) and compared to the same experimental reference data^{72,108}. **c** Proportion of coordination defects in the final amorphous structures (300 K) generated by melt-and-quench MD with quenching rates of 10^{13} K/s using several potentials. The number of wrongly-coordinated silicon atoms (defined as those with fewer or more than four neighbours) were related to the total number of silicon atoms, and the same was done for wrongly-coordinated oxygen atoms (fewer or more than two neighbours). **d** As panel **b**, but now for slower quenching simulations, including a “hybrid protocol” involving a CHIK quench at 10^{11} K s^{-1} and subsequent GAP annealing. **e** Excess energies for simulations using the hybrid quenching approach at different quench rates, comparing the structures directly after the CHIK quench as well as additional CHIK relaxation (blue), and after GAP annealing and relaxation (red). All energies were computed with the GAP for consistency and are referenced to α -quartz.

Table 2. Excess energies of small-scale structural models of amorphous silica^a.

	ΔE (meV/SiO ₂)				
	Single-point evaluation			Re-relaxed	
	SCAN	GAP	CHIK	GAP	CHIK
Munetoh quench	1,334 ± 79	1,314 ± 74	1,376 ± 30	657 ± 25	678 ± 30
Vashishta quench	310 ± 97	302 ± 82	423 ± 32	248 ± 65	315 ± 32
BKS quench	361 ± 19	364 ± 18	385 ± 15	256 ± 16	315 ± 15
CHIK quench	311 ± 20	317 ± 14	321 ± 16	226 ± 12	(321 ± 16) ^b
GAP quench	235 ± 15	234 ± 16	515 ± 15	(234 ± 16) ^b	399 ± 15

^aThe structure was obtained by fast quenching using different potentials (as mean and standard deviation for five independent simulations). We report $\Delta E = E_{\text{amo}} - E_{\alpha\text{-quartz}}$, where E_{amo} is the energy of a structure quenched and relaxed with the stated potential and then evaluated in a single-point computation (using either SCAN-DFT, GAP, or CHIK), or after re-relaxation (using either GAP or CHIK). ^bIn these cases, the initial relaxation had been done with the same potential, and therefore the ‘single-point evaluation’ values are repeated.

the experimental literature: the enthalpy difference according to literature is between 78 and 131 meV/SiO₂ for bulk silica glass⁵⁸, while the standard enthalpy of deposited silica glass films and ‘snow’, which are presumably more structurally disordered, can be up to 1 eV/SiO₂ higher⁵⁹.

In Table 2, we present an analysis of the energetics of small-scale amorphous structural models (192 atoms), prepared with the various potentials by melt-quench simulations and additional full structural optimisation. These results emphasise the range of disordered configurations that are explored by the different

potentials, with variations of more than 1 eV/SiO₂. The single-point SCAN DFT energy is lowest for the structures prepared using the GAP; this is perhaps not too surprising, because the GAP has been fitted to data at the same computational level. Still, the comparison across different methods is instructive—emphasising that the structures generated by the Munetoh potential are highly unfavourable (more than 1 eV/SiO₂ above α -quartz), in line with their large defect count (Fig. 5c). Additionally, the table emphasises the high accuracy of the GAP with respect to SCAN,

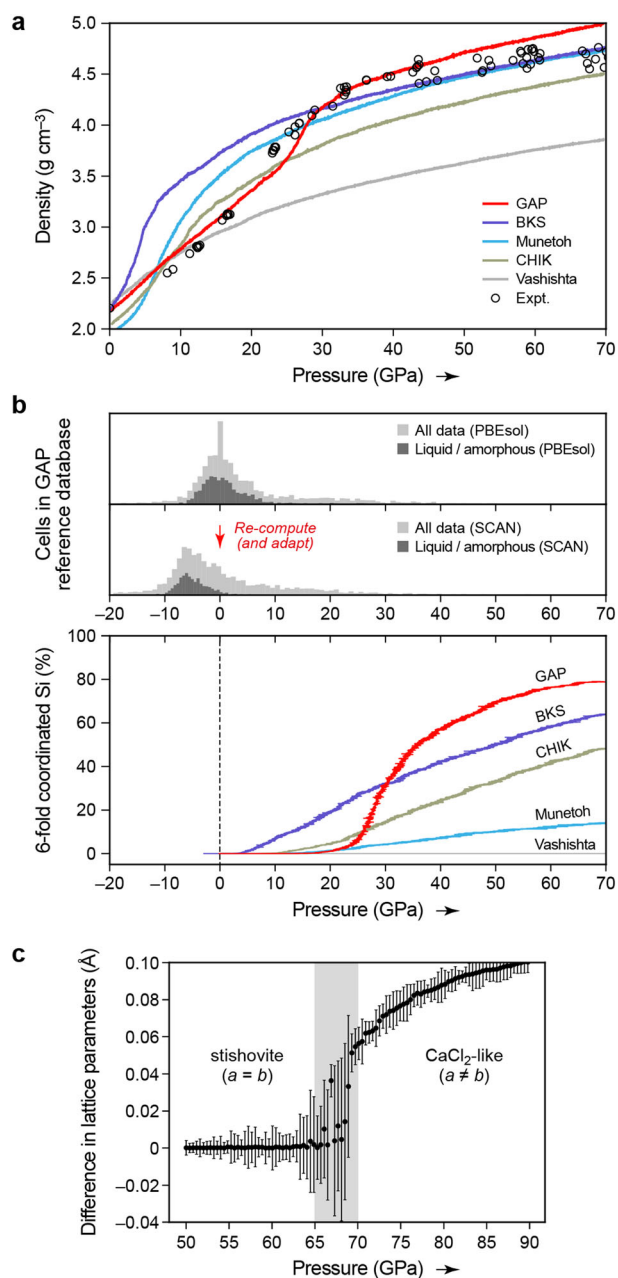


Fig. 6 Extrapolation to high pressure. **a** Density–pressure curve of amorphous silica simulated using different potentials and compared to experimental data by Marsh⁶¹. **b** Proportion of 6-fold coordinated Si atoms as dependent on the pressure. For the Munetoh, BKS, and CHIK potentials, this proportion increases nearly linearly; for the GAP, it increases rapidly at about 30 GPa and then more slowly. The histograms at the top of the plot show the number of unit cells in the GAP training set with given DFT pressures. Although there are some datapoints at high pressure, the amorphous and liquid phase training data in particular are limited to lower pressures—this effect is more pronounced upon re-computing the reference data with the SCAN functional. **c** Mean of the difference between the lattice parameters a and b referenced to a stishovite unitcell; the error is given by the standard deviation. MD simulations were performed at 300 K and at different pressures initially using a stishovite supercell. A transition from stishovite to a CaCl_2 -like structure can be seen at pressures between 65 and 70 GPa.

even when the higher-energy structures are evaluated in subsequent single-point computations.

Since the GAP and CHIK potentials have shown good performance for amorphous silica thus far, we used both to re-relax the other structures as well; these results are shown on the right-hand side of Table 2. Due to the additional relaxation, all structures become significantly more favourable. The mean excess energies for the CHIK-quenched structures even become slightly more favourable than those of their GAP-quenched counterparts (although the results are within one standard deviation of one another). This suggests that a combination of CHIK quenching and GAP relaxation might be promising for the modelling of amorphous silica.

To explore this question, we carried out simulations using a ‘hybrid’ protocol: quenching slowly with CHIK (10^{11} K s^{-1} , two orders of magnitude slower than in the previous tests), and then briefly annealing the quenched structure with the GAP. The first part is made possible by the comparatively low computational cost of the CHIK potential, and indeed leads to a structure factor in nearly quantitative agreement with experiment (blue line in Fig. 5d). Subsequent annealing with the GAP slightly lowers the FSDP again (red line in Fig. 5d), yet the predicted structure factor is much better than that from a fast GAP quench (Fig. 5b). The rationale for including the ML potential, now, is in the accurate description of the energetics: indeed, Fig. 5e shows a substantial gain in excess energy (the glassy structures becoming more favourable) once the GAP optimisation step is included.

Extrapolation to high pressure

We now test the ability of the GAP to extrapolate into unknown regions of structural space, i.e., those for which the potential has not been explicitly optimised. In the present case, we study the high-pressure behaviour; such simulations have recently been shown to be rewarding targets for ML potential-based modelling^{31–33}. High-pressure simulations have also been done for silica using empirical potentials, e.g., using compression to 25 GPa as one of the benchmarks⁶⁰.

In Fig. 6a, we show computed density–pressure diagrams for amorphous silica, obtained by compressing the same (GAP-generated) amorphous structure using various interatomic potential models. For comparison, we include data from Hugoniot shock experiments⁶¹. In these experiments, the samples are rapidly compressed. This allows for much better comparison with the simulation time scales accessible to MD in contrast to static experimental data. We find that the GAP is the only potential which captures the features of the density–pressure curve correctly: up to a density of 4.5 g/cm^3 , its predictions (red line) agree well with experimental data (circles). In contrast, none of the empirical potentials tested in the same way are able to fully reproduce the qualitative features of the experimental pressure–density curve.

More detailed information about the atomic structure can be obtained from Fig. 6b, which shows the proportion of 6-fold coordinated atoms as dependent on the external pressure. While this number grows rapidly in the GAP simulation at about 30 GPa, the CHIK, BKS, and Munetoh potentials predict an almost linear growth. The Vashishta potential appears to be unable to describe 6-fold coordinated atoms correctly and instead predicts a densified, mainly tetrahedral-like structure, with only a small amount of 5-fold coordinated silicon ($\approx 1\%$), and some oxygen atoms being 1- and 3-fold coordinated (each $\approx 5\%$). Experimental work indicated a phase transition from amorphous silica to stishovite at a pressure of 36 GPa⁶². Therefore, in compression

Table 3. Structural and energetic data of α -PbO₂ type silica^a.

		Expt.	SCAN	GAP
Ambient ^b	<i>a</i> (Å)	4.097(1)	4.064	4.078
	<i>b</i> (Å)	5.0462(9)	5.024	4.981
	<i>c</i> (Å)	4.4946(8)	4.488	4.455
	Si–O (Å)	1.742	1.760	1.749
		1.776	1.763	1.755
		1.855	1.816	1.812
	ΔE (meV/SiO ₂)	—	566	534
<i>p</i> = 129 GPa ^c	<i>a</i> (Å)	3.7277(2)	3.717	3.802
	<i>b</i> (Å)	4.6576(2)	4.655	4.596
	<i>c</i> (Å)	4.1609(3)	4.159	4.159
	Si–O (Å)	1.624	1.622	1.626
		1.655	1.649	1.657
		1.677	1.678	1.681

^aThe computed lattice parameters are determined for 0 K. ΔE is the computed energy difference to α -quartz.
^bExperimental data taken from ref. 68; computed data at zero external pressure.
^cExperimental data taken from ref. 69.

simulations, one might expect a rapid transition to a phase with mainly sixfold-coordinated silicon atoms at this pressure (and probably crystallisation on longer time-scales), rather than an almost linear increase in the count of sixfold-coordinated silicon atoms over a wide pressure range.

In a way similar to ref. 63, we include histograms at the top of Fig. 6b that indicate the regions of parameter space that are covered by the reference database. (In turn, the histograms show that the high-pressure regions are *not* well covered—which suggests that the ML potential is indeed extrapolating.) Interestingly, the influence of the DFT exchange–correlation functional on the stresses in the database is clearly visible. Although the databases do not coincide completely, their composition is largely similar. Presumably due to the change of the minimum-energy bond length and bond angles when moving from PBEsol to SCAN, the latter functional causes a shift of the pressures to more negative values. This leads to a distribution which is no longer centred at zero pressure. It therefore exacerbates the challenge for the potential which has not “seen” high-pressure disordered structures.

Returning to the crystalline phases, we finally test the GAP for very-high-pressure (‘post-stishovite’) polymorphs of silica. The sequence of transitions, viz. stishovite (space group $P4_2/mnm$) \rightarrow CaCl₂-type ($Pnmm$) \rightarrow α -PbO₂-type ($Pbcn$) \rightarrow pyrite-type ($Pa\bar{3}$), has been discussed, for example, in ref. 64.

Regarding the first step, we performed GAP-MD simulations at 300 K which indeed showed a tetragonal-to-orthorhombic transition (Fig. 6c), consistent with the change from rutile-type stishovite to the distorted-rutile-like CaCl₂-type. The analysis showed a transition region between 65 GPa and 70 GPa, somewhat higher than in experimental^{65,66} and other theoretical results^{54,64}. The observation of this transition is encouraging, since CaCl₂-type silica had not been explicitly included in the training database. However, the training data *do* include many instances of randomly distorted stishovite unit cells, and in this way, good performance for a specific type of distortion is to be expected.

For the next-higher-pressure polymorph, α -PbO₂-type silica, we list the lattice parameters in Table 3. Both at ambient and high pressure (129 GPa), the starting structure remained stable during GAP relaxation, and the computed lattice parameters agree well

with experimental data^{67–69} and the SCAN reference. GAP phonon calculations confirmed dynamical stability. Moreover, the computed energy of the relaxed structure (relative to α -quartz) agrees with SCAN to be within 6%.

Every potential will at some point reach its limit, and the limitation of the current GAP model becomes apparent when studying the highest-pressure polymorph listed above: the pyrite-type ($Pa\bar{3}$) form which becomes stable at about 200 GPa^{64,70}. In this structure, the silicon atoms occupy the corners and faces of the unit cell, and the oxygen atoms are found on the 8c Wyckoff site, at (x, x, x) and equivalent. Together with the cubic lattice parameter, a , the coordinate x therefore defines the structure⁶⁴. We started with $x = 0.348$ and relaxed the structure with an external pressure of 268 GPa. The GAP relaxation led to $x = 0.25$, corresponding to the fluorite type, which is incorrect for silica. In contrast, a SCAN relaxation at this pressure led to the correct structure with $x = 0.349$. This behaviour can be intuitively understood because the other previously mentioned high-pressure silica polymorphs are based on distorted close packing of anions (including stishovite, which is included in the training), whereas there is no close-packing of anions in the $Pa\bar{3}$ structure⁶⁴. The latter is therefore rather different from the training data and the current GAP model falls short of describing it, leading to what is presumably a false minimum. (If one were to study very-high-pressure phases in detail, one would of course wish to include more information about those phases in extended training databases; such studies are planned.)

DISCUSSION

A machine-learned interatomic potential for silica was developed that accurately describes various crystalline and amorphous phases. The question of what exactly makes an ‘accurate’ ML potential, however, is far from trivial, and may have different answers for different application cases. We assess the quality of our silica potential and others by a range of numerical measures that relate to (i) the crystalline phases, including the phase-diagram prediction; (ii) the amorphous phases; and (iii) the computational speed. Detailed definitions of these quality measures are given in Supplementary Material. We focus the following discussion on silica, but we emphasise that defining accuracy is a much wider-ranging challenge in ML potential development⁴⁴.

Figure 7 shows that the machine-learned GAP model overcomes many of the transferability limitations of established empirical interatomic potentials. In particular, it shows a high accuracy in the prediction of phase stability regions, enabled by the quality of the fit and also by the SCAN DFT functional on which it is based. The GAP performs well for lattice parameters and elastic constants, as do the BKS, CHIK, and Broughton potentials; however, only the GAP achieves a comprehensive description of the stability of the crystalline phases (including the α -quartz–coesite and coesite–stishovite coexistence lines in the phase diagram). We do not include the post-stishovite phases in this assessment, because they are beyond the scope of any of the potentials discussed—nonetheless, they have allowed us to test the extrapolation capability of the GAP, and may be of interest for further work on ML potential development.

For properties of amorphous silica, assessed by inspecting the structure factor and the energetics of fast-quenched structural models (and characterised in the lower part of the radar charts in Fig. 7), the CHIK, Vashishta, and BKS potentials show good predictions—partially better ones than the GAP. However, the GAP yields the highest number of ‘correctly’ (4- and 2-fold, respectively) coordinated atoms. There is also initial evidence of the GAP being able to describe the pressure–density relation of rapidly compressed amorphous silica and thus to extrapolate into unknown regions of configuration space (Fig. 6)—consistent with

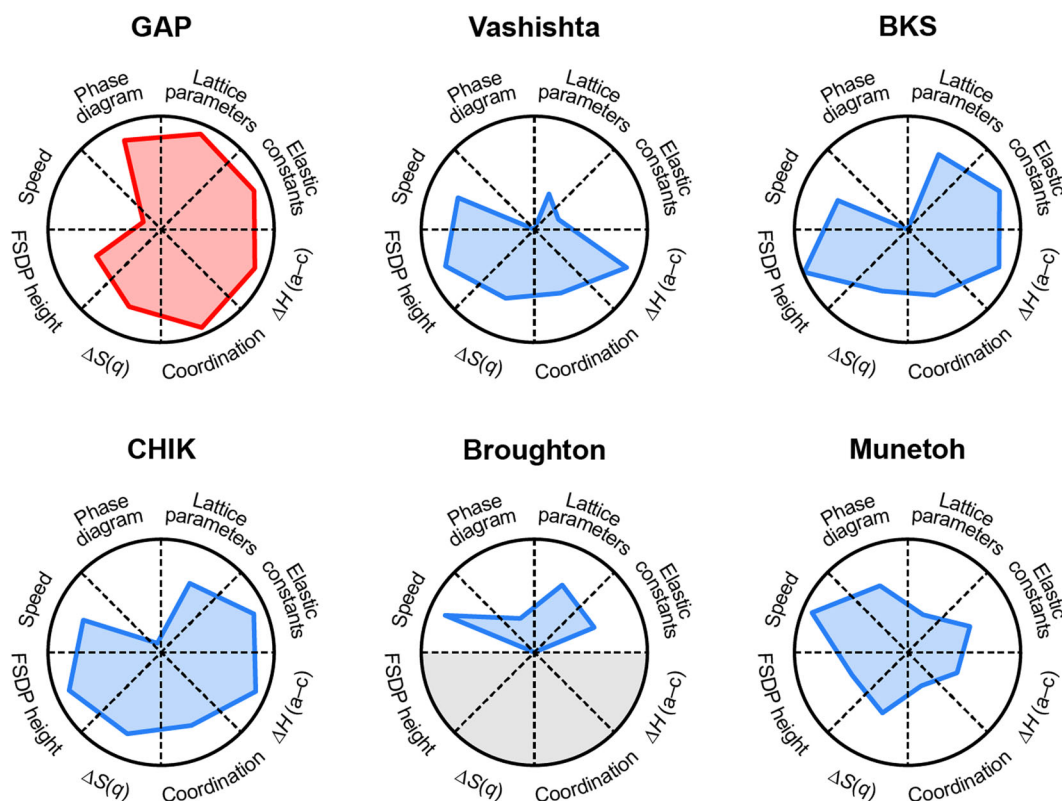


Fig. 7 Performance of various interatomic potentials in predicting properties of silica. The current GAP model describes all considered properties with high accuracy relative to experimental benchmarks; it outperforms all tested potentials in the description of the phase diagrams and lattice parameters, whereas for structure factors of amorphous silica, other potentials give slightly better agreement. For each category, the potentials are given a grade between 0 (worst) and 1 (best). Supplementary Table 3 describes how these values have been obtained.

the encouraging performance for the crystalline α -PbO₂-type polymorph (Table 3).

A remaining challenge is related to the detailed structure of amorphous silica models: specifically, the fact that the structure factor from purely GAP-based quench simulations deviates notably from experiment at small q . The reason why our current machine-learned model underperforms a range of empirical potentials in this regard—viz. CHIK, Vashishta, and BKS—remains unclear at present. A possible origin could be the lack of explicit long-range Coulomb interactions in the current version of our potential; such interactions are included, for example, in the CHIK potential which indeed performs very well for the amorphous phase. This assumption seems consistent with earlier arguments for the origin of the FSDP⁷¹, e.g., caused by intermediate-range ordering between [SiO₄] tetrahedra^{72,73}. However, it is not a priori clear that inclusion of long-range interactions is strictly required: for titania (neural-network ML potential; ref. 37) and hafnia (GAP; ref. 74), an explicit description of such long-range interactions was not necessary to reproduce the first peak in the structure factor. The role of long-range electrostatics for GAP models, assessed through inclusion of a baseline model, was recently discussed in ref. 75. An alternative hypothesis, namely, that the GAP-MD cooling rates are not sufficient, would be supported if the potential behaved better for the liquid phase. However, despite reasonable performance for the self-diffusion coefficients (Supplementary Fig. 4), the GAP also underestimates the height of the first peak for liquid silica (Supplementary Fig. 3); a comparable underestimation was seen with a different ML potential fitting framework⁴².

Even though the machine-learned potential described herein shows generally wide transferability, we emphasise that it is orders of magnitude slower than empirically fitted ones. The

message arising from the present work, we believe, is therefore more nuanced: rather than replacing existing silica potentials with machine-learned alternatives entirely, we envision this model as a *complement* to suitable empirically fitted potentials. To illustrate this point, we generated a large structural model of the ambient-pressure glass by combining slow quenching with the CHIK potential and subsequent GAP optimisation. The resulting structure is lower in energy than its purely CHIK-quenched counterpart, while retaining much of the excellent description of the structure factor that is enabled by CHIK. Such ‘hybrid’ protocols may be useful for future computational studies of silica.

In closing, we emphasise that the current version of the machine-learned potential is only trained for single-phase models of bulk SiO₂. For systems with surfaces, grain boundaries, or interfaces, further extensions to the reference database are expected to be required.

METHODS

Molecular dynamics

Molecular dynamics simulations were performed using the LAMMPS software⁷⁶, a Nosé–Hoover thermostat^{77,78} for NVT and additionally a Parrinello–Rahman barostat⁷⁹ for NPT simulations. We used a time step of 1 fs, a temperature damping constant of 100 fs, and a pressure damping constant of 1000 fs.

Amorphous-phase structural models were created from initially randomly placed atoms, which were additionally randomised at 6000 K (NVT, 10 ps) and then held at 4000 K (NPT with zero external pressure, 100 ps) to generate a melt. The melt was quenched to 300 K with a rate of 10¹³ K/s, unless noted otherwise. The resulting amorphous structure was then held for another 10 ps. ‘Hybrid’ quenches with a combination of CHIK and GAP

(Fig. 5d, e) used the same procedure for the CHIK potential, however, with quenching rates between 10^{13} K/s and 10^{11} K/s, and subsequent annealing for 20 ps using the GAP.

Density-functional theory computations

For DFT computations, the GPAW^{80,81} (in combination with ASE⁸²) and VASP^{83–86} codes employing the projector augmented-wave method^{87,88} were used. Early versions of the GAP were based on reference data computed using the PBEsol functional⁸⁹; later on, we used the SCAN functional⁴³, as detailed in the supplementary material. For GPAW, we used an energy cut-off of 700 eV and a k -spacing of 0.279 \AA^{-1} , for VASP, a higher energy cut-off of 900 eV and a denser k -spacing of 0.23 \AA^{-1} .

GAP fitting

The GAP model was fitted using a combination of 2-body descriptors and the Smooth Overlap of Atomic Positions (SOAP)⁹⁰, with a radial cut-off of 5 Å for each descriptor. The number of representative ('sparse') points was 15 (Si–Si 2-body), 15 (Si–O 2-body), 15 (O–O 2-body), and 4,060 (SOAP), respectively. Details concerning the construction of the reference database and the choice of fitting parameters are given in the supplementary material. A comprehensive review of the GAP methodology is found in ref. ⁴⁴.

Validation of potential models

Bulk moduli of crystalline phases were obtained from energy–volume data for hydrostatically deformed structures by fitting to the Birch–Murnaghan equation of state. Elements of the stiffness tensor were obtained by fitting the energy variation upon applying finite distortions. 21 strain modes were applied, using strain components (ϵ_{11} ϵ_{22} ϵ_{33} $2\epsilon_{23}$ $2\epsilon_{13}$ $2\epsilon_{12}$) with $\epsilon_{ij} \in [-1\%, 1\%]$. The curvature of the energy–strain curves was determined by a fourth-order fit. From this, we obtain 21 equations for the elastic constants. The independent elastic constants were determined by a fit of the overdetermined system of equations.

Phonon dispersions and free energies were computed within the frozen-phonon approximation as implemented in `phonopy`⁹¹. For α -quartz, coesite, and stishovite, the quasi-harmonic approximation (QHA) was used to determine the phase boundaries in a p – T diagram⁹².

Structure factors were obtained as weighted sums of Faber–Ziman partial structure factors determined over the last 5 ps (10 ps for hybrid simulations) of simulations containing 1,728 formula units. Coordination numbers were determined using OVITO with a bond-length cut-off of 2 Å⁹³. Single-point energies were computed using VASP for cells with 192 atoms (after relaxation of the ionic positions and the cell size with the respective potential).

For studying compression behaviour, the GAP-quenched amorphous structure containing 1,728 formula units was used as starting point. These structures were held for 20 ps using the different potentials and then compressed up to 70 GPa in 20 ps. Again, the coordination was determined using OVITO and a bond-length cut-off of 2 Å.

DATA AVAILABILITY

The potential parameter files and associated DFT reference data, as well as relevant structural data, are openly available in the Zenodo repository at <https://doi.org/10.5281/zenodo.6353684>. The identifier for the potential (given in the XML file) is GAP_2021_4_19_120_7_32_55_336.

Received: 7 October 2021; Accepted: 26 March 2022;

Published online: 28 April 2022

REFERENCES

- Heaney, P. J., Prewitt, C. T. & Gibbs, G. V. (eds.) *Silica: Physical Behavior, Geochemistry, and Materials Applications* (Berlin, Boston, 1994).
- Danel, J. S. & Delapierre, G. Quartz: a material for microdevices. *J. Microeng. Microeng.* **1**, 187–198 (1991).
- Parmenter, K. E. & Milstein, F. Mechanical properties of silica aerogels. *J. Non-Cryst. Solids* **223**, 179–189 (1998).
- Weitkamp, J. Zeolites and catalysis. *Solid State Ion.* **131**, 175–188 (2000).
- Cady, W. G. Piezoelectricity: an introduction to the theory and applications of electromechanical phenomena in crystals. International series in pure and applied physics (New York, London, 1946), 1st ed.
- Yeganeh-Haeri, A., Weidner, D. J. & Parise, J. B. Elasticity of α -cristobalite: a silicon dioxide with a negative Poisson's ratio. *Science* **257**, 650–652 (1992).
- Kingma, K. J., Cohen, R. E., Hemley, R. J. & Mao, H.-k. Transformation of stishovite to a denser phase at lower-mantle pressures. *Nature* **374**, 243–245 (1995).
- Yilmaz, B. & Müller, U. Catalytic applications of zeolites in chemical industry. *Top. Catal.* **52**, 888–895 (2009).
- Blanc, W. & Dussardier, B. Formation and applications of nanoparticles in silica optical fibers. *J. Opt.* **45**, 247–254 (2016).
- van Beest, B. W. H., Kramer, G. J. & van Santen, R. A. Force fields for silicas and aluminophosphates based on ab initio calculations. *Phys. Rev. Lett.* **64**, 1955–1958 (1990).
- Vollmayr, K., Kob, W. & Binder, K. Cooling-rate effects in amorphous silica: a computer-simulation study. *Phys. Rev. B* **54**, 15808–15827 (1996).
- Koziatek, P., Barrat, J. L. & Rodney, D. Short- and medium-range orders in as-quenched and deformed SiO₂ glasses: an atomistic study. *J. Non-Cryst. Solids* **414**, 7–15 (2015).
- Cowen, B. J. & El-Genk, M. S. On force fields for molecular dynamics simulations of crystalline silica. *Comput. Mater. Sci.* **107**, 88–101 (2015).
- Saika-Voivod, I., Sciortino, F., Grande, T. & Poole, P. H. Phase diagram of silica from computer simulation. *Phys. Rev. E* **70**, 061507 (2004).
- Flikkema, E. & Bromley, S. A new interatomic potential for nanoscale silica. *Chem. Phys. Lett.* **378**, 622–629 (2003).
- Carré, A., Horbach, J., Ispas, S. & Kob, W. New fitting scheme to obtain effective potential from Car-Parrinello molecular-dynamics simulations: application to silica. *EPL* **82**, 17001 (2008).
- Tangney, P. & Scandolo, S. An ab initio parametrized interatomic force field for silica. *J. Chem. Phys.* **117**, 8898–8904 (2002).
- Yasukawa, A. Using an extended Tersoff interatomic potential to analyze the static-fatigue strength of SiO₂ under atmospheric influence. *J. Ser. A Mech. Mater. Eng.* **39**, 313–320 (1996).
- Yu, J., Sinnott, S. B. & Phillpot, S. R. Charge optimized many-body potential for the Si-SiO₂ system. *Phys. Rev. B* **75**, 085311 (2007).
- Shan, T.-R. et al. Second-generation charge-optimized many-body potential for Si-SiO₂ and amorphous silica. *Phys. Rev. B* **82**, 235302 (2010).
- Vashishta, P., Kalia, R. K., Rino, J. P. & Ebbsjö, I. Interaction potential for SiO₂: a molecular-dynamics study of structural correlations. *Phys. Rev. B* **41**, 12197–12209 (1990).
- Nakano, A., Kalia, R. K. & Vashishta, P. First sharp diffraction peak and intermediate-range order in amorphous silica: finite-size effects in molecular dynamics simulations. *J. Non-Cryst. Solids* **171**, 157–163 (1994).
- Broughton, J. Q., Meli, C. A., Vashishta, P. & Kalia, R. K. Direct atomistic simulation of quartz crystal oscillators: bulk properties and nanoscale devices. *Phys. Rev. B* **56**, 611–618 (1997).
- Munetoh, S., Motooka, T., Moriguchi, K. & Shintani, A. Interatomic potential for Si–O systems using Tersoff parameterization. *Comput. Mater. Sci.* **39**, 334–339 (2007).
- Herzbach, D., Binder, K. & Müser, M. H. Comparison of model potentials for molecular-dynamics simulations of silica. *J. Chem. Phys.* **123**, 124711 (2005).
- Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
- Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
- Thompson, A., Swiler, L., Trott, C., Foiles, S. & Tucker, G. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J. Comput. Phys.* **285**, 316–330 (2015).
- Shapeev, A. V. Moment tensor potentials: a class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **14**, 1153–1173 (2016).
- Zhang, L., Han, J., Wang, H., Car, R. & E, W. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).
- Cheng, B., Mazzola, G., Pickard, C. J. & Ceriotti, M. Evidence for supercritical behaviour of high-pressure liquid hydrogen. *Nature* **585**, 217–220 (2020).
- Deringer, V. L. et al. Origins of structural and electronic transitions in disordered silicon. *Nature* **589**, 59–64 (2021).
- Zong, H. et al. Free electron to electride transition in dense liquid potassium. *Nat. Phys.* **17**, 955–960 (2021).
- Wang, C., Aoyagi, K., Wisesa, P. & Mueller, T. Lithium ion conduction in cathode coating materials from on-the-fly machine learning. *Chem. Mater.* **32**, 3741–3752 (2020).
- Wengert, S., Csányi, G., Reuter, K. & Margraf, J. T. Data-efficient machine learning for molecular crystal structure prediction. *Chem. Sci.* **12**, 4536–4546 (2021).

36. Artrith, N., Morawietz, T. & Behler, J. High-dimensional neural-network potentials for multicomponent systems: applications to zinc oxide. *Phys. Rev. B* **83**, 153101 (2011).
37. Calegari Andrade, M. F. & Selloni, A. Structure of disordered TiO₂ phases from ab initio based deep neural network simulations. *Phys. Rev. Mater.* **4**, 113803 (2020).
38. Sivaraman, G. et al. Machine-learned interatomic potentials by active learning: amorphous and liquid hafnium dioxide. *npj Comput. Mater.* **6**, 104 (2020).
39. Timmermann, J. et al. IrO₂ surface complexions identified through machine learning and surface investigations. *Phys. Rev. Lett.* **125**, 206101 (2020).
40. Novikov, I. S. & Shapcev, A. V. Improving accuracy of interatomic potentials: more physics or more data? A case study of silica. *Mater. Today Commun.* **18**, 74–80 (2019).
41. Balyakin, I. A., Rempel, S. V., Ryltsev, R. E. & Rempel, A. A. Deep machine learning interatomic potential for liquid silica. *Phys. Rev. E* **102**, 052125 (2020).
42. Kobayashi, K., Nagai, Y., Itakura, M. & Shiga, M. Self-learning hybrid Monte Carlo method for isothermal-isobaric ensemble: application to liquid silica. *J. Chem. Phys.* **155**, 034106 (2021).
43. Sun, J., Ruzsinszky, A. & Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **115**, 036402 (2015).
44. Deringer, V. L. et al. Gaussian process regression for materials and molecules. *Chem. Rev.* **121**, 10073–10141 (2021).
45. Rowe, P., Deringer, V. L., Gasparotto, P., Csányi, G. & Michaelides, A. An accurate and transferable machine learning potential for carbon. *J. Chem. Phys.* **153**, 034702 (2020).
46. Deringer, V. L., Caro, M. A. & Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.* **31**, 1902765 (2019).
47. Deringer, V. L. & Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Phys. Rev. B* **95**, 094203 (2017).
48. Sosso, G. C., Miceli, G., Caravati, S., Behler, J. & Bernasconi, M. Neural network interatomic potential for the phase change material GeTe. *Phys. Rev. B* **85**, 174103 (2012).
49. Shen, Y., Jester, S. B., Qi, T. & Reed, E. J. Nanosecond homogeneous nucleation and crystal growth in shock-compressed SiO₂. *Nat. Mater.* **15**, 60–65 (2016).
50. George, J., Hautier, G., Bartók, A. P., Csányi, G. & Deringer, V. L. Combining phonon accuracy with high transferability in Gaussian approximation potential models. *J. Chem. Phys.* **153**, 044104 (2020).
51. Dorner, B., Grimm, H. & Rzany, H. Phonon dispersion branches in α -quartz. *J. Phys. C: Solid State Phys.* **13**, 6607–6613 (1980).
52. Wehinger, B. et al. Lattice dynamics of α -cristobalite and the boson peak in silica glass. *J. Phys.: Condens. Matter* **27**, 305401 (2015).
53. Bosak, A. et al. Lattice dynamics of stishovite from powder inelastic x-ray scattering. *Geophys. Res. Lett.* **36**, L19309 (2009).
54. Tsuchiya, T., Caracas, R. & Tsuchiya, J. First principles determination of the phase boundaries of high-pressure polymorphs of silica. *Geophys. Res. Lett.* **31**, L11610 (2004).
55. Bohlen, S. R. & Boettcher, A. L. The quartz–coesite transformation: a precise determination and the effects of other components. *J. Geophys. Res. Solid Earth* **87**, 7073–7078 (1982).
56. Akaogi, M., Yusa, H., Shiraishi, K. & Suzuki, T. Thermodynamic properties of α -quartz, coesite, and stishovite and equilibrium phase relations at high pressures and high temperatures. *J. Geophys. Res. Solid Earth* **100**, 22337–22347 (1995).
57. Lane, J. M. D. Cooling rate and stress relaxation in silica melts and glasses via microsecond molecular dynamics. *Phys. Rev. E* **92**, 012320 (2015).
58. Richet, P., Bottinga, Y., Denielou, L., Petitet, J. P. & Tequi, C. Thermodynamic properties of quartz, cristobalite and amorphous SiO₂: drop calorimetry measurements between 1000 and 1800 K and a review from 0 to 2000 K. *Geochim. Cosmochim. Acta* **46**, 2639–2658 (1982).
59. Huffman, M., Navrotsky, A. & Pintchovski, F. S. Thermochemical and spectroscopic studies of chemically vapor-deposited amorphous silica. *J. Electrochem. Soc.* **133**, 164 (1986).
60. Sundararaman, S., Huang, L., Ispas, S. & Kob, W. New optimization scheme to obtain interaction potentials for oxide glasses. *J. Chem. Phys.* **148**, 194504 (2018).
61. Marsh, S. P. LASS Shock Hugoniot Data. University of California Press, Los Angeles (1980).
62. Tracy, S. J., Turneure, S. J. & Duffy, T. S. In situ X-ray diffraction of shock-compressed fused silica. *Phys. Rev. Lett.* **120**, 135702 (2018).
63. Bartók, A. P. & Csányi, G. Gaussian approximation potentials: a brief tutorial introduction. *Int. J. Quantum Chem.* **115**, 1051–1057 (2015).
64. Oganov, A. R., Gillan, M. J. & Price, G. D. Structural stability of silica at high pressures and temperatures. *Phys. Rev. B* **71**, 064104 (2005).
65. Ono, S., Hirose, K., Murakami, M. & Isshiki, M. Post-stishovite phase boundary in SiO₂ determined by in situ x-ray observations. *Earth Planet. Sci. Lett.* **197**, 187–192 (2002).
66. Fischer, R. A. et al. Equations of state and phase boundary for stishovite and CaCl₂-type SiO₂. *Am. Mineral.* **103**, 792–802 (2018).
67. Dera, P., Prewitt, C. T., Boctor, N. Z. & Hemley, R. J. Characterization of a high-pressure phase of silica from the Martian meteorite Shergotty. *Am. Mineral.* **87**, 1018–1023 (2002).
68. El Goresy, A. et al. Seifertite, a dense orthorhombic polymorph of silica from the Martian meteorites Shergotty and Zagami. *Eur. J. Mineral.* **20**, 523–528 (2008).
69. Zhang, L. et al. In-situ crystal structure determination of seifertite SiO₂ at 129 GPa: Studying a minor phase near Earth's core–mantle boundary. *Am. Mineral.* **101**, 231–234 (2016).
70. Taek Park, K., Terakura, K. & Matsui, Y. Theoretical evidence for a new ultra-high-pressure phase of SiO₂. *Nature* **336**, 670–672 (1988).
71. Elliott, S. R. Medium-range structural order in covalent amorphous solids. *Nature* **354**, 445–452 (1991).
72. Mei, Q., Benmore, C. J., Sen, S., Sharma, R. & Yarger, J. L. Intermediate range order in vitreous silica from a partial structure factor analysis. *Phys. Rev. B* **78**, 144204 (2008).
73. Wright, A. C. Crystalline-like ordering in melt-quenched network glasses? *J. Non-Cryst. Solids* **401**, 4–26 (2014).
74. Sivaraman, G. et al. Experimentally driven automated machine-learned interatomic potential for a refractory oxide. *Phys. Rev. Lett.* **126**, 156002 (2021).
75. Staacke, C. G. et al. On the role of long-range electrostatics in machine-learned interatomic potentials for complex battery materials. *ACS Appl. Energy Mater.* **4**, 12562–12569 (2021).
76. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
77. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).
78. Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
79. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
80. Enkovaara, J. et al. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *J. Phys.: Condens. Matter* **22**, 253202 (2010).
81. Mortensen, J. J., Hansen, L. B. & Jacobsen, K. W. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B* **71**, 035109 (2005).
82. Larsen, A. H. et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys.: Condens. Matter* **29**, 273002 (2017).
83. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
84. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
85. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
86. Kresse, G. & Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251–14269 (1994).
87. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
88. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
89. Perdew, J. P. et al. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).
90. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
91. Togo, A. & Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1–5 (2015).
92. Togo, A., Chaput, L., Tanaka, I. & Hug, G. First-principles phonon calculations of thermal expansion in Ti₃SiC₂, Ti₃AlC₂, and Ti₃GeC₂. *Phys. Rev. B* **81**, 174301 (2010).
93. Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool. *Model. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).
94. Heyliger, P., Ledbetter, H. & Kim, S. Elastic constants of natural quartz. *J. Acoust. Soc. Am.* **114**, 644–650 (2003).
95. Levien, L., Prewitt, C. T. & Weidner, D. J. Structure and elastic properties of quartz at pressure. *Am. Mineral.* **65**, 920–930 (1980).
96. Liu, L.-g. Bulk moduli of SiO₂ polymorphs: Quartz, coesite and stishovite. *Mech. Mater.* **14**, 283–290 (1993).
97. Kirfel, A. & Will, G. Ending the “P₂/a coesite” discussion. *Z. Kristallogr.* **167**, 287–292 (1984).
98. Keskar, N. R. & Chelikowsky, J. R. Structural properties of nine silica polymorphs. *Phys. Rev. B* **46**, 1–13 (1992).
99. Diaz-Cabañas, M.-J. & Barrett, P. A. Synthesis and structure of pure SiO₂ chabazite: the SiO₂ polymorph with the lowest framework density. *Chem. Commun.* **13**, 1881–1882 (1998).

100. Downs, R. T. & Palmer, D. C. The pressure behavior of a cristobalite. *Am. Mineral.* **79**, 9–14 (1994).
101. Léger, J.-M., Haines, J. & Chateau, C. The high-pressure behaviour of the “moganite” polymorph of SiO₂. *Eur. J. Mineral.* **13**, 351–359 (2001).
102. Miehe, G. & Graetsch, H. Crystal structure of moganite: a new structure type for silica. *Eur. J. Mineral.* **4**, 693–706 (1992).
103. Kato, K. & Nukui, A. Die Kristallstruktur des monoklinen Tief-Tridymits. *Acta Crystallogr. Sect. B* **32**, 2486–2491 (1976).
104. Bernstein, N., Csányi, G. & Deringer, V. L. De novo exploration and self-guided learning of potential-energy surfaces. *npj Comput. Mater.* **5**, 99 (2019).
105. Bechmann, R. Elastic and piezoelectric constants of alpha-quartz. *Phys. Rev.* **110**, 1060–1061 (1958).
106. Weidner, D. J. & Carleton, H. R. Elasticity of coesite. *J. Geophys. Res.* **82**, 1334–1346 (1977).
107. Weidner, D. J., Bass, J. D., Ringwood, A. E. & Sinclair, W. The single-crystal elastic moduli of stishovite. *J. Geophys. Res. Solid Earth* **87**, 4740–4746 (1982).
108. Mei, Q., Benmore, C. J. & Weber, J. K. R. Structure of liquid SiO₂: a measurement by high-energy X-ray diffraction. *Phys. Rev. Lett.* **98**, 057802 (2007).

ACKNOWLEDGEMENTS

The research was supported by the Bundesministerium für Bildung und Forschung (BMBF) within the project FESTBATT under Grant No. 03XP0174A. Calculations for this research were conducted on the Lichtenberg high performance computer of the TU Darmstadt. L.C.E. acknowledges support from the Erasmus+ programme for a research visit to Oxford University. J.R. and L.C.E. acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, Grant no. RO 4542/4-1 and STU 611/5-1). V.L.D. acknowledges a Leverhulme Early Career Fellowship.

AUTHOR CONTRIBUTIONS

L.C.E. performed all calculations, the construction of the database and the fitting and validation of the potential. J.R. and K.A. supervised the density-functional theory calculations and the study of elastic and thermodynamic properties. V.L.D. supervised the construction of the database, the fitting of the potential, and the study of the amorphous structures. All authors contributed to the design of the research and to data analysis and interpretation. All authors wrote the paper.

FUNDING

Open Access funding enabled and organized by Projekt DEAL.

COMPETING INTERESTS

The authors declare that there are no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-022-00768-w>.

Correspondence and requests for materials should be addressed to Jochen Rohrer, Karsten Albe or Volker L. Deringer.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022