Machine-learning potentials for nanoscale simulations of tensile deformation and fracture in ceramics

Check for updates

Shuyao Lin [1,2] ✉, Luis Casillas-Trujillo[2], Ferenc Tasnádi [2], Lars Hultman [2], Paul H. Mayrhofer [1], Davide G. Sangiovanni [2] & Nikola Koutná [1,2]

Machine-learning interatomic potentials (MLIPs) offer a powerful avenue for simulations beyond length and timescales of ab initio methods. Their development for investigation of mechanical properties and fracture, however, is far from trivial since extended defects—governing plasticity and crack nucleation in most materials—are too large to be included in the training set. Using TiB_2 as a model ceramic material, we propose a training strategy for MLIPs suitable to simulate mechanical response of monocrystals until failure. Our MLIP accurately reproduces ab initio stresses and fracture mechanisms during room-temperature uniaxial tensile deformation of TiB_2 at the atomic scale ($\approx 10^3$ atoms). More realistic tensile tests (low strain rate, Poisson's contraction) at the nanoscale ($\approx 10^4$–10^6 atoms) require MLIP up-fitting, i.e., learning from additional ab initio configurations. Consequently, we elucidate trends in theoretical strength, toughness, and crack initiation patterns under different loading directions. As our MLIP is specifically trained to modelling tensile deformation, we discuss its limitations for description of different loading conditions and lattice structures with various Ti/B stoichiometries. Finally, we show that our MLIP training procedure is applicable to diverse ceramic systems. This is demonstrated by developing MLIPs which are subsequently validated by simulations of uniaxial strain and fracture in TaB_2, WB_2, ReB_2, TiN, and Ti_2AlB_2.

Simulations of materials' mechanical response—including (i) intrinsic strength and toughness, (ii) nucleation of extended defects (e.g., dislocations, stacking faults, cracks) and their implications for (iii) plasticity and fracture mechanisms —require length and time scales beyond limits of ab initio methods ($\approx 10^3$ atoms, $\ll$ ns)[1–4]. The *go-to* approach in most cases would be classical Molecular Dynamics (MD), allowing to access atomistic pathways for deformation and fracture in nanoscale systems ($\approx 10^6$ atoms) and "realistic" operation conditions (e.g., ultra-high temperatures, times up to μs). However, a severe problem of MD is that the necessary interatomic potentials do not exist for most engineering materials or are limited in accuracy and transferability with respect to temperatures, phases, and defective structures (see e.g., refs. 5–7).

A powerful avenue for MD simulations on multiple time and length scales with near ab initio accuracy is the application of machine learning interatomic potentials[8,9] (MLIPs), in case of no ambiguity just "potentials"). MLIPs learn the atomic energy (or other atomic properties like forces) from a descriptor that characterizes local atomic environments in an ab initio training set[10]. Compared to conventional ab initio calculations, MD with MLIPs can achieve a speed up of as much as 5 orders of magnitude[11,12]. Previous studies showed examples of MLIPs' transferability with respect to defects (e.g., grain boundaries[13], dislocation structures[14,15]) and phases[16,17] (e.g., Ni-Mo phase diagram illustrating superior performance of a MLIP over a classical potential[18]). Recently, Tasnádi et al.[19] have demonstrated high accuracy of MLIP-predicted elastic constants for TiAlN ceramics, hence, have set the stage for MLIP development beyond linear elastic regime.

Based on the parametrization of local structural properties, MLIPs can be fitted employing different formalisms: spectral neighbor analysis potentials (SNAP)[20], neural networks potentials (NNP)[21], Gaussian approximation potentials (GAP)[22], moment tensor potentials (MTP)[23], linearized interatomic potentials[24], or atom cluster expansion (ACE) potentials[10]. Benchmarks for some of these parametrizations have been published in the case of carbon[25] or graphene[26], but are missing for chemically complex materials.

[1]Technische Universität Wien, Institute of Materials Science and Technology, Vienna, Austria. [2]Linköping University, Department of Physics, Chemistry, and Biology (IFM), Linköping, Sweden. ✉e-mail: shuyao.lin@tuwien.ac.at

Fundamental challenges related to MLIP developments include (i) efficient training dataset generation, (ii) training strategies for simulations beyond length scales feasible for ab initio methods, and (iii) assessing the MLIPs' reliability over different length scales. Point (ii) closely relates to successively improving the MLIP's predictive power by up-fitting/active learning[27]. Here an important concept is the extrapolation/Maxvol grade (γ[28]) expressing the extent of MLIP's extrapolation on any structure containing given chemical elements (irrespective of the phase and stoichiometry). Readily available in current ACE and MTP implementations[29–31], γ allows selecting sufficiently "new" environments to expand the training set. Besides γ, other methods used for selecting configurations are stratified sampling or random selection (for detailed discussion of *pros* and *cons* see refs. [32–35]). Besides identifying new environments, γ can serve to indicate the MLIP's reliability during MD simulations. This is particularly advantageous at scales where the direct validation by ab initio calculations is not possible.

Our work pursues the development of MLIPs suitable to simulate atomic-to-nanoscale deformation of ceramics, providing both methodological and materials science discussion. The model material, TiB_2, is a widely studied system, representative of ultra-high temperature ceramics (UHTCs). Exhibiting high hardness and resistance to corrosion, abrasive and erosive wear[36,37], UHTCs are suitable to protect tools and machining components under extreme operation conditions[38–41]. TiB_2, which crystallizes in the AlB_2-type phase[42–44] (α, P6/mmm), exemplifies outstanding mechanical properties[45,46] of UTHCs. It exhibits high hardness of 41–53 GPa[45,47,48], has a melting point of 3500 K[49] and mature synthesis technologies[50,51]. Insights into mechanical behaviour of TiB_2 and other diborides have been offered by ab initio calculations[52–54] and recently also by molecular dynamics with classical empirical potentials (TiB_2[55,56], ZrB_2[57], HfB_2[58]). To date, however, no MLIP capable of predicting mechanical response of UTHCs until fracture has been reported.

Using the MTP formalism, we propose a general training strategy for development of MLIPs targeted to model tensile deformation and fracture of single-crystal ceramics at finite temperatures. The extrapolation grade, γ, is exploited to iteratively improve our MLIP and also as a mean of validation. Specifically, γ values calculated during MD tensile tests are discussed alongside with statistical errors of relevant physical observables, such as time-averaged stresses derived from equivalent ab initio molecular dynamics (AIMD) simulations. The newly-developed MLIP allows to describe tensile deformation in TiB_2 supercells with $\approx 10^3$–10^6 atoms, thus

also providing a basis for analyzing size effects on mechanical properties and fracture patterns. Furthermore, we test transferability of our MLIP to other loading conditions and phases, as well as applicability of our general training strategy to other ceramics, exemplified by TaB_2, WB_2, ReB_2, TiN, and Ti_2AlB_2.

Results and discussion

We aim to develop MLIPs targeted to simulations of tensile deformation in TiB_2. Although the mechanical properties of ceramics are typically assessed by nanoindenation, microcantilever bending, or micropillar compression experiments, MD simulations of tensile loading can be directly compared to results of ab initio calculations. We simulate deformation along low-index [0001], [10$\bar{1}$0], and [$\bar{1}2\bar{1}0$] directions that are parallel or normal to the typical growth direction of hexagonal transition metal diborides[36] and have been considered in previous ab initio studies of surface energies and mechanical properties[54].

Training procedure and fitting initial MLIPs

Our general training procedure is described below (Procedure 1) and schematically depicted in Fig. 1a (for further details, see the Methods). Throughout this work, training configurations (i.e., structures labelled by ab initio total energies, forces acting on each ion, and six stress tensor components) are generated by finite-temperature AIMD calculations. To avoid over-representation, a training set is initialized by a small fraction of randomly selected AIMD snapshots and iteratively expanded using the concept of the extrapolation grade[28], γ.

The extrapolation grade $\gamma \in \mathbf{R}^+$ (see equations 9–1 in ref. [28]) expresses the degree of extrapolation when MLIP predicts atomic properties of a certain configuration. Specifically, $\gamma \leq 1$ means interpolation and $\gamma > 1$ extrapolation. Hence, the higher the γ, the larger is the uncertainty on predicted energies, forces, and stresses[28]. In practice, one sets an extrapolation threshold (γ_{thr}). Configurations with $\gamma > \gamma_{thr}$ are added to the training set to improve transferability of the potential. Procedure 1 uses $\gamma_{thr} = 2$, motivated by our tests (In particular, during each iteration (i) in the step (4) of Procedure 1, the maximum γ is correlated with errors of energies, forces and stresses (quantified via common regression model evaluation metrics, MAE, RMSE, R^2, see e.g., ref. [59]) for the TS_i (fitting errors) and the VS (validation errors). With $\gamma_{thr} = 2$, we are limited mainly by the accuracy of the underlying ab initio training data.) as well as by previous work by the
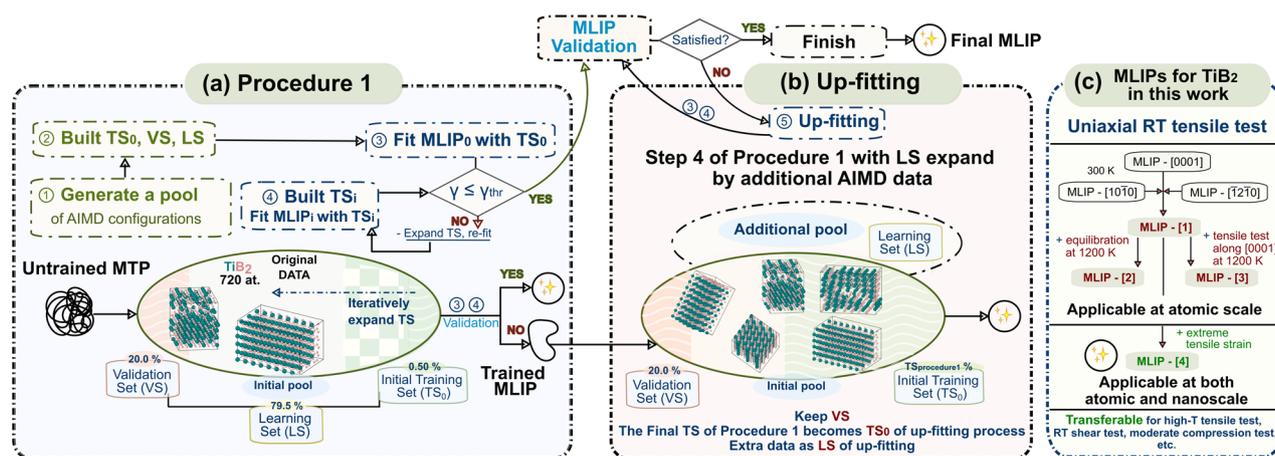
Fig. 1 | Our MLIP training strategy. Schematic visualization of **a** our general training procedure (Procedure 1, Section Training procedure and fitting initial MLIPs), and **b** up-fitting (Section MLIPs' validation against atomic scale tensile tests and MLIPs' up-fitting for nanoscale tensile tests), both applicable to any MLIP formalism with the extrapolation/Maxvol grade (γ) quantification, particularly MTP (used in this work) and ACE. **c** An overview of the here-developed MLIPs: MLIP-[0001], MLIP-[10$\bar{1}$0], MLIP-[$\bar{1}2\bar{1}0$] (applicable for atomic scale tensile loading, Sections Training procedure and fitting initial MLIPs and MLIPs' validation against atomic scale tensile tests); MLIP-[1], MLIP-[2], MLIP-[3] (intermediate MLIPs produced by up-fitting, Section MLIPs' up-fitting for nanoscale tensile tests); and MLIP-[4] (applicable for both atomic and nanoscale tensile loading, Section MLIPs' up-fitting for nanoscale tensile tests--Otherloading conditions and MLIP's transferability). Details of the training datasets are given in the corresponding sections and additionally summarized in the Methods under "Development of machine-learning interatomic potentials". both.
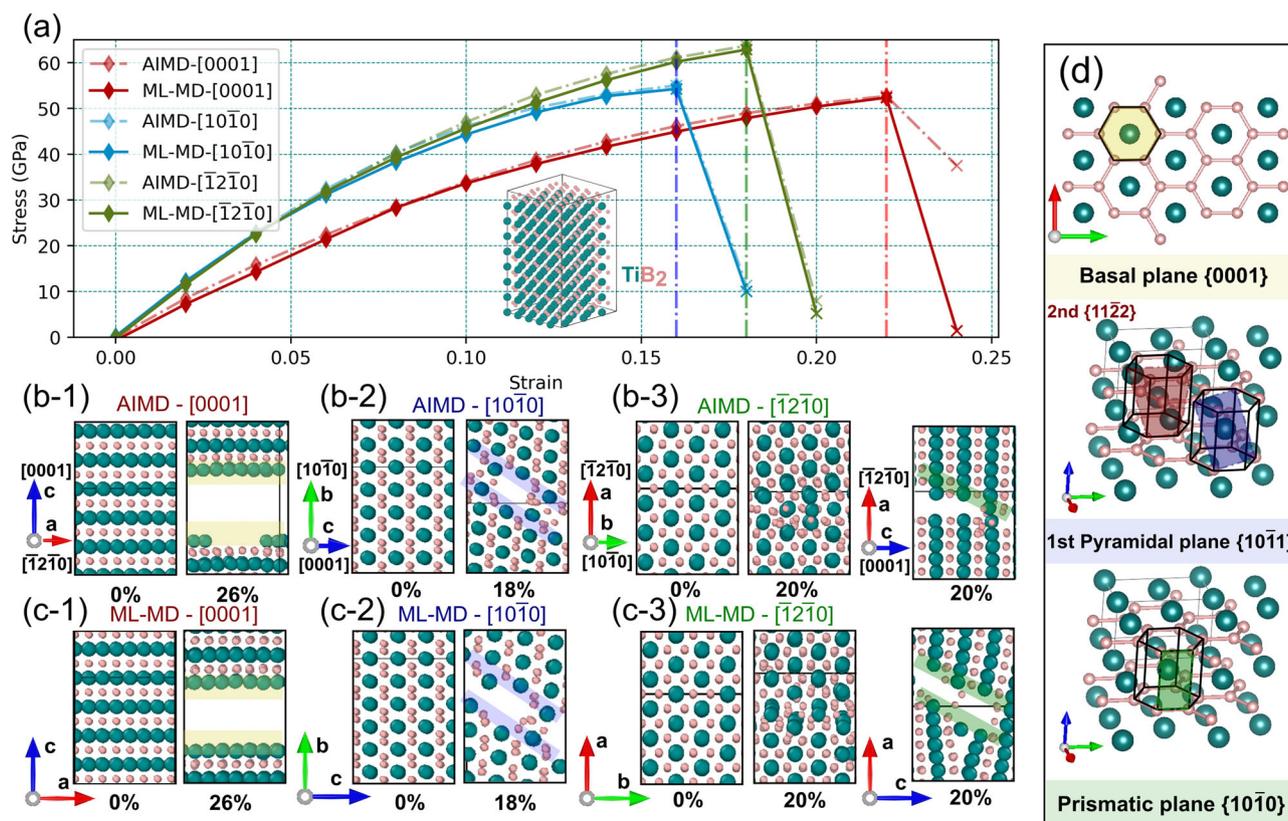
Fig. 2 | Validation of the here-developed MLIPs (MLIP-[0001], MLIP-[10$\bar{1}$0], and MLIP-[$\bar{1}$2$\bar{1}$0]) against atomic scale room-temperature AIMD tensile tests. **a** Comparison of AIMD (dash-dotted line) and ML-MD (solid line) stress/strain curves for TiB$_2$ subject to [0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0] tensile deformation at 300 K, using a 720-atom supercell with dimensions of $\approx (1.52 \times 1.58 \times 2.57)$ nm^3. Only the stress component in the loaded direction is plotted. **b, c** Snapshots of the fracture point in AIMD (b-1, b-2, b-3) and ML-MD (c-1, c-2, c-3). **d** Illustration of fracture surfaces (see also ref. [100]).

MTP developers[60] and by one of us[12] showing that such value corresponds to *accurate* extrapolation[60] (*near-interpolation*[12]).

Procedure 1. MLIP training
  (1) Generate a pool of AIMD configurations.
  (2) Divide the pool into an initial training set (TS$_0$), a learning set (LS), and a validation set (VS) by randomly selecting 0.5%, 79.5%, and 20% of non-overlapping configurations.
  (3) Fit an initial MLIP (MLIP$_0$, trained on TS$_0$). If γ of all configurations in the LS and VS is below $\gamma_{thr} = 2$, exit. Else, build TS$_1$ by adding maximum 10% of configurations from the LS to TS$_0$, selected by the Maxvol algorithm[28] based on their extrapolation grade. Fit a new MLIP (MLIP$_1$, trained on TS$_1$).
  (4) While γ of all configurations in the LS and VS is above $\gamma_{thr} = 2$, build TS$_i$ by adding maximum 10% of configurations from the LS to TS$_{i-1}$, selected by the Maxvol algorithm based on their extrapolation grade. Fit a new MLIP (MLIP$_i$, trained on TS$_i$).

Employing Procedure 1 and the MTP formalism, we fit three MLIPs: MLIP-[0001], MLIP-[10$\bar{1}$0], and MLIP-[$\bar{1}$2$\bar{1}$0]. The training uses snapshots from room-temperature AIMD simulations for a 720-atom TiB$_2$ supercell, uniaxially elongated in the [0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0] crystallographic direction, respectively, with a strain step of 2% (for details of AIMD simulations, see the Methods). The entire pool of AIMD data consists of $\approx 120,000$ configurations, where each loading condition ([0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0]) represents $\approx 1/3$.

The final training sets (the last TS$_i$ in the step (4) of Procedure 1) of MLIP-[0001], MLIP-[10$\bar{1}$0], and MLIP-[$\bar{1}$2$\bar{1}$0] contain 181, 155, and 180 configurations, respectively. The fitting and validation errors, quantified through the residual mean square error (RMSE[59]), of total energies, forces, and stresses do not exceed 2.6 meV atom^{-1}, 0.11 eV Å^{-1}, and 0.30 GPa,

respectively. As follows from Procedure 1, $\gamma < \gamma_{thr} = 2$ for all configurations in the learning set.

MLIPs' validation against atomic scale tensile tests
Following evaluation of the fitting and validation errors (Section Training procedure and fitting initial MLIPs), further validation steps consist in using the above developed MLIPs to run MD simulations (ML-MD) of TiB$_2$ subject to tensile deformation.

During ML-MD, the MLIP's reliability is assessed via:
  (i) Comparison with AIMD predictions of physical, mechanical properties, and fracture mechanisms. In particular, we monitor time-averaged stresses, theoretical strength and toughness, and cleavage on different crystallographic planes.
  (ii) The extrapolation grade. At each ML-MD time step, we calculate γ of the corresponding configuration. Values exceeding *reliable* extrapolation signalise that the MLIP requires up-fitting (i.e., expanding the TS by additional configurations and going back to step (4) of Procedure 1; see Fig. 1b, c).

As suggested by MTP developers[60], we consider $\gamma_{reliable} \leq 10$ as reliable extrapolation. Such choice allows us to develop MLIPs with accuracy similar to the underlying ab initio training set.

Figure 2 a depicts stress/strain curves derived from room-temperature AIMD and ML-MD tensile tests, in which TiB$_2$ supercell ($\approx 10^3$ atoms, $\approx 1.5^3$ nm^3) is loaded in the [0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0] direction, respectively. Each deformation is simulated with a MLIP trained to the respective loading condition. Excellent quantitative agreement between AIMD and ML-MD results indicates reliability of our MLIPs. Specifically, time-averaged stresses in ML-MD differ from AIMD values by 0.07–1.94 GPa, yielding statistical errors RMSE ≈ 1.02 GPa, $R^2 \approx 0.9997$. Stresses

normal to the loaded direction—not vanishing due to the omission of Poisson's effect in both AIMD and ML-MD simulations—are also used for assessing MLIPs' reliability. After fracture, the [0001] stress component recorded in AIMD does not drop to zero. The effect is due to long-range electrostatic effects, which are absent in ML-MD. The extrapolation grade during all ML-MD simulations remains low ($\gamma \leq 5 < \gamma_{\text{reliable}}$), thus suggesting reliable extrapolation.

The calculated stress/strain curves allow us to evaluate TiB$_2$'s theoretical tensile strength and toughness along different crystal directions. We report results of the *ultimate* tensile strength, which corresponds to the global stress maximum during the tensile test[61]. We define the tensile toughness as the integrated stress/strain area. The property describes the ability of an initially defect-free material at absorbing mechanical energy until failure.

The theoretical tensile strengths of TiB$_2$ obtained by ML-MD are 63.7, 55.0, and 52.7 GPa for the $[\bar{1}2\bar{1}0]$, $[10\bar{1}0]$, and $[0001]$ directions, respectively. These differ from AIMD values by maximum 0.8 GPa (0.99%). The ML-MD predicted toughness along $[\bar{1}2\bar{1}0]$, $[10\bar{1}0]$, and $[0001]$ reaches 4.3, 3.1, and 4.3 GPa, respectively, differing from AIMD values by maximum 0.029 GPa (0.67%). Note, however, that theoretical tensile strength and toughness are affected by the supercell size. Strength and toughness values saturate over nm lengthscales (see following sections).

AIMD and ML-MD tensile-testing of TiB$_2$ reveal very similar fracture mechanisms (Fig. 2b–c). Specifically, the fracture surface formed during [0001] deformation almost perfectly aligns with (0001) basal planes (Fig. 2(b-1),(c-1),d). Tensile loading along the[$10\bar{1}0$] direction opens a void diagonally across Ti/B$_2$ layers (Fig. 2(b-2),(c-2)). The fracture surface is parallel to the second order pyramidal planes of the {$11\bar{2}2$} family (Fig. 2d). For the $[\bar{1}2\bar{1}0]$ loading condition, fracture planes are approximately parallel to the {$10\bar{1}0$} prismatic planes (Fig. 2(b-3),(c-3),d). The TiB$_2$ fracture planes predicted in our simulations are consistent with experimentally-characterized slip planes in single-crystal TiB$_2$[50,62].

MLIPs' up-fitting for nanoscale tensile tests

As discussed in the previous section, MLIP-[0001], MLIP-[$10\bar{1}0$], and MLIP-[$\bar{1}2\bar{1}0$] provide reliable description of TiB$_2$'s response to uniaxial tensile loading at the atomic scale. This is indicated by low extrapolation grades ($\gamma \leq 5 < \gamma_{\text{reliable}}$) as well as stress/strain curves and fracture mechanisms in agreement with AIMD results.

As a next step, we carry out tensile tests at the nanoscale. Methodological differences between atomic and nanoscale simulations are listed below.

- **Atomic scale tensile tests** (presented in the previous section) employ 720-atom supercells with dimensions of $\approx (1.5 \times 1.6 \times 2.6)$ nm^3. Strain is incremented at steps of 2% with fixed lattice vectors normal to the loaded direction.
- **Nanoscale tensile tests** employ supercells with four different sizes: $S1$ (12,960 at.), $S2$ (141,120 at.), $S3$ (230,400 at.), and $S4$ (432,000 at.), where $S1 \approx 5^3$ nm^3 and $S4 \approx 15^3$ nm^3. These simulations impose a continuous and homogeneous increase in strain (rate 50 Å s^{-1}) and account for Poisson's contraction.

Employing MLIP-[0001], MLIP-[$10\bar{1}0$], and MLIP-[$\bar{1}2\bar{1}0$] for room-temperature nanoscale tensile tests results in unphysical dynamics (losing atoms) and rapidly increasing extrapolation grades ($\gamma \gg 10^3 \gg \gamma_{\text{reliable}}$) when approaching the fracture point. This indicates that deformation is controlled by formation of extended crystallographic defects, which are absent in atomic scale simulations. Thus, to enable a description of TiB$_2$'s fracture at the nanoscale, our MLIPs require up-fitting (Fig. 1b). Generally, this is a non-trivial task[63,64], since structures causing large γ cannot be directly treated by ab initio calculations.

Below, we describe up-fitting steps leading to MLIP-[4] (schematically depicted in Fig. 1c). We shall see that MLIP-[4] enables reliable simulations of TiB$_2$ tensile deformation at the nanoscale.

Table 1 | Validation of the here-developed MLIP (MLIP-[4], equivalent results are produced by intermediate MLIP-[1–3]) against theoretical and experimental (exp.) lattice constants

	No. of atoms	T	a	c	Source
DFT	720	0	3.027	3.213	This work
ML-MD	720	0	3.027	3.213	This work
AIMD	720	300	3.035	3.218	This work
ML-MD	720	300	3.036	3.218	This work
ML-MD	$(13{-}430) \cdot 10^3$	300	3.036	3.217	This work
Exp.	-	300	3.032	3.229	Ref. 92
Exp.	-	300	3.029	3.229	Ref. 93
Exp.	-	300	3.021	3.230	Ref. 94
AIMD	720	1200	3.056	3.249	This work
ML-MD	720	1200	3.056	3.249	This work
ML-MD	$(13{-}430) \cdot 10^3$	1200	3.047	3.239	This work

The a and c values for TiB$_2$ (in Å) are shown at temperature T (K). Experimental values, refs. 92 and 93,94, are for TiB$_2$ powder and thin films, respectively.

- We produce MLIP-[1] by up-fitting MLIP-[0001], where the LS is expanded by final TSs of MLIP-[$10\bar{1}0$] and MLIP-[$\bar{1}2\bar{1}0$]. MLIP-[1] accurately simulates atomic scale tensile properties, but does not well describe nanoscale deformation and fracture in TiB$_2$ ($\gamma \gg \gamma_{\text{reliable}}$).
- We up-fit MLIP-[1] using three different LSs, producing MLIP-[2], MLIP-[3], and MLIP-[4]. MLIP-[2] and MLIP-[3] learn from AIMD snapshots of TiB$_2$ equilibrated at 1200 K (MLIP-[2]), and sequentially elongated in the [0001] direction until cleavage (MLIP-[3]). MLIP-[4] learns from AIMD snapshots of TiB$_2$ elongated by 150% in the [0001] direction, initializing atoms at ideal lattice sites and equilibrating at 300 and 1200 K under fixed volume and shape. Such large strain quickly induces fracture, thus providing additional information for training MLIP on highly deformed lattice environments and surface properties.MLIP-[2], MLIP-[3], and MLIP-[4] all provide results in agreement with AIMD tensile tests at the atomic scale ($\gamma \leq 5 < \gamma_{\text{reliable}}$, Supplementary Fig. 2). Fracture mechanisms and elastic constants are also correctly reproduced. At the nanoscale, MLIP-[2] and MLIP-[3] exhibit lower γ than MLIP-[1]. However, the sought improvement ($\gamma \leq \gamma_{\text{reliable}}$) is achieved only by MLIP-[4], which will be used to carry out nanoscale ML-MD simulations of TiB$_2$ deformation.

The reliability of MLIP-[4] is indicated by low γ values, realistic description of structural changes during nanoscale tensile tests (see Supplementary Fig. 3, and following section), and excellent agreement with ab initio and experimental values of TiB$_2$ lattice parameters (Table 1).

To explain why MLIP-[4] enables nanoscale tensile tests, one needs to analyse the corresponding training set, TS(MLIP-[4]). In Fig. 3, we visualize selected characteristics of TS(MLIP-[4]) in comparison to the training set of MLIP-[1], TS(MLIP-[1]), where the latter is not applicable to simulate TiB$_2$'s fracture at the nanoscale. The radial distribution function (RDF, Fig. 3a) and bond angle distribution analysis (Supplementary Fig. 4) suggest minor geometrical differences between structures contained in TS(MLIP-[1]) and TS(MLIP-[4]). Their total energy and stress distribution, however, differ significantly (Fig. 3b). In particular, TS(MLIP-[4]) contains atomic configurations with higher total energy and higher total energy in combination with higher stress in principal crystallographic axes, which are missing in TS(MLIP-[1]). An illustration of structures from the two training sets is given in Fig. 3c. The chosen snapshots indicate that TS(MLIP-[4]) provides a variety of atomic environments relevant for simulations of non-stoichiometry, locally amorphous regions, and surfaces, which are likely to be helpful also for nanoscale simulations.
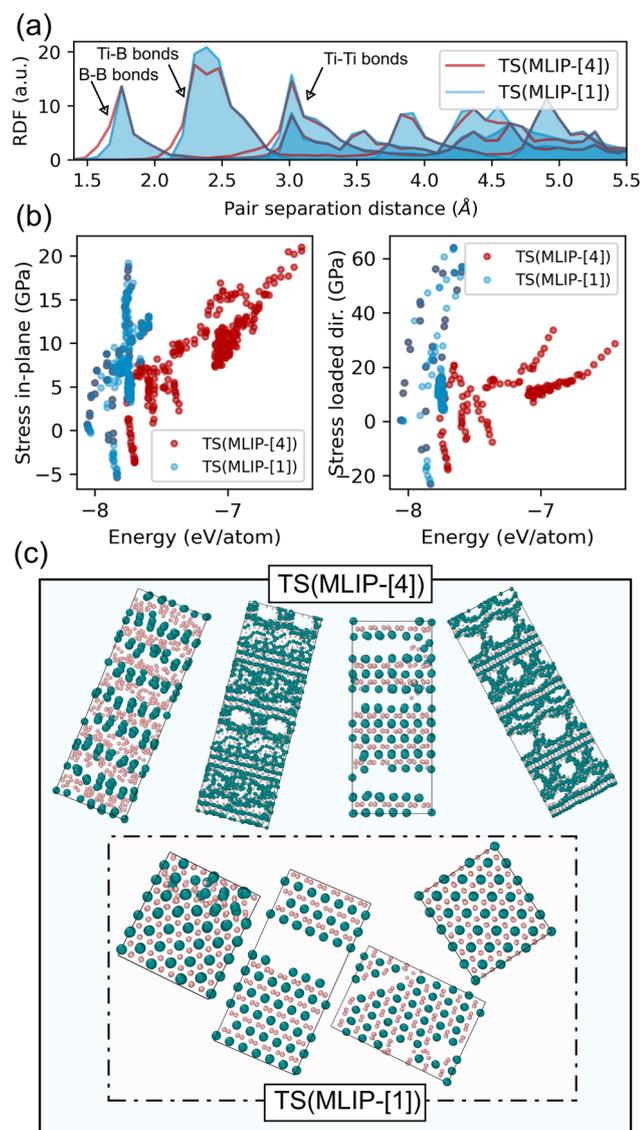
Fig. 3 | Qualitative differences between configurations in the training sets (TSs) of the here-developed MLIPs suitable only for the atomic scale ML-MD tensile tests (MLIP-[1]) or for both atomic and nanoscale ML-MD tensile tests (MLIP-[4]). a Radial distribution function (RDF, with 5.5 Å cutoff) for B–B, Ti–B, and Ti–Ti bonds (integrated over all configurations). **b** Stress components (in-plane and in the loaded direction) vs. total energy of all configurations in the training set. **c** Snapshots of representative structures from the two training sets.

4 Size effects in tensile response of TiB$_2$

Equipped with the above developed MLIP-[4], in this section we discuss TiB$_2$'s response to room-temperature uniaxial tensile loading from atomic to nanoscale. Recall that an important difference between our atomic and nanoscale tensile tests is that the former disregard Poisson's contraction.

The stress/strain curves calculated by ML-MD at room temperature are depicted in Fig. 4. For strains below $\approx 10\%$, the tensile stresses computed for each loading direction remain unaltered by an increase in supercell sizes from $\approx 10^3$ to 10^6 atoms. Such overlap indicates consistency in elastic constants derived from atomic and nanoscale models. As already seen in atomic scale simulations, the basal plane is elastically isotropic, which is shown by the same initial slope of stress/strain curves for [$\bar{1}2\bar{1}0$] and [$10\bar{1}0$] elongation and in line with experimental reports for hexagonal crystals[65]. Due to Poisson's contraction, however, differences in stress/strain curves emerge beyond the linear-elastic regime. A shrinkage of the lattice parameters

normal to the applied tensile strain yields Poisson's ratio ($\nu \approx 0.127$ (The Poisson's contraction was calculated as $\nu = -\frac{d\varepsilon_{compressed}}{d\varepsilon_{elongated}}$, where the $d\varepsilon_{compressed}$ ($d\varepsilon_{elongated}$) is the lattice parameter shrinkage (increment) orthogonal (parallel) to the loading direction. The presented value is an average of Poisson's ratios for both in-plane directions.)) consistent with the value obtained from elastic constants ($\nu \approx 0.113$, see Table 2). Approaching the fracture point, differences between TiB$_2$'s tensile behavior at atomic and nanoscale become more apparent. For example, while atomic-scale simulations indicate that the [0001] direction is the softest, tensile tests at the nanoscale show that [$10\bar{1}0$] elongation returns the lowest strength value (Table 3). Even more pronounced size effects are expected for ceramics which exhibit plastic behavior upon tensile loading.

The TiB$_2$'s theoretical strength and toughness calculated during nanoscale simulations at room temperature remain essentially unchanged for supercell sizes increasing from $\approx 5^3$ nm^3 to $\approx 15^3$ nm^3 (see S1–S4 results in Table 3). Specifically, the [$\bar{1}2\bar{1}0$] direction exhibits the highest tensile strength (≈ 56 GPa), followed by the [0001] direction (≈ 54 GPa), and the [$10\bar{1}0$] direction (≈ 51 GPa). The [0001] direction exhibits the highest toughness (≈ 4.80 GPa), followed by the [$\bar{1}2\bar{1}0$] direction (≈ 3.37 GPa), and the [$10\bar{1}0$] direction (≈ 2.78 GPa).

Besides characterizing directional dependence of tensile strength and toughness in dislocation-free monocrystals, ML-MD nanoscale simulations also provide insights into crack nucleation and growth mechanisms. These are illustrated by Figs. 5–7. Results of ML-MD atomic scale simulations are included for comparison.

During atomic-scale simulations of [0001] elongation, all atoms in TiB$_2$ vibrate close to their ideal lattice sites until a sudden brittle cleavage induces the formation of two surfaces almost perfectly parallel with (0001) basal planes (Fig. 5, row 1). At the nanoscale, fracture is initiated by opening of voids accompanied by local necking which produces lattice re-orientations (Fig. 5, row 2 and 3). Rapid void coalescence and fraying of ligaments results in corrugated fractured surfaces, predominantly with (0001) orientation. Following the stress release, inner parts of the crystal relax back to the ideal TiB$_2$ lattice sites.

The S1 supercell yields in only one region (Fig. 5, row 2). The larger S2, S3, and S4 supercells do not fracture in two pieces but nucleate cracks with size of few nm. For S4 supercells, the phenomenon is depicted in Fig. 5, row 3. The fractured surfaces do not align only with the basal planes but also with the {$10\bar{1}1$} first order pyramidal planes (see notation in Fig. 2d). Volumetric strain analysis (Fig. 5d, e) highlights locally increased tensile strain concentration surrounding small voids (see TiB$_2$ slice at $\approx 27\%$ strain) due to decreased interplanar spacings between Ti and B layers (predominantly due to [0001] compression) above and below the voids. The large size of S4 models allows cracks to propagate along different directions, thus offering a detailed view of fracture patterns.

For the [$10\bar{1}0$] tensile test, size effects in fracture mechanisms are compared in Fig. 6. At the atomic scale, two voids open diagonally across Ti/B layers (Fig. 6, row 1). At the nanoscale, we observe nucleation of V-shaped cracks, as illustrated for the S1 and the S4 supercell (Fig. 6, row 2 and 3, and panels c, d), where S4 additionally reveals lattice rotation near the V-shaped defects. We infer that loading in the direction of strong covalent B–B bonds most often induces crack deflection and fracture at {$11\bar{2}2$} family of surfaces parallel to the second order pyramidal planes.

Changing to the [$\bar{1}2\bar{1}0$] tensile deformation, atomic scale simulations predict fracture along {$10\bar{1}0$} prismatic planes (Fig. 7, row 1). This is underpinned also by nanoscale ML-MD (Fig. 7, row 2 and 3), suggesting that crack growth often occurs both orthogonally and diagonally across Ti/B layers (see the dashed line with arrow in Fig. 7e).

A direct comparison between experimental and ML-MD results of TiB$_2$ mechanical properties and preferred fracture planes would require synthesis and tensile testing of TiB$_2$ monocrystals. Unfortunately, TiB$_2$ ceramics are typically synthesized as thin films on substrates, which renders measurements of tensile strength essentially unfeasible. Nevertheless, as
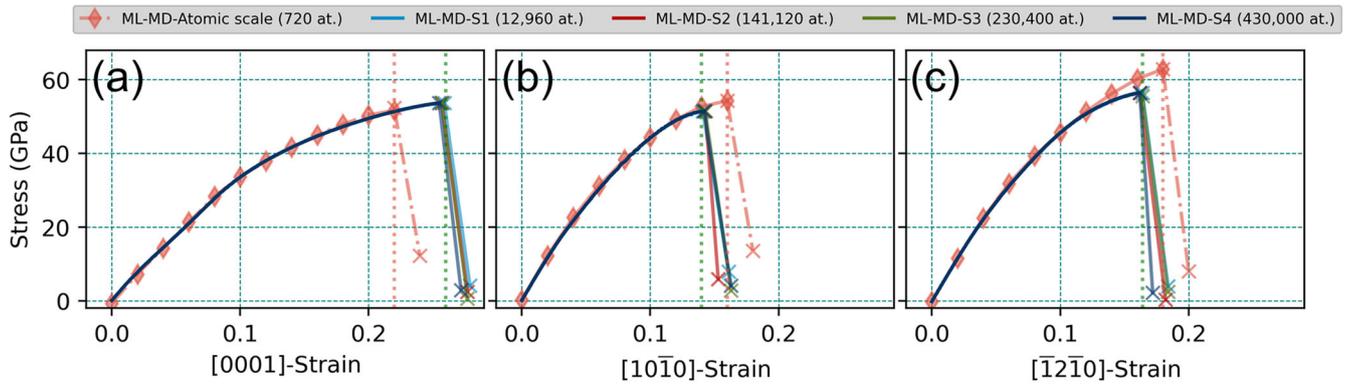
Fig. 4 | Comparison of AIMD and ML-MD stress/strain curves calculated for TiB$_2$ at room temperature. The ML-MD tensile stresses, obtained using MLIP-[4], are plotted as a function of TiB$_2$ elongation parallel to **a** [0001], **b** [10$\bar{1}$0], and **c** [$\bar{1}$2$\bar{1}$0] directions. The orange diamonds correspond to atomic scale ML-MD simulations (720 at), while the solid lines correspond to nanoscale ML-MD simulations (12,960–430,000 at), as defined at the beginning of this section. Note that the theoretical strength of defect-free crystal models represents an ideal upper bound of strength attainable in actual ceramics. Much lower stresses are expectable in experiments due to, e.g., nanostructural defects.

Table 2 | Validation of the here-developed MLIP (MLIP-[1]) against theoretical and experimental (exp.) room-temperature elastic constants, C_{ij}

	No. of atoms	T	C_{11}	C_{33}	C_{44}	C_{12}	C_{13}	E	B	G	ν	Source
DFT	192	300	640	446	251	62	91	574	244	260	0.106	Ref. 95
AIMD	720	300	588	430	252	79	111	547	244	243	0.126	This work
ML-MD	720	300	588	409	261	85	98	554	236	246	0.113	This work
Exp.	—	300	660	432	260	48	93	565	244	266	0.099	Ref. 96
Exp.	—	300	588	503	238	72	84	575	249	255	0.114	Ref. 97

The C_{ij} values for TiB$_2$ (in GPa, at temperature T (K)) are shown together with the polycrystalline bulk modulus, B (in GPa), shear modulus, G (in GPa), Young's modulus, E (in GPa), and Poisson's ratio, ν, compared to reference ab initio and experimental (exp.) data. ref. 96 and ref. 97 is for TiB$_2$ single and polycrystal, respectively. AIMD and ML-MD elastic constants were evaluated following ref. 86, based on a second-order polynomial fit of the [0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0] stress/strain data (C_{11}, C_{12}, C_{13}, C_{33}) and of the (0001)[$\bar{1}$2$\bar{1}$0], (10$\bar{1}$0)[$\bar{1}$2$\bar{1}$0], and (10$\bar{1}$0)[0001] shear stress/strain data (C_{44}), assuming strains up to 4%. For details see the Methodology section. Note that values from ref. 95 were based on 0 K ab initio calculations coupled to phonon-theory assessments of thermal expansion.

Table 3 | Size and temperature effects on mechanical properties of TiB$_2$

No. of atoms	T (K)	Dimensions (nm)			Strength (GPa)			Toughness (GPa)			Fracture strain (%)		
		a	b	c	[0001]	[10$\bar{1}$0]	[$\bar{1}$2$\bar{1}$0]	[0001]	[10$\bar{1}$0]	[$\bar{1}$2$\bar{1}$0]	[0001]	[10$\bar{1}$0]	[$\bar{1}$2$\bar{1}$0]
720	300	1.51	1.58	2.57	52.72	55.01	63.69	4.33	3.11	4.32	22.0	16.0	18.0
12,960 (S1)	300	4.55	4.73	5.15	53.71	51.36	56.40	4.83	2.77	3.38	26.4	14.3	16.8
141,120 (S2)	300	10.63	11.05	10.30	53.69	51.44	56.41	4.81	2.78	3.37	26.4	14.3	16.8
230,400 (S3)	300	12.14	12.63	12.87	53.71	51.43	56.38	4.82	2.78	3.37	26.3	14.3	16.8
432,000 (S4)	300	15.18	15.79	15.45	53.67	51.47	56.42	4.80	2.78	3.37	26.2	14.2	16.8
720	1200	1.53	1.59	2.60	43.87	45.21	51.34	3.27	2.29	3.07	20.0	16.0	16.0
432,000 (S4)	1200	15.26	15.86	15.52	43.27	41.30	44.67	3.45	1.96	2.32	21.6	12.8	14.1

The directional response to strain is assessed for supercells with sizes ranging from the atomic to the nanoscale. The mechanical properties of TiB$_2$ are extracted from ML-MD stress/strain data in Fig. 4.

mentioned above, the fracture planes observed in ML-MD simulations are consistent with slip systems known to operate in TiB$_2$ at room temperature[50,62]. It is also worth noting that the high hardness measured for [0001]-textured TiB$_2$ thin films, 41–53 GPa[45,47,48], is consistent with large strength values predicted for TiB$_2$ by atomistic simulations.

In general, the fracture strength measured for hard ceramics is one-to-two orders of magnitude lower than the theoretical strength. The discrepancy is due to premature cracking initiated at native structural defects in actual materials. For example, microcantilever bending experiments conducted on hard polycrystalline Ti-Al-N evidence intergranular fracture, which severely limits the material strength to values below 10 GPa[66]; much smaller than corresponding theoretical strength results[67]. Nevertheless, prior to crack growth, the tensile stress accumulated at structural flaws of ceramics can locally reach several tens of GPa[68]. This reconciles stress states predicted by atomistic simulations with real mechanical-testing experiments. Likewise, the elongation withstood by defect-free ceramic models during tensile testing simulations is indicative of strain locally produced (nm lengthscale) in specimens subject to load (see fracture strains in Fig. 4).

Other loading conditions and MLIP's transferability

Our MLIP (MLIP-[4]) has been specifically developed and optimized to target atomic to nanoscale simulations of TiB$_2$ subject uniaxial
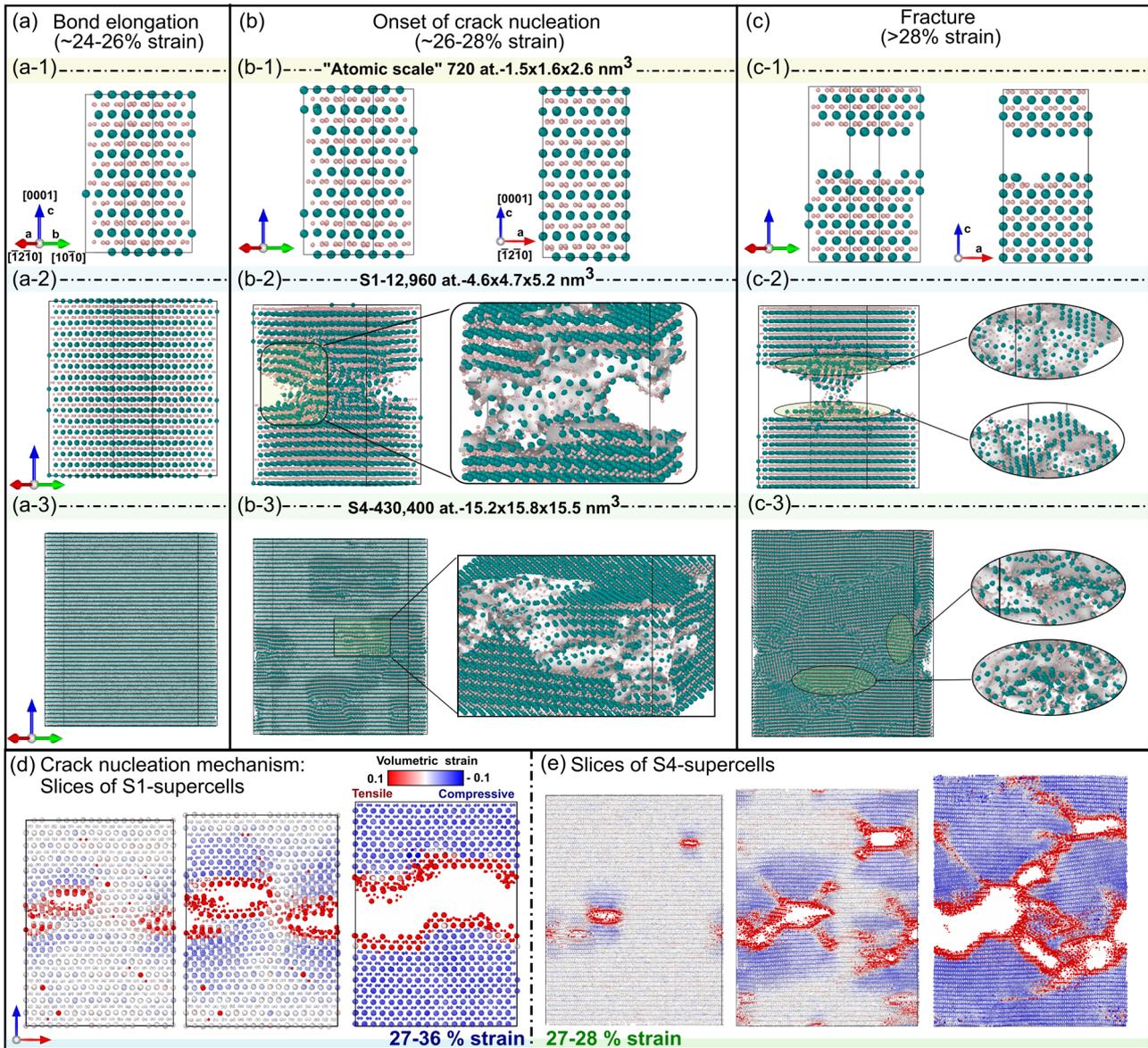
Fig. 5 | Representative ML-MD snapshots of TiB$_2$ strained along [0001] at 300K. Upper (x-1), middle (x-2), and lower (x-3) panels show results of simulations over different lengthscales. Key deformation stages: **a** bond elongation in the loaded direction, **b** onset of crack nucleation, and **c** fracture. Thin slices of the nanoscale **d** S1 and **e** S4 supercells color-coded based on volumetric strain (using the corresponding equilibrium structure as reference). Red (blue) regions denote high tensile (compressive) strain.

tensile loading until fracture. Examination of the underlying training set (Fig. 3c) indicated a variety of atomic environments including, e.g., different Ti/B stoichiometries, locally amorphous regions, or surface structures. In this section, we discuss the MLIP's transferability to the description of loading conditions, phases, and chemical environments for which it has not been explicitly trained on. In most cases, we conclude that up-fitting is necessary to guarantee quantitative agreement with ab initio results.

Accuracy of the predicted observables (e.g., shear strengths or surface energies) is presented in the context of extrapolations grades, allowing to identify types of configurations beneficial for further up-fitting, thus, broadening the MLIP's applicability.

- **High-temperature tensile deformation of TiB$_2$.** Since TiB$_2$ is an UHTC (see the Introduction), modelling its mechanical behaviour at elevated temperature is of high practical relevance. Here we choose 1200 K which is close to the highest anti-oxidation temperature of TiB and TiC reported experimentally[69].

Atomic scale [0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0] tensile tests at 1200 K show excellent quantitative agreement with AIMD simulations at the same temperature (see Supplementary Fig. 5). Specifically, differences from AIMD-calculated stresses are 0.01–3.54 GPa, resulting in statistical errors RMSE ≈ 1.85 GPa, R^2 ≈ 0.9997. Extrapolations grades indicate reliable extrapolation ($\gamma \leq 7 < \gamma_{reliable}$).

TiB$_2$'s theoretical tensile strength at 1200 K decreases by about 17–19% compared to 300 K. For tensile toughness, our simulations predict ≈ 25% decrease. Fracture mechanisms remain qualitatively unchanged with respect to 300 K.

- **Room-temperature shear deformation of TiB$_2$.** Simulations of shear deformation provide useful insights for understanding of how dislocations nucleate and move in generally brittle UHTCs[70,71]. Furthermore, as diborides typically crystallize in layered structures (α, γ, ω), which correspond to different stacking of transition metal planes, shear deformation may induce plastic deformation via faulting, twinning, and phase transformation[72]. Based on experimental characterization of
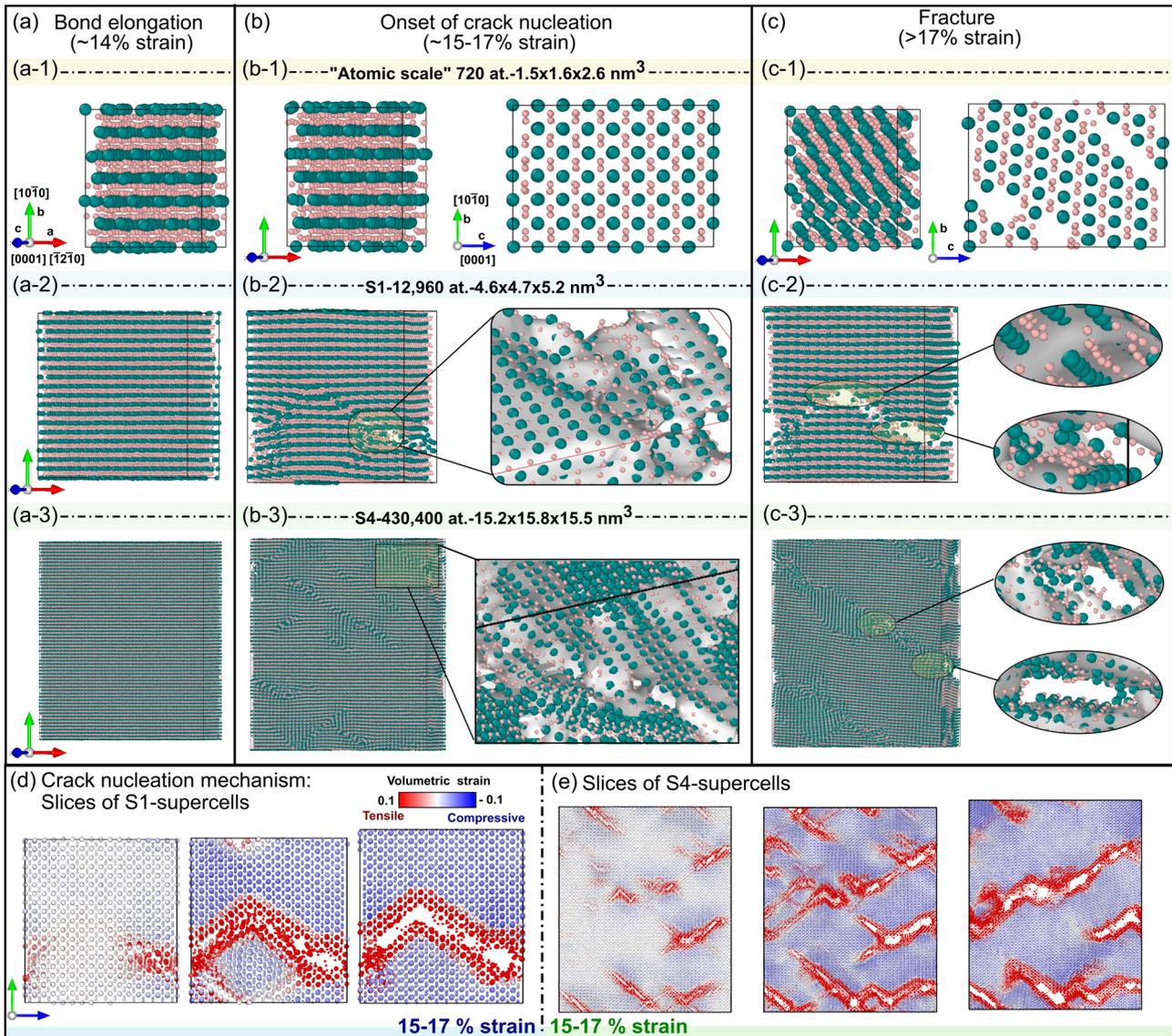
Fig. 6 | Representative ML-MD snapshots of TiB$_2$ strained along [10$\bar{1}$0] at 300K. Upper (x-1), middle (x-2), and lower (x-3) panels show results of simulations over different lengthscales. Key deformation stages: **a** bond elongation in the loaded direction, **b** onset of crack nucleation, and **c**) fracture. Thin slices of the nanoscale **d** S1 and **e** S4 supercells color-coded based on volumetric strain (using the corresponding equilibrium structure as reference). Red (blue) regions denote high tensile (compressive) strain.

room-temperature slip in single-crystal TiB$_2$[50,62], we simulate shear deformation along (0001)[$\bar{1}2\bar{1}0$], (10$\bar{1}$0)[$\bar{1}2\bar{1}0$], and (10$\bar{1}$0)[0001] slip systems.

Stress evolution during atomic scale ML-MD shear deformation (Fig. 8a) agrees well with equivalent AIMD simulations. This is particularly the case for strains below ≈ 20%, where stresses differ from AIMD values by 0.01–5.08 GPa (yielding statistical errors RMSE ≈ 3.72 GPa, R^2 ≈ 0.9993) and γ is close to reliable extrapolation ($\gamma < 14$). This is a good result, if one considers that the training set did not contain sheared configurations. Shear strains above ≈ 20% induce notably larger discrepancies in stresses (differing from AIMD by 5.4–8.5 GPa) and increased γ ($\gamma \leq 26$). The main reason is that lattice slip—responsible for a partial stress release during shearing—does not occur at the same strain step.

The shear strengths predicted by ML-MD for (0001)[$\bar{1}2\bar{1}0$], (10$\bar{1}$0)[$\bar{1}2\bar{1}0$], and (10$\bar{1}$0)[0001] deformation, (49, 57, and 51 GPa), are ≈ 8% lower than AIMD values (58, 72, and 68 GPa). Nevertheless, shear-induced structural changes observed during ML-MD correctly reproduce AIMD results (Fig. 8b–d).

When subject to (0001)[$\bar{1}2\bar{1}0$] shearing, TiB$_2$ undergoes slip on the basal plane. The mechanism—activated for (0001)[$\bar{1}2\bar{1}0$] shear strain of ≈ 24% – restores atoms to their ideal lattice sites (Fig. 8b). (10$\bar{1}$0)[$\bar{1}2\bar{1}0$] shearing induces plastic flow on both ($\bar{1}2\bar{1}0$)[10$\bar{1}$0] and (10$\bar{1}$0)[$\bar{1}2\bar{1}0$] slip systems. The mechanisms are activated at strains near 30% and 50% (Fig. 8c). Both are accompanied by displacements of Ti and B atoms from ideal TiB$_2$ lattice sites. Similar to the results in (Fig. 8c), shearing along (10$\bar{1}$0)[0001] activates different slip systems (Fig. 8d). TiB$_2$(0001) lattice layers glide along [10$\bar{1}$0] at ≈ 30% strain. A further increase in strain to ≈ 50% induces glide of (10$\bar{1}$0) planes along the [0001] direction. The latter process results in significant displacements of B atoms and formation of stacking faults, as indicated by horizontal green lines in Fig. 8d (panels on the right).

The excellent agreement between ML-MD and AIMD stress/strain curves within the elastic shear response, accompanied by reasonably good agreement near TiB$_2$ yielding, suggests that up-fitting our MLIP[4] to accurately model shear deformation would benefit from adding configurations near TiB$_2$'s shear instabilities to the training set.
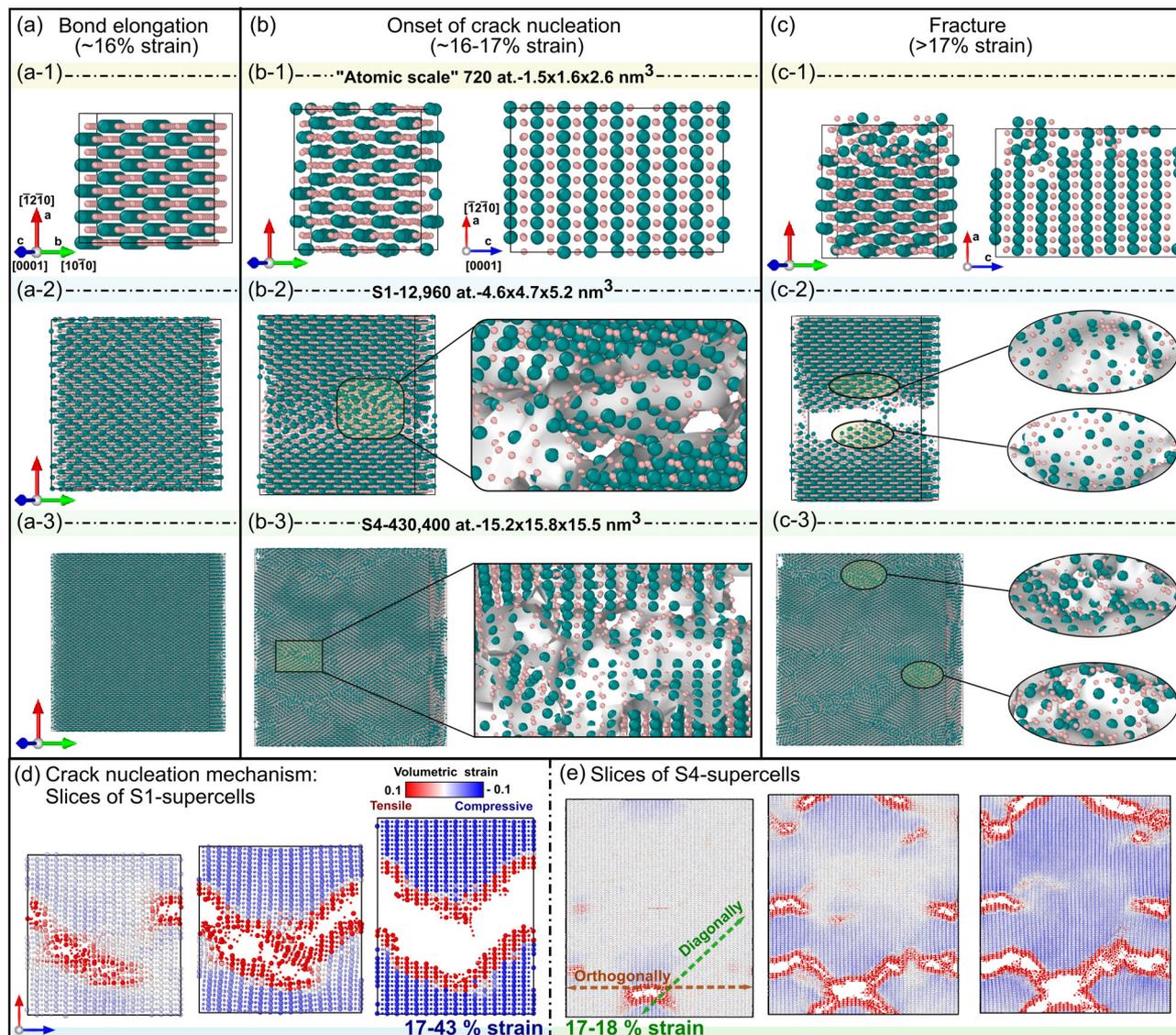
Fig. 7 | Representative ML-MD snapshots of TiB$_2$ strained along $[\bar{1}2\bar{1}0]$ at 300K. Upper (x-1), middle (x-2), and lower (x-3) panels show results of simulations over different lengthscales. Key deformation stages: **a** bond elongation in the loaded direction, **b** onset of crack nucleation, and **c** fracture. Thin slices of the nanoscale **d** S1 and **e** S4 supercells color-coded based on volumetric strain (using the corresponding equilibrium structure as reference). Red (blue) regions denote high tensile (compressive) strain.

- **Room-temperature volumetric compression of TiB$_2$.** Training on snapshots of compressed structures may be important not only for simulations of e.g., uniaxial compression or nanoindentation, but also in order to correctly account for Poisson's contraction during tensile deformation. For TiB$_2$, the relatively low Poisson's ratio (Table 2) is manifested by a rather small lateral shrinkage of the supercell during nanoscale tensile testing simulations. Here we simulate severe volumetric compression of TiB$_2$ at room temperature. *All* lattice vectors are compressed by up to 12% to show rapidly growing γ values and suggest how to improve MLIP reliability by up-fitting.

  As shown in Fig. 8e, f, compression-induced stresses along main crystallographic directions are indeed extremely large. In AIMD, they exceed 50 GPa for a 5% compression and reach $\approx$ 150 GPa for a 10% compression. Our MLIP yields satisfactory agreement with AIMD for volumetric compression of 1–2%, with stress tensor components differing from ab initio values by less than 1.83 GPa (9.57%) and γ indicating reliable extrapolation ($\gamma \leq 10 = \gamma_{\text{reliable}}$). A further increase in compression to 10%, however, causes increasing deviations from AIMD stresses and $\gamma \approx 10^2 – 10^3$.

We illustrate the effect of up-fitting by producing a new MLIP (MLIP-[4]$_{Plus}$) which learns from AIMD snapshots of 10% volumetrically-compressed TiB$_2$ (added to the LS of MLIP-[4]). This not only greatly improves accuracy for the 10% compression (stress differences are maximum 0.79 GPa (0.48%) and $\gamma \leq 2$) but also within the entire tested compression range (see red data points in Fig. 8e, f).

- **Surface energies of TiB$_2$.**

  Although our training set did not contain ideal surfaces, environments describing TiB$_2$'s fracture may facilitate reasonable predictions for energies of low-index surfaces. To test this hypothesis, we evaluate the energies of formation, E_{surf}, of (0001), $(\bar{1}2\bar{1}0)$, and $(10\bar{1}0)$ surfaces, i.e., orthogonal to simulated tensile loading directions (Section MLIPs'validation against atomic scale tensile tests–Size effects in tensile response of TiB$_2$).

  The MLIP-predicted E_{surf} are consistent with equivalently produced ab initio values (Table 4). The differences are relatively small: 0.03 J m^{-2} (1.40%) for $E_{\text{surf}}(0001)$, 0.04 J m^{-2} (1.68%) for $E_{\text{surf}}(\bar{1}2\bar{1}0))$, and 0.13 J m^{-2} (5.75%) for $E_{\text{surf}}(10\bar{1}0)$, as underlined also by low extrapolation grades ($\gamma < \gamma_{\text{reliable}}$). ML-MS and DFT
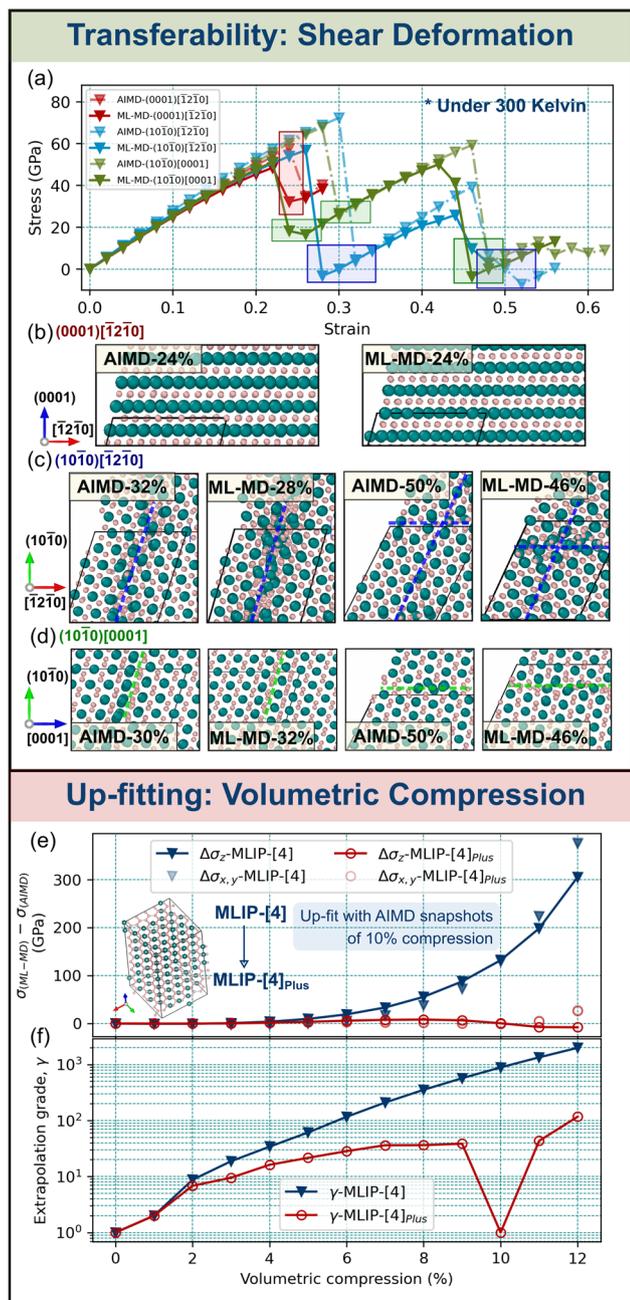
Fig. 8 | Illustration of simulations to which the here-developed MLIP (MLIP-4) is transferable (a–d) or for which it requires up-fitting (e–f). a Comparison of AIMD (dash-dotted line) and ML-MD (solid line) stress/strain curves for TiB$_2$ subject to $(0001)[\bar{1}2\bar{1}0]$, $(10\bar{1}0)[\bar{1}2\bar{1}0]$, and $(10\bar{1}0)[0001]$ room-temperature atomic-scale shear deformation. **b–d** Representative snapshots at strain steps marked by shaded rectangles in **a**. The dashed lines in **b–d** guide the eye for slip directions described in the text. **e** Differences in ML-MD and AIMD stresses ($\sigma_{(\text{ML-MD})} - \sigma_{(\text{AIMD})}$), resolved in the basal plane and [0001] direction ($\sigma_{x,y}$ and σ_z) of TiB$_2$ subject to room-temperature volumetric compression, plotted as a function of the compression percentage. **f** Blue and red data points indicate maximum extrapolation grades returned by MLIP-[4] and its up-fitted version, MLIP-[4]$_{Plus}$, during TiB$_2$ compression.

calculations of the present work indicate that the TiB$_2$($\bar{1}2\bar{1}0$) surface is energetically more stable than the basal plane. This is surprising, given that formation of ($\bar{1}2\bar{1}0$) surfaces requires breaking strong, covalent B-B bonds. That the basal plane of TiB$_2$ is not the one with lowest energy has also been indicated by previous DFT tudies[73,74].

Table 4 | Transferability of the here-developed MLIP (MLIP-[4]) in molecular statics (MS) calculations of surface energies, E_{surf} (J m^{-2}), for low-index surfaces of TiB$_2$

	Surface	E_{surf}	Source
ML-MS	(0001)	3.80	This work
DFT	(0001)	3.80	This work
DFT	(0001)	4.50–4.72	Ref. 98
DFT	(0001)	4.22–4.24	Ref. 73
DFT	(0001)	4.14	Ref. 99
DFT	(0001)	≈ 4.20	Ref. 74
ML-MS	(10$\bar{1}$0)	3.98	This work
DFT	(10$\bar{1}$0)	4.12	This work
DFT	(10$\bar{1}$0)	≈ 4.10	Ref. 74
ML-MS	($\bar{1}$2$\bar{1}$0)	3.42	This work
DFT	($\bar{1}$2$\bar{1}$0)	3.57	This work
DFT	($\bar{1}$2$\bar{1}$0)	4.20–4.83	Ref. 73
DFT	($\bar{1}$2$\bar{1}$0)	≈ 4.02	Ref. 74

Our calculations consider stoichiometric TiB$_2$ surfaces, i.e., with 1:2 Ti-to-B ratio.

Nevertheless, MLIP up-fitting on higher-accuracy DFT data would be needed to verify the relationship among TiB$_2$ surface energies.

- **Off-stoichiometric TiB$_2$ structures and other phases**. Our MLIP was trained to snapshots of TiB$_2$ (AlB$_2$-type phase, P6/mmm) with a perfect stoichiometry (speaking of the entire supercell). Visualization of the training set (Fig. 3c), however, indicates the presence of atomic environments with various Ti-to-B ratios as well as bond lengths and angles different from those in TiB$_2$. This may be useful for simulations of e.g., vacancy-containing TiB$_2$ structures commonly reported by synthesis[36] or other phases in the Ti–B phase diagram[75].

  To investigate transferability to other phases, we use MS calculations to find the ground-state of known phases from the Ti–B phase diagram[75,76]: Ti$_2$B (tetragonal, I4/mcm), Ti$_3$B$_4$ (orthorhombic, Immm), and TiB (orthorhombic Pnma) (The supercell sizes are always ≈ 700 atoms, i.e., comparable to that used for TiB$_2$.). Additionally, we equilibrate the TiB$_2$ phase with Ti, B, or combined Ti and B vacancies: Ti$_{36}$B$_{71}$, Ti$_{35}$B$_{72}$, and Ti$_{35}$B$_{70}$. For all calculations, extrapolation grades ($\gamma \approx 10^2$–10^4) are far beyond *reliable* extrapolation. In terms of total energies (E_{tot}), and lattice parameters (a, b, c), the largest deviation from ab initio values is exhibited by Ti$_3$B$_4$ (10% and 2.5% differences on E_{tot} and c, respectively).

  Simulations of other stoichiometries and phases therefore require up-fitting (not necessarily due to poor accuracy but especially due to high uncertainty, $\gamma \gg \gamma_{reliable}$). To illustrate the up-fitting effect, MLIP-[4] learns from additional ab initio snapshots: from 0 K equilibration of a 780-atom Ti$_3$B$_4$ supercell. Prior to up-fitting, equilibration of Ti$_3$B$_4$ yields $\gamma \geq 10^4$. Afterwards, $\gamma \leq 5 < \gamma_{reliable}$ and E_{tot}, a and c deviate from ab initio values by 4.18%, 0.07%, and 0.73%.

Viability of our training strategy for modelling tensile deformation in other ceramics

To illustrate general applicability of the proposed training strategy, we develop MLIPs for 5 other ceramic systems. Once more, we emphasize that the developed MLIPs are primarily targeted to simulations of uniaxial tensile loading at room temperature. The chosen materials are

- hexagonal α-TaB$_2$[72], which serves as an example of changing the transition metal while keeping the same crystal structure,
- hexagonal ω-WB$_2$[72] and γ-ReB$_2$[72], i.e., examples of changing the transition metal as well as the phase,
- cubic NaCl-type TiN[67], which serves as example of ceramic system with different non-metal species, different lattice symmetry, and bonding

character that is less covalent but more ionic and metallic than TiB_2 and diborides in general,

- orthorhombic, nanolaminated Ti_2AlB_2 (a MAB phase[77]), which is example of ternary system with different crystal structure and mixture of ceramic-like and metallic-like bonding.

Training data generation, up-fitting, and validation follow steps carried out for TiB_2 in Sections Training procedure and fitting initial MLIPs–MLIPs' up-fitting for nanoscale tensile tests. Again, the validation consists in (i) evaluating fitting and validation errors with respect to the training set and a meaningful validation set, (ii) comparing the predicted physical properties with equivalently calculated ab initio values, and (iii) monitoring the extrapolation grade during ML-MD simulations.

Fitting and validation errors of energies, forces, and stresses (RMSE < 3 meV atom^{-1}, < 0.15 eV Å^{-1}, and < 0.25 GPa, respectively) are similar to those evaluated for TiB_2 in Section Training procedure and fitting initial MLIPs and close to the accuracy of the AIMD training set. Extrapolation grades for all configurations in the validation set are below the *accurate* extrapolation threshold ($\gamma < 2$).

Figure 9 exemplifies validation of ML-MD tensile-stress/elongation curves against corresponding atomic-scale AIMD results. The intention is to demonstrate applicability of our training approach to different ceramic systems. Hence, we omit discussion of deformation and fracture mechanisms which, however, are consistent with AIMD results. The time-averaged stresses recorded during ML-MD tensile tests differ from the corresponding AIMD values by 0–3.3 GPa, yielding statistical errors RMSE ≈ 2.0 GPa, $R^2 \approx 0.99$. The discrepancy is partly due to stochastic stress fluctuations, which may also onset fracture at slightly different strains in independent MD runs (compare ML-MD and AIMD results for TaB_2, ReB_2, and TiN in Fig. 9). Previous molecular dynamics tensile-testing investigations demonstrated that the statistical uncertainties on TiN elongation at fracture are comparable to the strain increment[68].

The theoretical tensile strengths of TaB_2, WB_2, ReB_2, TiN, and Ti_2AlB_2 predicted by ML-MD are 40.3, 50.8, 76.2, 36.7, and 16.5 GPa, respectively. These values differ from those obtained via AIMD by maximum 8%. The corresponding ML-MD and AIMD tensile toughness values deviate by maximum 5%. Extrapolation grades during all ML-MD simulations indicate reliable extrapolation ($\gamma \leq 5 < \gamma_{reliable} \approx 10$) and remain of similar magnitude also during ML-MD tensile tests on supercells with $S1$-size (see second bullet point in Section MLIPs' up-fitting for nanoscale tensile tests).

The results in Fig. 9 suggest general applicability of our approach for development of MLIPs able to describe tensile deformation and fracture in hard ceramics. Specifically, our MLIPs correctly reproduce AIMD results of stress/strain curves and fracture mechanisms in different ceramic systems subject to tensile deformation. Extrapolation grades during all ML-MD tensile tests (both atomic and nanoscale) are of similar magnitude as for TiB_2, hence indicating reliable extrapolation and realistic deformation and fracture processes.

Summary and outlook

We proposed a strategy for the development of MLIPs specifically trained to description of deformation and fracture in tensile-loaded ceramic monocrystals. TiB_2 served as a model ceramic system. Training data generation, fitting, and validation procedure were performed within the moment tensor potential (MTP) formalism. MLIP-based molecular dynamics tensile-testing investigations have been carried out from the atomic scale ($\approx 10^3$ atoms) to the nanoscale ($\approx 10^4$–10^6 atoms). Furthermore, we discussed the MLIP's transferability to, e.g., description of TiB_2 subject to different loading conditions or different Ti-B phases, as well as the viability of the here-proposed training strategy for developing MLIPs of other ceramics.

Key findings are summarized below.

MLIP development:

1. MLIPs for simulations of tensile deformation until fracture can be trained following the scheme in Fig. 1, based on snapshots from finite-temperature AIMD simulations of sequentially elongated single-crystal models with sizes of $\approx 10^3$ atoms. An analogous training approach may be applicable to other loading conditions (e.g., shearing) and ceramics in different stoichiometries and crystalline phases (e.g., diborides, nitrides, MAB phases).

2. The applicability of MLIPs to description of tensile deformation and Poisson's contraction at the nanoscale requires up-fitting. This is due to, e.g., nucleation of extended defects being hindered by the small size of AIMD supercells used for generation of training sets. We propose to generate additional ab initio data by room-temperature and elevated-temperature (1200 K) AIMD: imposing a large strain along one lattice vector, initializing atoms at ideal lattice sites, and equilibrating the supercell under fixed volume and shape.

3. MLIPs fitted to room-temperature tensile dataset may be transferable to simulate other loading conditions; here we show examples of high-temperature tensile deformation and room-temperature shear deformation at the atomic scale. Contrarily, up-fitting is certainly required for simulations of volumetric compression, other phases and stoichiometries.

Predictions forTiB_2:

1. Our calculations indicate elastic isotropy of TiB_2's basal plane at 300 and 1200 K.
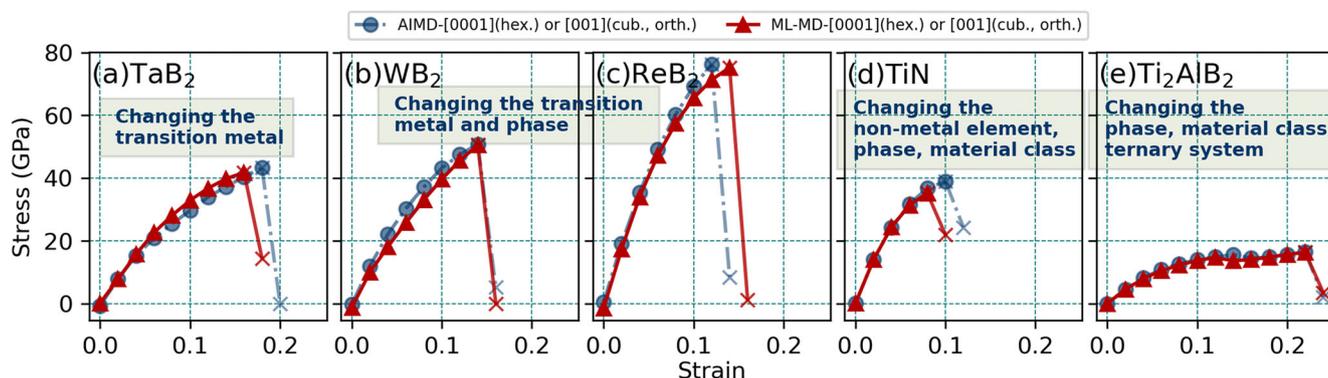
Fig. 9 | Generality of our MLIP training strategy for ceramic materials. a α-TaB_2, **b** ω-WB_2, **c** γ-ReB_2, **d** NaCl-type TiN, and **e** the orthorhombic Ti_2AlB_2 MAB phase, as illustrated by stress/strain curves for room-temperature uniaxial tensile deformation. Specifically, the [0001] and the [001] loading directions are chosen as representative examples for hexagonal systems (TaB_2, WB_2, ReB_2) and for the cubic and the orthorhombic systems (TiN, Ti_2AlB_2), respectively. The supercell sizes and computational setup are equivalent to the atomic scale tensile tests for TiB_2, defined by the first bullet point in Section MLIPs' up-fitting for nanoscale tensile tests. Note that the stress values should not be over-interpreted, as they were obtained for atomic scale supercells (to make a fair comparison with ab initio data) and---in case of negligible size effects---are the ideal upper bounds attainable by a perfect single crystal.

2. The directional dependence of mechanical properties in (initially) dislocation-free supercells qualitatively changes from the atomic to the nanoscale. However, all predicted properties rapidly saturate for supercell sizes increasing from $\approx 10^4$ to $\approx 10^6$ atoms. At 300 K, theoretical tensile strengths during [0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0] deformation reach 51–56 GPa.

3. Nanoscale MD simulations provide insights into crack nucleation and growth mechanisms. Subject to [0001] tensile loading, Ti/B$_2$ layer delamination induces opening of nm-sized voids which rapidly coalesce, inducing formation of few-nm-size cracks inside the material. Fracture surfaces align predominantly with basal planes, {0001}, and first order pyramidal planes, {10$\bar{1}$1}.

4. Considering deformation within the basal plane, [10$\bar{1}$0] tensile test (i.e., loading in the direction of strong B–B bonds), most often induces crack deflection, formation of V-shaped defects, and fracture at {11$\bar{2}$2} family of surfaces. Contrarily, the [$\bar{1}$2$\bar{1}$0] tensile deformation induces fracture at {10$\bar{1}$0} prismatic planes.

The example of TiB$_2$ together with additional ML-MD tensile tests done on other ceramics (TaB$_2$, WB$_2$, ReB$_2$, TiN, Ti$_2$AlB$_2$) indicate the viability of the here-proposed MLIP training strategy. Our approach may be extendable also to other MLIP formalisms. The predictions of nanoscale deformation and fracture in TiB$_2$ presented in this work may aid interpretation of future mechanical-testing experiments. Several previous studies have already demonstrated the importance of MD simulations for elucidating microscopy observations refs. 78–80. Follow-up work could focus on MLIP up-fitting for modelling more complex problems as, e.g., Mode-I loading of native flaws and nanosized cracks.

Methods
Ab initio calculations
Zero Kelvin ab initio calculations as well as finite-temperature Born-Oppenheimer ab initio molecular dynamics (AIMD) were carried out using VASP[81] together with the projector augmented wave (PAW)[82] method and the Perdew-Burke-Ernzerhof exchange-correlation functional revised for solids (PBEsol)[83]. All AIMD calculations employed plane-wave cut-off energies of 300 eV and Γ-point sampling of the reciprocal space.

Supercells. The model of TiB$_2$ was based on the AlB$_2$-type structure (P6/mmm). The 720-atom (240 Ti + 480 B) supercell—used to generate the training/learning/validation dataset—had size of $\approx (1.5 \times 1.6 \times 2.6)$ nm^3, with x, y, z axes chosen to satisfy the following crystallographic relationships: $x \parallel$ [10$\bar{1}$0], $y \parallel$ [$\bar{1}$2$\bar{1}$0], $z \parallel$ [0001]. Similar supercells—with 720 atoms (240 M + 480 B)—were used in Section Viability of our training strategy for modelling tensile deformation in other ceramics for TaB$_2$, WB$_2$, and ReB$_2$, where the latter two are in the ω and γ phase[72], respectively. Cubic (fcc, Fm$\bar{3}$m) TiN[67] was modelled in a 360-atom (180 Ti + 180 N) supercell, with x, y, and z axes aligned with the [100], [010], and [001] directions. The orthorhombic (Cmcm) Ti$_2$AlB$_2$[77] was modelled in a 720-atom supercell (288 Ti + 144 Al + 288 B), oriented in the same way as the TiN supercell.

Equilibration of TiB$_2$ at 300 and 1200 K was performed in 2 steps: (i) 10 ps AIMD isobaric-isothermal (NPT) simulation with Parrinello-Rahman barostat[84] and Langevin thermostat, and (ii) a 2 ps for 300 K, 4 ps for 1200 K AIMD run within the canonical (NVT) ensemble based on Nosé-Hoover thermostat, using time-averaged lattice parameters from the NPT simulation (see the Supplementary Fig. 1). TaB$_2$, WB$_2$, ReB$_2$, TiN, and Ti$_2$AlB$_2$ were equilibrated at 300 K using the same approach.

Computational setup for simulations of TiB$_2$'s [0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0] tensile deformation and (0001)[$\bar{1}$2$\bar{1}$0], (10$\bar{1}$0)[$\bar{1}$2$\bar{1}$0], and (10$\bar{1}$0)[0001] shear deformation (all with Γ point sampling) followed refs. 61,67,85. Specifically, the equilibrated supercell was elongated or sheared in the desired direction using strain increments of 2%. Poisson's contraction effects were not considered in AIMD simulations. The supercells are equilibrated for 3 ps at each deformation step. Stress tensor components were calculated as averages of the

final 0.5 ps. The same approach was used to simulate tensile deformation in TaB$_2$, WB$_2$, ReB$_2$, TiN, and Ti$_2$AlB$_2$.

Room-temperature elastic constants, C_{ij}, of TiB$_2$ were evaluated following ref. 86, based on a second-order polynomial fit of the [0001], [10$\bar{1}$0], and [$\bar{1}$2$\bar{1}$0] stress/strain data ($C_{11}, C_{12}, C_{13}, C_{33}$) and of the (0001)[$\bar{1}2\bar{1}$0], (10$\bar{1}$0)[$\bar{1}2\bar{1}$0], and (10$\bar{1}$0)[0001] shear stress/strain data (C_{44}), considering strains between 0 and 4%.

Simulations of TiB$_2$'s room-temperature volumetric compression were carried out for a 720-atom TiB$_2$ supercell maintained at 300 K (for 2 ps) within the NVT ensemble. The surface energies were calculated at zero Kelvin using a 60-atom TiB$_2$ supercell (with a $3 \times 3 \times 1$ **k**-mesh and cut-off energy of 300 eV) together with a 10 Å vacuum layer. The supercells were fully relaxed until forces on atoms were below 10^{-2} eV Å^{-1} and the total energy was converged with accuracy of 10^{-5} eV per supercell. Other ground-state and higher-energy structures from the Ti–B phase diagram (Ti$_2$B, Ti$_3$B$_4$, TiB, TiB$_{12}$, etc.) were fully relaxed at 0 K starting from lattice parameters and atomic positions from the Materials Project[76].

Development of machine-learning interatomic potentials (MLIPs)
We used the moment tensor potential (MTP) formalism, as implemented in the `mlip-2` package[87]. Training data generation and general workflow are detailed in the Section Training procedure and fitting initial MLIPs and Fig. 1. Training/learning/validation sets included only equilibrated configurations: the initial part (5%) of NVT rus was discarded.

MLIPs were fitted based on the 16g MTPs (referring to the highest degree of polynomial-like basis functions in the analytic description of the MTP[23]), using the Broyden-Fletcher-Goldfarb-Shanno method[88] with 1500 iterations and 1.0, 0.01 and 0.01 weights for total energy, stresses and forces in the loss functional. A cutoff radius of 5.5 Å was employed, similar to other recent MLIP studies[19,89]. Tests using larger cutoffs, 7.4 and 10.0 Å did not show notable changes in accuracy. Expansion of a training set by selection of configurations from a learning set (LS), was done using the `select add` command of the `mlip-2` package. Specifically, all configurations in the LS were ordered by their extrapolation grade (γ[28]) and maximum 15 from the upper 20% was selected to expand the training set.

Details of MLIPs developed in this work (summarized in Fig. 1c) are given below.

- MLIP-[0001], (MLIP-[10$\bar{1}$0], and MLIP-[$\bar{1}$2$\bar{1}$0]): trained on AIMD snapshots of TiB$_2$ subject to room-temperature tensile loading in the [0001] ([10$\bar{1}$0], [$\bar{1}$2$\bar{1}$0]) direction. See Section Training procedure and fitting initial MLIPs and MLIPs' validation against atomic scale tensile tests.
- MLIP-[1]: up-fitting MLIP-[0001], learning from the final TSs of MLIP-[10$\bar{1}$0] and MLIP-[$\bar{1}$2$\bar{1}$0]. See Section MLIPs' up-fitting for nanoscale tensile tests.
- MLIP-[2] and MLIP-[3]: up-fitting MLIP-[1], learning from AIMD snapshots of TiB$_2$ equilibrated at 1200 K (MLIP-[2]), and sequentially elongated in the [0001] direction until cleavage (MLIP-[3]). See Section MLIPs' up-fitting for nanoscale tensile tests.
- MLIP-[4]: up-fitting MLIP-[1], learning from AIMD snapshots of TiB$_2$ elongated by 150% in the [0001] direction, initializing atoms at ideal lattice sites and equilibrating at 300 and 1200 K under fixed volume and shape. See Section MLIPs' up-fitting for nanoscale tensile tests–Otherloading conditions and MLIP's transferability.

Molecular dynamics with MLIPs (ML-MD)
ML-MD calculations were performed with the LAMMPS code[90] interfaced with `mlip-2` package[87], which allows using MTP-type MLIPs (specified in the `pair_style` command). Additionally, the active learning state file (`state.als`, for details see the `mlip-2` documentation (https://gitlab.com/ashapeev/mlip-2-paper-supp-info)) was used to output the extrapolation grade, γ, values during the simulations.

Computational setup of atomic scale ML-MD tensile and shear tests (at 300 or 1200 K) was equivalent to AIMD. Stress tensor components and

elastic constants were calculated in the same way as described above in case of AIMD.

Nanoscale ML-MD tensile tests at 300 or 1200 K used supercells with 12,960 atoms (S1); 141,120 atoms (S2); 230,400 atoms (S3); and 432,000 atoms (S4); with dimensions of about $(1.5 \times 1.6 \times 2.6)$ nm^3, $(4.6 \times 4.7 \times 5.1)$ nm^3, $(10.6 \times 11.0 \times 10.3)$ nm^3, $(12.1 \times 12.6 \times 12.9)$ nm^3, and $(15.2 \times 15.8 \times 15.4)$ nm^3, respectively. Prior to simulating mechanical deformation, the supercells were equilibrated for 5 ps at the target temperature using the isobaric-isothermal (NPT) ensemble coupled to the Nosé-Hoover thermostat with a 1 fs time step. Tensile loading was simulated by deforming the supercell with a constant strain rate (50 Å s^{-1}), accounting for lateral contraction (Poisson's effect) in the NPT ensemble.

Atomic scale volumetric compression simulations used supercell sizes and deformation approach equivalent to what was described above for AIMD. Surface structures and other Ti–B phases were fully relaxed at 0 K using conjugate gradient energy minimization in molecular statics (MS).

Visualization and structural analysis

The OVITO package[91] allowed us to visualize and analyze selected AIMD and ML-MD trajectories. In particular, we looked at (i) *Radial pair distribution functions* (with a cut-off radius of 5.5 Å), (ii) *Elastic strain maps* and (iii) *Atomic strain maps* (with cut-off radius of ± 0.1 Å). For details see the OVITO documentation.

Data availability

The related files are compressed as " Available_Data.zip " and submitted as part of Supplementary Materials, the details description can be found at the end of " Supplementary_Materials.pdf ", as well as " READ_ME.txt " in the compressed folder. The more specific explanation and help will be made available upon request.

References

1. Bianchini, F., Glielmo, A., Kermode, J. & De Vita, A. Enabling QM-accurate simulation of dislocation motion in γ-Ni and α-Fe using a hybrid multiscale approach. *Phys. Rev. Mater.* **3**, 043605 (2019).
2. Zhao, Y. Understanding and design of metallic alloys guided by phase-field simulations. *Npj Comput. Mater.* **9**, 94 (2023).
3. Sivaraman, G. et al. Machine-learned interatomic potentials by active learning: amorphous and liquid hafnium dioxide. *Npj Comput. Mater.* **6**, 104 (2020).
4. Fiedler, L. et al. Predicting electronic structures at any length scale with machine learning. *Npj Comput. Mater.* **9**, 115 (2023).
5. Rassoulinejad-Mousavi, S. M. & Zhang, Y. Interatomic potentials transferability for molecular simulations: a comparative study for platinum, gold and silver. *Sci. Rep.* **8**, 2424 (2018).
6. Bianchini, F., Kermode, J. & De Vita, A. Modelling defects in Ni–Al with EAM and DFT calculations. *Model. Simul. Mat. Sci. Eng.* **24**, 045012 (2016).
7. Proville, L., Rodney, D. & Marinica, M.-C. Quantum effect on thermally activated glide of dislocations. *Nat. Mater.* **11**, 845–849 (2012).
8. Deringer, V. L., Caro, M. A. & Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.* **31**, 1902765 (2019).
9. Zuo, Y. et al. Performance and cost assessment of machine learning interatomic potentials. *J. Phys. Chem. A* **124**, 731–745 (2020).
10. Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104 (2019).
11. Smith, J. S., Isayev, O. & Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **8**, 3192–3203 (2017).
12. Shapeev, A. V., Podryabinkin, E. V., Gubaev, K., Tasnádi, F. & Abrikosov, I. A. Elinvar effect in β-Ti simulated by on-the-fly trained moment tensor potential. *N. J. Phys.* **22**, 113005 (2020).
13. Nishiyama, T., Seko, A. & Tanaka, I. Application of machine learning potentials to predict grain boundary properties in fcc elemental metals. *Phys. Rev. Mater.* **4**, 123607 (2020).
14. Deng, F., Wu, H., He, R., Yang, P. & Zhong, Z. Large-scale atomistic simulation of dislocation core structure in face-centered cubic metal with Deep Potential method. *Comput. Mater. Sci.* **218**, 111941 (2023).
15. Mori, H. & Ozaki, T. Neural network atomic potential to investigate the dislocation dynamics in bcc iron. *Phys. Rev. Mater.* **4**, 040601 (2020).
16. Rosenbrock, C. W. et al. Machine-learned interatomic potentials for alloys and alloy phase diagrams. *Npj Comput. Mater.* **7**, 24 (2021).
17. Zong, H., Pilania, G., Ding, X., Ackland, G. J. & Lookman, T. Developing an interatomic potential for martensitic phase transformations in zirconium by machine learning. *Npj Comput. Mater.* **4**, 48 (2018).
18. Li, X.-G. et al. Quantum-accurate spectral neighbor analysis potential models for Ni-Mo binary alloys and fcc metals. *Phys. Rev. B* **98**, 094104 (2018).
19. Tasnádi, F., Bock, F., Tidholm, J., Shapeev, A. V. & Abrikosov, I. A. Efficient prediction of elastic properties of Ti$_{0.5}$Al$_0$. 5N at elevated temperature using machine learning interatomic potential. *Thin Solid Films* **737**, 138927 (2021).
20. Thompson, A. P., Swiler, L. P., Trott, C. R., Foiles, S. M. & Tucker, G. J. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J. Comput. Phys.* **285**, 316–330 (2015).
21. Behler, J. Constructing high-dimensional neural network potentials: a tutorial review. *Int. J. Quantum Chem.* **115**, 1032–1050 (2015).
22. 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).
23. Shapeev, A. V. Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **14**, 1153–1173 (2016).
24. Seko, A., Togo, A. & Tanaka, I. Group-theoretical high-order rotational invariants for structural representations: Application to linearized machine learning interatomic potential. *Phys. Rev. B* **99**, 214108 (2019).
25. Qamar, M., Mrovec, M., Lysogorskiy, Y., Bochkarev, A. & Drautz, R. Atomic cluster expansion for quantum-accurate large-scale simulations of carbon. *J. Chem. Theory Comput.* **19**, 5151–5167 (2023).
26. Rowe, P., Csányi, G., Alfe, D. & Michaelides, A. Development of a machine learning potential for graphene. *Phys. Rev. B* **97**, 054303 (2018).
27. Cohn, D. A., Ghahramani, Z. & Jordan, M. I. Active learning with statistical models. *J. Artif. Intell.* **4**, 129–145 (1996).
28. Podryabinkin, E. V. & Shapeev, A. V. Active learning of linearly parametrized interatomic potentials. *Comput. Mater. Sci.* **140**, 171–180 (2017).
29. Gubaev, K., Podryabinkin, E. V. & Shapeev, A. V. Machine learning of molecular properties: Locality and active learning. *Tj. Chem. Phys.* **148**, 241727 (2018).
30. Lysogorskiy, Y., Bochkarev, A., Mrovec, M. & Drautz, R. Active learning strategies for atomic cluster expansion models. *Phys. Rev. Mater.* **7**, 043801 (2023).
31. Lysogorskiy, Y. et al. Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon. *npj Comput. Mater.* **7**, 1–12 (2021).

32. Luo, G. A review of automatic selection methods for machine learning algorithms and hyper-parameter values. *Netw. Model. Anal. Health Inform.* **5**, 1–16 (2016).

33. Salem, M., Cowan, M. J. & Mpourmpakis, G. Predicting segregation energy in single atom alloys using physics and machine learning. *ACS omega* **7**, 4471–4481 (2022).

34. Fang, J. et al. Machine learning accelerates the materials discovery. *Mater. Today Commun*. 104900 (2022).

35. Ribeiro, F. & Gradvohl, A. L. S. Machine learning techniques applied to solar flares forecasting. *Astron. Comput.* **35**, 100468 (2021).

36. Magnuson, M., Hultman, L. & Högberg, H. Review of transition-metal diboride thin films. *Vacuum* **196**, 110567 (2022).

37. Holleck, H. Material selection for hard coatings. *J. Vac. Sci. Technol.* **4**, 2661–2669 (1986).

38. Golla, B. R., Mukhopadhyay, A., Basu, B. & Thimmappa, S. K. Review on ultra-high temperature boride ceramics. *Prog. Mater. Sci.* **111**, 100651 (2020).

39. Wang, C., Akbar, S., Chen, W. & Patton, V. Electrical properties of high-temperature oxides, borides, carbides, and nitrides. *J. Mater. Sci.* **30**, 1627–1641 (1995).

40. Sevik, C., Bekaert, J., Petrov, M. & Milošević, M. V. High-temperature multigap superconductivity in two-dimensional metal borides. *Phys. Rev. Mater.* **6**, 024803 (2022).

41. Wiedemann, R., Oettel, H. & Jerenz, M. Structure of deposited and annealed TiB_2 layers. *Surf. Coat. Technol.* **97**, 313–321 (1997).

42. Hofmann, W. & Jäniche, W. Die struktur von aluminiumborid AlB_2. *Z. f.ür. Physikalische Chem.* **31**, 214–222 (1936).

43. Eorgan, J. & Fern, N. Zirconium diboride coatings on tantalum. *JOM* **19**, 6–11 (1967).

44. Norton, J. T., Blumenthal, H. & Sindeband, S. Structure of diborides of titanium, zirconium, columbium, tantalum and vanadium. *JOM* **1**, 749–751 (1949).

45. Mikula, M. et al. Mechanical properties of superhard TiB_2 coatings prepared by DC magnetron sputtering. *Vacuum* **82**, 278–281 (2007).

46. Geng, J. et al. Microstructural and mechanical anisotropy of extruded in-situ TiB_2/2024 composite plate. *Mater. Sci. Eng.* **687**, 131–140 (2017).

47. Zhang, T. F., Gan, B., Park, S.-m, Wang, Q. M. & Kim, K. H. Influence of negative bias voltage and deposition temperature on microstructure and properties of superhard TiB_2 coatings deposited by high power impulse magnetron sputtering. *Surf. Coat. Technol.* **253**, 115–122 (2014).

48. Fuger, C. et al. Revisiting the origins of super-hardness in TiB_{2+z} thin films–impact of growth conditions and anisotropy. *Surf. Coat. Technol.* **446**, 128806 (2022).

49. Munro, R. G. Material properties of titanium diboride. *J. Res. Natl Inst. Stan.* **105**, 709 (2000).

50. Chen, Z. et al. Room-temperature deformation of single crystals of ZrB_2 and TiB_2 with the hexagonal AlB_2 structure investigated by micropillar compression. *Sci. Rep.* **11**, 14265 (2021).

51. Zhang, T. et al. Preparation of highly-dense TiB_2 ceramic with excellent mechanical properties by spark plasma sintering using hexagonal TiB_2 plates. *Mater. Res. Express* **6**, 125055 (2019).

52. Zhou, Y., Wang, J., Li, Z., Zhan, X. & Wang, J. First-principles investigation on the chemical bonding and intrinsic elastic properties of transition metal diborides TMB_2 (TM = Zr, Hf, Nb, Ta, and Y). *Ultra-High Temperature Ceramics: Materials for Extreme Environment Applications* 60–82 (2014).

53. Dai, F.-Z. & Zhou, Y. Effects of transition metal (TM = Zr, Hf, Nb, Ta, Mo, W) elements on the shear properties of TMB_2s: A first-principles investigation. *Comput. Mater. Sci.* **117**, 266–269 (2016).

54. Zhang, X., Luo, X., Li, J., Hu, P. & Han, J. The ideal strength of transition metal diborides TMB_2 (TM = Ti, Zr, Hf): plastic anisotropy and the role of prismatic slip. *Scr. Mater.* **62**, 625–628 (2010).

55. Attarian, S. & Xiao, S. Development of a 2NN - MEAM potential for TiB system and studies of the temperature dependence of the nanohardness of TiB_2. *Comput. Mater. Sci.* **201**, 110875 (2022).

56. Attarian, S. *Multiscale modeling of Ti/TiB composites*. Ph.D. thesis, The University of Iowa (2021).

57. Timalsina, B. *Development of Eam and Rf-MEAM Interatomic Potential for Zirconium Diboride*. Ph.D. thesis, Missouri State University (2021).

58. Daw, M. S., Lawson, J. W. & Bauschlicher Jr, C. W. Interatomic potentials for zirconium diboride and hafnium diboride. *Comput. Mater. Sci.* **50**, 2828–2835 (2011).

59. Chicco, D., Warrens, M. J. & Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **7**, e623 (2021).

60. Podryabinkin, E., Garifullin, K., Shapeev, A. & Novikov, I. MLIP-3: Active learning on atomic environments with moment tensor potentials. *J. Chem. Phys.* **159**, 084112 (2023).

61. Koutná, N. et al. Atomistic mechanisms underlying plasticity and crack growth in ceramics: a case study of AlN/TiN superlattices. *Acta Mater.* **229**, 117809 (2022).

62. Nakano, K., Imura, T. & Takeuchi, S. Hardness anisotropy of single crystals of IVa-diborides. *Jpn. J. Appl. Phys.* **12**, 186 (1973).

63. Hodapp, M. & Shapeev, A. In operando active learning of interatomic interaction during large-scale simulations. *Mach. Learn. Sci. technol.* **1**, 045005 (2020).

64. Freitas, R. & Cao, Y. Machine-learning potentials for crystal defects. *MRS Commun.* **12**, 510–520 (2022).

65. Paul, B. et al. Plastic deformation of single crystals of CrB_2, TiB_2 and ZrB_2 with the hexagonal AlB_2 structure. *Acta Mater.* **211**, 116857 (2021).

66. Waldl, H. et al. Evolution of the fracture properties of arc evaporated $Ti_{1-x}Al_xN$ coatings with increasing Al content. *Surf. Coat. Technol.* **444**, 128690 (2022).

67. Sangiovanni, D., Tasnádi, F., Johnson, L., Odén, M. & Abrikosov, I. Strength, transformation toughening, and fracture dynamics of rocksalt-structure $Ti_{1-x}Al_x$N (0 x 0.75) alloys. *Phys. Rev. Mater.* **4**, 033605 (2020).

68. Sangiovanni, D. G. et al. Descriptor for slip-induced crack blunting in refractory ceramics. *Phys. Rev. Mater.* **7**, 103601 (2023).

69. Jiao, Y., Huang, L. & Geng, L. Progress on discontinuously reinforced titanium matrix composites. *J. Alloy. Compd.* **767**, 1196–1215 (2018).

70. Dub, S. et al. Mechanical properties of single crystals of transition metals diborides TMB_2 (TM= Sc, Hf, Zr, Ti). experiment and theory. *J. Superhard Mater.* **39**, 308–318 (2017).

71. Lei, J. et al. Synthesis and high-pressure mechanical properties of superhard rhenium/tungsten diboride nanocrystals. *ACS nano* **13**, 10036–10048 (2019).

72. Leiner, T. et al. On energetics of allotrope transformations in transition-metal diborides via plane-by-plane shearing. *Vacuum* **215**, 112329 (2023).

73. Sun, W., Dai, F., Xiang, H., Liu, J. & Zhou, Y. General trends in surface stability and oxygen adsorption behavior of transition metal diborides (TMB_2). *J. Mater. Sci. Technol.* **35**, 584–590 (2019).

74. Gan, Q. et al. Robust hydrophobic materials by surface modification in transition-metal diborides. *ACS Appl. Mater. Interfaces* **13**, 58162–58169 (2021).

75. Murray, J., Liao, P. & Spear, K. The B-Ti (boron-titanium) system. *Bull. Alloy Phase Diagr.* **7**, 550–555 (1986).

76. Jain, A. et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).

77. Carlsson, A., Rosen, J. & Dahlqvist, M. Theoretical predictions of phase stability for orthorhombic and hexagonal ternary MAB phases. *Phys. Chem. Chem. Phys.* **24**, 11249–11258 (2022).

78. Koutná, N. et al. Atomistic mechanisms underlying plasticity and crack growth in ceramics: a case study of AlN/TiN superlattices. *Acta Mater.* **229**, 117809 (2022).

79. Salamania, J. et al. Elucidating dislocation core structures in titanium nitride through high-resolution imaging and atomistic simulations. *Mater. Des.* **224**, 111327 (2022).

80. Chen, Z. et al. Atomic insights on intermixing of nanoscale nitride multilayer triggered by nanoindentation. *Acta Mater.* **214**, 117004 (2021).

81. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).

82. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).

83. Perdew, J. P. et al. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).

84. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

85. Sangiovanni, D., Mellor, W., Harrington, T., Kaufmann, K. & Vecchio, K. Enhancing plasticity in high-entropy refractory ceramics via tailoring valence electron concentration. *Mater. Des.* **209**, 109932 (2021).

86. Sangiovanni, D. G. et al. Temperature-dependent elastic properties of binary and multicomponent high-entropy refractory carbides. *Mater. Des.* **204**, 109634 (2021).

87. Novikov, I. S., Gubaev, K., Podryabinkin, E. V. & Shapeev, A. V. The MLIP package: moment tensor potentials with MPI and active learning. *Mach. learn.: sci. technol.* **2**, 025002 (2020).

88. Fletcher, R.*Practical methods of optimization* (John Wiley & Sons, 2013).

89. Erhard, L. C., Rohrer, J., Albe, K. & Deringer, V. L. A machine-learned interatomic potential for silica and its relation to empirical models. *Npj Comput. Mater.* **8**, 90 (2022).

90. Thompson, A. P. et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.* **271**, 108171 (2022).

91. Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO–the open visualization tool. *Model. Simul. Mat. Sci. Eng.* **18**, 015012 (2009).

92. Chen, L. et al. A facile one-step route to nanocrystalline TiB_2 powders. *Mater. Res. Bull.* **39**, 609–613 (2004).

93. Mukaida, M., Goto, T. & Hirai, T. Preferred orientation of TiB_2 plates prepared by CVD of the $TiCl_{4+}$ B_2 H_6 system. *J. Mater. Sci.* **26**, 6613–6617 (1991).

94. Kelesoglu, E. & Mitterer, C. Structure and properties of TiB_2 based coatings prepared by unbalanced DC magnetron sputtering. *Surf. Coat. Technol.* **98**, 1483–1489 (1998).

95. Xiang, H., Feng, Z., Li, Z. & Zhou, Y. Temperature-dependence of structural and mechanical properties of TiB_2: A first principle investigation. *J. Appl. Phys.***117** (2015).

96. Spoor, P. et al. Elastic constants and crystal anisotropy of titanium diboride. *Appl. Phys. Lett.* **70**, 1959–1961 (1997).

97. Amulele, G. M. & Manghnani, M. H. Compression studies of TiB_2 using synchrotron X-ray diffraction and ultrasonic techniques. *J. Appl. Phys.* **97**, 023506 (2005).

98. Guan, C. & Zhu, H. Theoretical insights into the behaviors of sodium and aluminum on the cathode titanium diboride surfaces. *Comput. Mater. Sci.* **211**, 111535 (2022).

99. Clayton, J. et al. Deformation and failure mechanics of boron carbide–titanium diboride composites at multiple scales. *JOM* **71**, 2567–2575 (2019).

100. Fan, H. & El-Awady, J. A. Molecular dynamics simulations of orientation effects during tension, compression, and bending deformations of magnesium nanocrystals. *J. Appl. Mech.* **82**, 101006 (2015).

Acknowledgements

Author contributions

S.L. carried out the simulations, analyzed and visualized the data, and wrote the first manuscript draft; L.C.T. and F.T. advised on the methodology, data analysis, and visualization; L.H. and P.H.M. helped to improve the manuscript and to discuss the results with respect to state-of-the art experiments, they also provided financial support; D.G.S. and N.K. designed the project and shaped the manuscript; N.K. provided day-to-day supervision to S.L. All authors commented on the manuscript and helped to revise it.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-024-01252-3.

Correspondence and requests for materials should be addressed to Shuyao Lin.

Reprints and permissions information is available at http://www.nature.com/reprints