






This work was carried out in whole or in part within the framework of the NOMATEN Center of Excellence, supported from the European Union Horizon 2020 research and innovation programme (Grant Agreement No. 857470) and from the European Regional Development Fund via the Foundation for Polish Science International Research Agenda PLUS programme (Grant No. MAB PLUS/2018/8).

This is a copy of the publication which appeared in: *Modelling and Simulation in Materials Science and Engineering*, vol 29 (2021) 055001 (15pp), published on: 12 May 2021.

DOI: [10.1088/1361-651X/abf152](https://doi.org/10.1088/1361-651X/abf152)

Computational study of crystal defect formation in Mo by a machine learning molecular dynamics potential

F J Domínguez-Gutiérrez^{1,4,*}, J Byggmästar²,
K Nordlund², F Djurabekova^{2,3} and U von Toussaint¹

¹ Max-Planck Institute for Plasma Physics, Boltzmannstrasse 2, 85748 Garching, Germany

² Department of Physics, University of Helsinki, Helsinki, PO Box 43, FIN-00014, Finland

³ Helsinki Institute of Physics, Helsinki, Finland

E-mail: javier.dominguez@ipp.mpg.de

Received 3 November 2020, revised 21 January 2021

Accepted for publication 24 March 2021

Published 12 May 2021



CrossMark

Abstract

In this work, we study the damage in crystalline molybdenum material samples due to neutron bombardment in a primary knock-on atom (PKA) range of 0.5–10 keV at room temperature. We perform classical molecular dynamics (MD) simulations using a previously derived machine learning (ML) interatomic potential based on the Gaussian approximation potential (GAP) framework. We utilize a recently developed software workflow for fingerprinting and visualizing defects in damaged crystal structures to analyze the Mo samples with respect to the formation of point defects during and after a collision cascade. As a benchmark, we report results for the total number of Frenkel pairs (a self-interstitial atom and a single vacancy) formed and atom displacements as a function of the PKA energy. A comparison to results obtained using an embedded atom method (EAM) potential is presented to discuss the advantages and limits of the MD simulations utilizing ML-based potentials. The formation of Frenkel pairs follows a sublinear scaling law as ξ^b where b is a fitting parameter and $\xi = E_{\text{PKA}}/E_0$ with E_0 as a scaling factor. We found that the $b = 0.54$ for the

⁴Present address: NOMATEN Centre of Excellence, National Center for Nuclear Research, ul. A. Sołtana 7, 05-400 Swierk/Otwock, Poland.

* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

GAP MD results and $b = 0.667$ for the EAM simulations. Although the average number of total defects is similar for both methods, the MD results show different atomic geometries for complex point defects, where the formation of crowdions by the GAP potential is closer to the DFT-based expectation. Finally, ion beam mixing results for GAP MD simulations are in a good agreement with experimental mixing efficiency data. This indicates that the modeling of atom relocation in cascades by machine learned potentials is suited to interpret the corresponding experimental findings.

Keywords: molybdenum, MD simulations, ion beam mixing, materials modeling, machine learning methods

(Some figures may appear in colour only in the online journal)

1. Introduction

The design of next generation of fusion machines needs experimental exploration of different plasma facing materials (PFM) candidates and the support and validation of numerical modeling [1, 2]. Molybdenum has been selected as a candidate for a PFM due to its high melting point, good resistance to deformation, and low sputtering yield under plasma irradiation [3]. Mo is also used for diagnostic mirrors in fusion machines to deal with the harsh plasma environment and fluxes of neutrals and neutron radiation [1, 4]. In order to guide PFM experiments and to understand the mechanism of the interaction of plasma and neutrons with materials, atomistic simulations based on the molecular dynamics (MD) method can be performed [5–7]. They can also serve to save laboratory and financial resources for carrying out the PFM experiments.

Transition metals and alloys are traditionally modeled by the embedded atom method (EAM) potentials in MD simulations [5, 8–10], reproducing many material properties in good agreement with those measured experimentally. However, EAM and other traditional potentials are limited to fixed functional forms [1, 8] and can wrongly model some point defects that are energetically unstable, or lack physical meaning in material damaging processes [11]. For this reason, interatomic potentials developed by using machine learning (ML) methods are now increasingly used to perform MD simulations with an accuracy close to density functional theory (DFT) [12–14]. These have the advantage that they can be systematically improved toward the accuracy of the DFT training data set. The goal of the present work is to numerically model the damage in crystalline Mo samples due to irradiation in a fusion reactor by utilizing the recently developed ML interatomic potential by Byggmästar *et al* [15]. Hence, we perform MD simulations to emulate neutron bombardment at intermediate primary knock-on atom (PKA) energies, providing an understanding about the modeling of the re-crystallization process after the collision cascade, which has been an issue for numerical simulations based on fixed functional forms [11, 12].

Our paper is organized as follows: in section 2 we briefly discuss the theory to develop the machine learned (ML) potential [15] for Mo, and the software workflow for fingerprinting and visualizing defects in damaged crystal structures (FaVAD) [16, 17] that is applied to quantify and classify the damage in Mo samples [18]. Our results for the total number of points defects, Frenkel pairs and atomic displacement are presented in section 3. We examine the limitations and advantages of our new ML interatomic potential by comparing to MD simulations results obtained by EAM potentials, as well as the ion beam mixing comparison between the machine learned MD simulation results and those reported experimentally. Finally, in section 4, we provide concluding remarks.

2. Methods

Interatomic potentials based on ML methods are not restricted to an analytical form and can be systematically improved toward the accuracy of the training data set. In order to model collision cascades, the ML potential must be able to treat realistic short-range dynamics defined by its repulsive part. In addition, the correct structure of the liquid phase and re-crystallization process should be well described, to accurately emulate atomic mixing together with defect creation and annihilation during the collision cascade. In this work, we use the ML interatomic potential for molybdenum that was recently developed [15] within the Gaussian approximation potential (GAP) framework [13, 19]. Here, the total energy of a system of N atoms is expressed as

$$E_{\text{tot}} = \sum_{i < j}^N V_{\text{pair}}(r_{ij}) + \sum_i^{N_d} E_{\text{GAP}}^i, \quad (1)$$

where V_{pair} is a purely repulsive screened Coulomb potential, and E_{GAP} is the ML contribution. E_{GAP} is constructed using a two-body and the many-body smooth overlap of atomic positions (SOAP) descriptor [13]. N_d is the number of descriptor environments for the N -atom system (i.e. number of pairs for the two-body descriptor and number of atoms for the many-body descriptor). The ML part of the potential is given by

$$E_{\text{GAP}}^i = \delta_{2b}^2 \sum_j^{M_{2b}} \alpha_{j,2b} K_{2b}(\vec{q}_{i,2b}, \vec{q}_{j,2b}) + \delta_{mb}^2 \sum_j^{M_{mb}} \alpha_{j,mb} K_{mb}(\vec{q}_{i,mb}, \vec{q}_{j,mb}), \quad (2)$$

where $\delta_{2b,mb}^2$ are prefactors that set the energy ranges of the ML predictions; $K_{2,mb}$ is the kernel function representing the similarity between the atomic environment of the i th and j th atoms; α is a coefficient obtained from the fitting process; and \vec{q} is the normalized descriptor vector (DV) of the local atomic environment of the i th atom (see section 2.2). In the computation of the ML potential the descriptors for two bodies, 2b, are utilized to take into account most of the interatomic bond energies, while the many-body, mb, contributions are treated by the SOAP descriptor. More details about the computation of the ML potentials for Mo can be found in reference [15].

2.1. MD simulations

ML based MD simulations are performed to model neutron bombardment processes at different PKA energy values to analyze damage in crystalline materials.

We first define a simulation box as a pristine Mo crystalline sample based on a body-centered-cubic (bcc) unit cell with a lattice constant of $a = 3.16 \text{ \AA}$ according to DFT calculations for computing the GAP potentials and those reported in the literature [20]. This value is slightly higher than the experimental measurement [21]. In table 1 we list the size of the numerical boxes used for each PKA value. Then, the numerical sample is prepared by a process of energy optimization and thermalization to 300 K using the Langevin thermostat, with the time constant of 100 fs [22]. The room temperature is used in our work to perform numerical simulations as close as possible to the experiments of material damaging [1]. The MD simulation is started by assigning a kinetic energy (KE) to a Mo atom located at the center of the numerical sample in a range of 0.5–10 keV of PKA. For each PKA energy value, the projectile travels on ten different crystal orientation: $\langle 001 \rangle$, $\langle 110 \rangle$, $\langle 111 \rangle$, and 7 cases for $\langle r_1 r_2 r_3 \rangle$, where r_i are random numbers uniformly distributed in an interval of $[0, 1]$. The velocity Verlet integration

Table 1. Size of the numerical boxes in nm as a function of the impact energy (PKA energy), which are used in the MD simulations. The lattice constant of the bcc Mo sample for GAP is $a = 3.168$ Å at 300 K.

| PKA (keV) | Num. atoms | Box size (nm) |
|-----------|------------|-----------------------|
| 0.5–2 | 25 392 | (7.24, 7.24, 7.55) |
| 5 | 55 800 | (9.51, 9.51, 9.82) |
| 10 | 104 044 | (11.72, 11.72, 12.04) |

Table 2. Physical properties of molybdenum (cohesion energy, E_{coh} ; melting temperature, T_{melt} ; SIA and vacancy formation energies, E_{f} ; and vacancy migration energy, $E_{\text{mig}}^{\text{vac}}$) obtained by GAP [15] and EAM [10], as reported in the literature [15], and their comparison to experimental measurements [25].

| | EAM | GAP | Expt. |
|---|---------------|---------------|-----------|
| E_{coh} (eV atom ⁻¹) | -6.82 | -6.288 | -6.821 |
| T_{melt} (K) | 3080 ± 20 | 2750 ± 10 | 2895 |
| $E_{\text{f}}^{(111)}$ (eV) | 7.19 | 7.56 | — |
| $E_{\text{f}}^{(110)}$ (eV) | 6.95 | 7.61 | — |
| $E_{\text{f}}^{(100)}$ (eV) | 7.18 | 8.99 | — |
| $E_{\text{f}}^{\text{octa}}$ (eV) | 7.56 | 9.00 | — |
| $E_{\text{f}}^{\text{tetra}}$ (eV) | 7.35 | 8.44 | — |
| $E_{\text{f}}^{\text{vac}}$ (eV) | 2.55 | 2.84 | 3.0–3.24 |
| $E_{\text{mig}}^{\text{vac}}$ (eV) | 1.28 | 1.28 | 1.35–1.62 |

algorithm is utilized to model the collision cascade, which is performed for 6 ps, followed by an additional relaxation time of 4 ps. An electronic stopping correction has been included in our MD simulations due to the high PKA energy range considered in this work. The electronic stopping powers S_e [23] were obtained from SRIM-2013 [24] using the default values for the molybdenum.

To explore the advantages and limitations of the new GAP interatomic potential, we compare results to those obtained by the EAM [9, 10] for the same simulation conditions and a lattice constant of $a = 3.147$ Å. This MD potential is denoted as EAM in this work and has previously been applied to study the sputtering of single-crystalline Mo surfaces by Mo_n ($n = 1, 2, 4$) projectiles in the total energy range of 0.125–4 keV. In table 2, we compare physical properties of Mo obtained by GAP and EAM as a benchmark for our MD simulations.

The MD simulations were done in the High Performance Computing Center of the Max Planck Institute and the Institutional Cluster of the Stony Brook University by using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [26]. The Quantum Mechanics and Interatomic Potential package [27] is used as an interface to implement machine learned interatomic potentials based on GAP [13]. In order to ensure accurate time integration in the high-energy collision dynamics, we use an adaptive timestep that is implemented in LAMMPS, with a maximum (t_{max}) of 4 fs and ensuring that the maximum displacement of the Mo atoms per step is less than 0.01 Å.

2.2. Identification of point defects and vacancies

The damage in the Mo sample is analyzed by a software workflow for fingerprinting and visualizing defects in damaged crystal structures (FaVAD) [16, 17], where the local atomic

environment of the i th atom of the material sample is described by a DV, $\vec{\xi}^i$. Here, the atom density around the i th atom is expressed by a sum of a truncated Gaussian density functions with the difference vector \vec{r}^{ij} between the atoms i and j , entering the exponent [19],

$$\rho^i(\vec{r}) = \sum_j^{\text{neigh.}} \exp\left(-\frac{|\vec{r} - \vec{r}^{ij}|^2}{2\sigma_{\text{atom}}^2}\right) f_{\text{cut}}(|\vec{r}^{ij}|), \quad (3)$$

$$= \sum_{nlm}^{NLM} c_{nlm}^{(i)} g_n(r) Y_{lm}(\hat{r}), \quad (4)$$

which is then approximated in terms of spherical harmonic functions, $Y_{lm}(\hat{r})$, and a set of basis functions in radial directions $g_n(r)$ as $c_{nlm}^{(i)} = \langle g_n Y_{lm} | \rho^i \rangle$ [18, 19]. The sum over the order m of the squared modulus of the coefficients c_{nlm} is invariant under rotations around the central atom [28]. It is given by

$$\vec{\xi}_k^i = \left\{ \sum_m (c_{nlm}^i)^* c_{n'l'm}^i \right\}_{n,n',l}, \quad (5)$$

where c_{nlm}^{*} denotes the complex conjugate of c_{nlm} . Here each component k of the vector $\vec{\xi}$ corresponds to one of the index triplets $\{n, n', l\}$. The normalized DV $\vec{q}^i = \vec{\xi}^i / |\vec{\xi}^i|$ is used throughout this work and calculated within the multi-body descriptor framework called ‘SOAP’, which implements equations (3) to (5) with the GAP [13].

Once the DVs of all the atoms of the damaged material are computed, we calculate the distance between the two corresponding DVs, $d = d(\vec{q}^i, \vec{q}^j)$ to the atomic local environment of a defect free and thermalized material sample [16, 18] as

$$d^M(T) = \sqrt{(\vec{q}^i - \vec{v}(T))^T \Sigma^{-1}(T) (\vec{q}^i - \vec{v}(T))}, \quad (6)$$

where $\vec{v}(T) = \frac{1}{N} \sum_{i=1}^N \vec{q}^i(T)$ is the mean DV of the defect free sample; and Σ is the associated co-variance matrix of the DV components [18, 29]. This calculation allows us to identify atoms outside of lattice positions that are then classified as point defects. The identification of vacancies is done by defining a numerical sampling grid of $N = N_x \times N_y \times N_z$ points with N_x, N_y, N_z being the number of equispaced points in the x, y , and z directions, respectively [18]. This is followed by a computation of the nearest neighbor distance between the position of the damaged sample atoms and the sampling grid points. Points where the distance to the nearest atom exceeds a given threshold describe the spatial volume of the identified vacancy [16, 17].

3. Results and discussion

In order to analyze the damage in the Mo sample, we first perform a single MD simulation at 500 eV on the $\langle 001 \rangle$ velocity direction, with the GAP and EAM potentials for 10 ps of simulation time. The final frame of the simulation contains the information of the point defects formed at the lowest PKA energy value of our study. Then, we analyze the damaged Mo sample by computing the DVs of all the Mo atoms with FaVAD, followed by the comparison to a defect free sample that is thermalized to 300 K.

In figure 1 we present results for the distance difference, $d^M(300 \text{ K})$, between the DVs of the damaged Mo materials and the set of reference DVs obtained by using FaVAD. The MD simulation performed with GAP presents two Mo atoms with the largest distance difference,

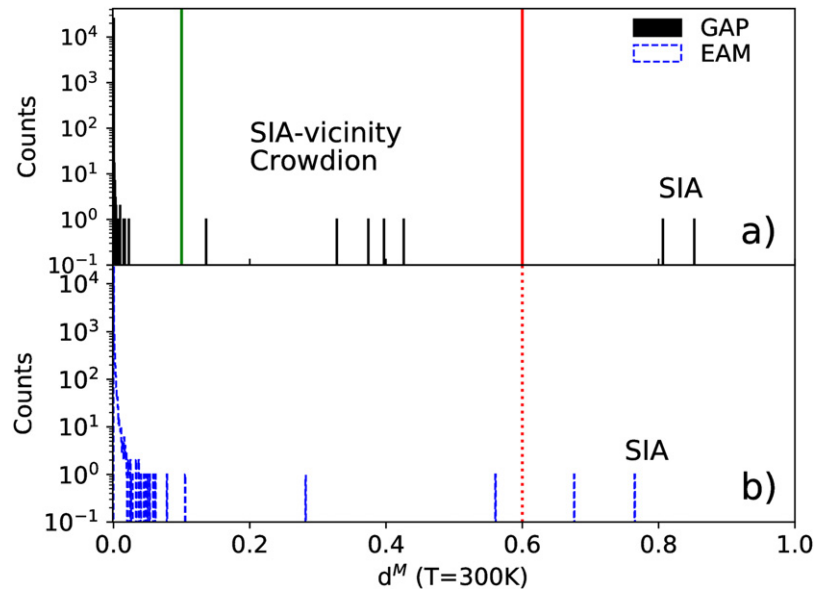


Figure 1. Distance difference between the DVs for the damaged Mo sample and those for the defect free and thermalized sample for the GAP in (a) and EAM in (b) potentials. We apply FaVAD to a bombarded sample at a PKA energy of 500 eV on the $\langle 001 \rangle$ velocity direction. We notice that GAP potentials modeled the formation of 2 crowdions, while the EAM potentials model the presence of only one crowdion defect after the collision cascade.

that allows us to set a threshold at 0.6 for further analyses. A Mo atom with distance difference bigger than a value of 0.6 is quantified as a self-interstitial atom (SIA) for all the MD simulations. We also identify Mo atoms in the vicinity of the SIA with a distance difference in the range of 0.1 to 0.6. Although these atoms are not quantified as SIAs, they provide information about the atomic arrangement of more complex point defects like a crowdion in this case, where four atoms share three lattice sites. For the EAM case, two Mo atoms are also identified as potential SIA by FaVAD. However, the atomic geometry of this point defects is associated to a dumbbell defect where two atoms share a lattice site, observed in the figure 1(b) where only a couple of Mo atoms have a distance difference in the range of 0.2 to 0.6. The GAP and EAM MD simulations report the same number of SIA, but the atomic geometry of the modeled defects is different. Besides that, the re-crystallization of the Mo sample is well modeled by the GAP potentials, where the majority of the Mo atoms have an atomic local environment similar to the the defect-free and thermalized Mo sample, which is noted in the upper panel of the same figure.

3.1. Crystal defects formation as a function of the simulation time

The interatomic MD potentials need to be capable to model the formation of point defects at different PKA energies assigned to the projectile. More complex defects can be found for high PKA energy values. For this reason, we increase the features of FaVAD to compute the DVs of all the atoms of the sample at different time steps of MD simulations. Point defects can be identified by comparing their DVs to those of the defect-free sample thermalized to 300 K. In order to show this new feature, we perform MD simulations at PKA energies of 10 keV and

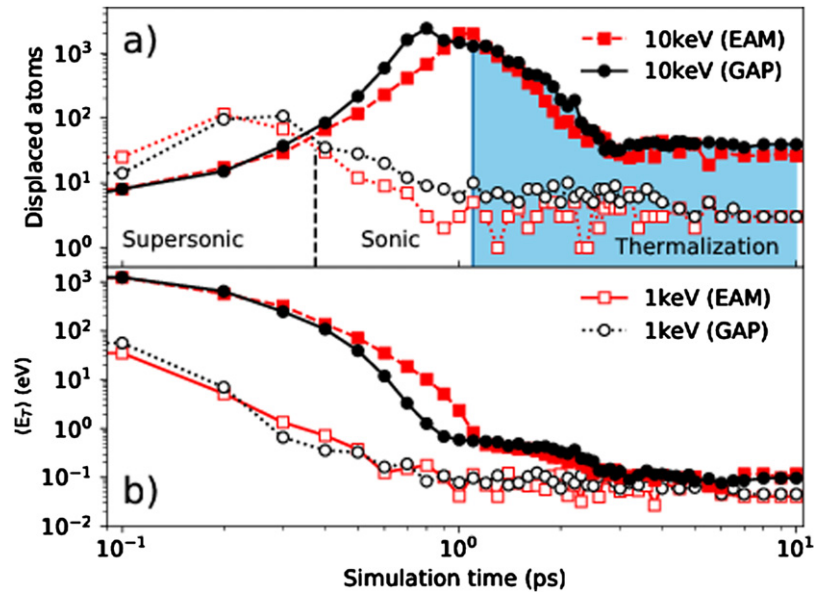


Figure 2. (a) Number of displaced atoms as a function of the simulation time for 1 and 10 keV PKA's in the $\langle 001 \rangle$ velocity direction, by using the GAP and EAM potentials. At 10 keV, three phases are identified during collision dynamics associated with the shockwave velocity where the destructive phase shows particles traveling at supersonic velocities. Then the re-crystallization of the Mo sample occurs during the thermalization phase. (b) The average of the KE of displaced Mo atoms. We notice a similitude between the two MD potentials.

1 keV on the $\langle 001 \rangle$ velocity direction. The information of the damage of the material is obtained as an output data from the MD simulation at a time step of $\Delta t_d = 0.1$ ps for 1–3 ps, where the collision cascade mechanisms mainly happen [30]; a $\Delta t_d = 0.5$ ps for 3–6 ps; and a final $\Delta t_d = 1.0$ ps for relaxing the damaged sample.

In figure 2(a), we present results for the quantification of point defects formed as a function of the simulation time at 1 (empty symbols) and 10 (solid symbols) keV of PKA with the GAP and EAM potentials. FaVAD is applied to identify and quantify the point defects with a distance difference threshold of $d^M = 0.6$ for all cases, for which value is observed that Mo atoms have the highest probability to be considered as actual defects [11]. Although the profiles presented by the MD simulations at 10 keV are similar, the maximum number of displaced atoms is located at 0.8 and 1.1 ps for the GAP and EAM respectively. At the end of the MD simulations, a total of 42 displaced atoms is reported for the GAP modeling and 35 Mo atoms are identified as defects by utilizing the EAM potential. The ML based MD potential produces a majority of crowdion defects rather than dumbbells, as will be discussed later. At a PKA of 1 keV, the dynamics presented by the displacement atoms is notable different at 1–7 ps where the re-crystallization of the material sample is carried out. Being well modeled by the GAP framework (figure 2(a)). However, both MD potentials report the same number of three stable point defects at the end of the MD simulation.

In figure 2(a), three phases of a collision cascade are identified for cascades at a PKA energy of 10 keV by computing the average velocity, $\langle V(t) \rangle$, of the total identified displaced atoms,

N_D , as a function of the simulation time as

$$\langle V(t) \rangle = \frac{1}{N_D} \sum_{i=1}^{N_D} \sqrt{v_x^i(t)^2 + v_y^i(t)^2 + v_z^i(t)^2}, \quad (7)$$

where v_x , v_y , and v_z are the instantaneous velocities of the i th displaced atom taken from the output data of the MD simulations. By considering the longitudinal velocity of sound ν_{Mo} in Mo that varies from 5.4 to 6.25 km s⁻¹ at a temperature of 300 K [31], we can define a supersonic collision cascade phase where $\langle V(t) \rangle > \nu_{Mo}$. This stage is closely related to the ballistic phase [6]. Thus, at the beginning of the MD simulation (0.1–0.4 ps) the average velocity of the displaced atoms is 1.5 to 3 times bigger than ν_{Mo} , which is a similar phase range reported for iron samples [32] and Fe–Ni alloys [30]. In this supersonic phase, highly energetic atoms start colliding and transferring KE to their nearest neighbor atoms. At this lapse of time, atoms with supersonic velocities create stable point defects in the sample, then some of them can be identified as Frenkel pairs, for example. More Mo atoms are displaced from their lattice position and the starting KE of the projectile is distributed among the Mo atoms, followed by loss of KE by the Mo atoms that were already displaced. This process leads the Mo sample to the sonic phase in the time lapse of 0.4 to 1.1 ps. The sonic wave does not create stable point defects, leading to a liquid phase inside the Mo sample which is well modeled by the GAP potential due to the inclusion of liquid samples at different pressures in the training data [15]. For collisions at 10 keV, we noticed that the number of displaced atoms at the limit between the supersonic and sonic phases is proportional to the total number of SIA at the end of the simulation time, as expected and reported in the literature for damage in materials [30]. The last phase of the collision simulation is called the thermalization phase and defined when the average velocity of the displaced atoms is $\langle V(t) \rangle < \nu_{Mo}$. In this phase the material sample is re-crystallized and cools down to its initial room temperature.

In figure 2(b) we report results for the average KE of displaced Mo atom as a function of the simulation time for PKA energy values of 1 and 10 keV with the GAP and EAM potentials. The KE of the i th displaced Mo atom is calculated as $E_K^i = (m/2) [v_x^i(t)^2 + v_y^i(t)^2 + v_z^i(t)^2]$ with m as the Mo mass. Followed by the computation of the average KE of all the displaced Mo atoms as $\langle E \rangle = (1/N_D) \sum_i^{N_D} E_K^i$. We notice that at the beginning of the MD simulation the projectile transfers its KE to its nearest neighbor atoms which starts the supersonic shock-wave. After 1.1 ps of simulation time, the average KE is almost constant, which represents the thermalization process of the sample for all the MD simulations at both PKA energy values. However, EAM and GAP model the expansion of the sonic wave and the energy landscape of defects differently, which leads to the formation of different atomic geometries for the identified SIAs and Mo atoms in their respective vicinity at the end of the MD simulation. The formation of dumbbell SIA defects is more common in the EAM potential.

3.2. Classification and quantification of crystal defects as a function of the PKA energy

We perform MD simulations in a PKA energy range of 0.5 to 10 keV for several velocity directions with the GAP and EAM potentials. The comparison of the modeling of the formation of point defects after collision cascade modeled by these potentials provide an insight of the advantages and limits of the GAP MD potentials over traditional ones (EAM potentials). In figures 3 and 4 we present the average number of point defects, $\langle PD(E_p) \rangle$, and its standard

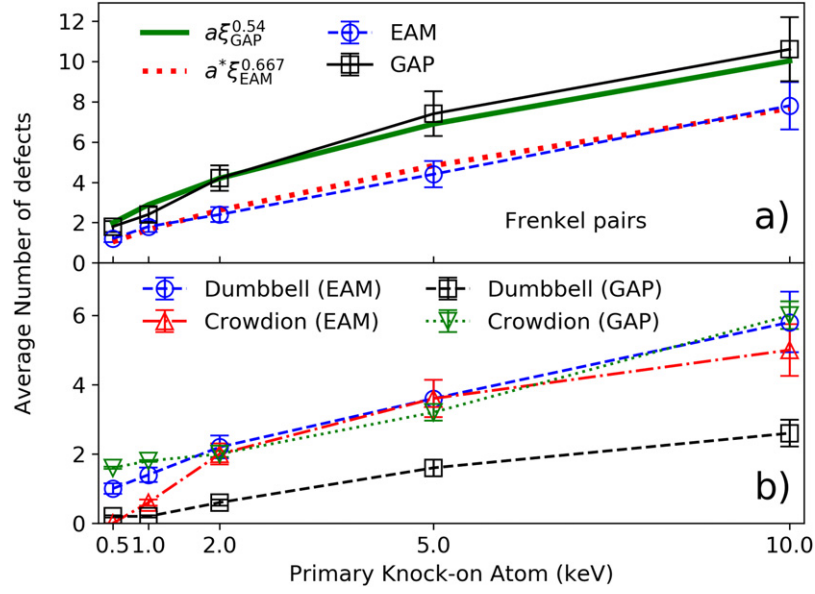


Figure 3. Average number of point defects formed after collision cascade as a function of the PKA energy. Total number of Frenkel pairs in (a), and crowdions and dumbbells in (b). Data points are connected by lines for visual guidance. We compare results obtained by GAP with those for the EAM MD potentials. A fitting curve is included to the total number of Frenkel pairs as: $a\xi^b$ with $\xi = E_{\text{PKA}}/E_0$ where $E_0 = 1$ keV, $a = 2.88$ and $b = 0.54$ for the GAP (green solid line) and $a^* = 1.65$ and $b = 0.667$ for the EAM (dotted red line), with a correlation factor of 0.99 for both methods. The number of Frenkel pair increases roughly with the square root of E_{PKA} .

deviation, $\sigma(E_p)$ as a function of the PKA, E_p , which are calculated as:

$$\langle \text{PD}(E_p) \rangle = \frac{1}{N_T} \sum_{i=1}^{N_T} N_i(E_p | \langle r_1 r_2 r_3 \rangle)$$

$$\sigma(E_p) = \sqrt{\frac{1}{N_T - 1} \sum_{i=1}^{N_T} (N_i(E_p | \langle r_x r_y r_z \rangle) - \langle \text{PD}(E_p) \rangle)^2}, \quad (8)$$

with $N_i(E_p | \langle r_1 r_2 r_3 \rangle)$ as the number of defects for a given velocity direction, $\langle r_1 r_2 r_3 \rangle$, and N_T as the total number of MD simulations performed. The defects are identified by FaVAD as a function of the PKA energy.

Figure 3(a) shows the average number of Frenkel pairs as a function of the PKA energy, modeled by the GAP and EAM potentials. This value is related to the average number of single vacancies found in the damaged Mo sample. FaVAD has to identify the formation of a stable SIA and a vacancy to quantify this kind of defect. Usually Mo atoms with a $d^M(T) > 0.6$ are identified by FaVAD as SIA and quantified by using equation (8), and Mo atoms in their vicinity indicate the formation of a crowdion or dumbbell defects. The average number of Frenkel pairs can be fitted to a scaling law proposed by Stoller *et al* [33] that quantifies the number of point defects formed in damaged samples as: $a\xi^b$, where $\xi = E_{\text{PKA}}/E_0$ with $E_0 = 1$ keV as a scaling factor, and a and b are fitting parameters. In our case, a good approximation to the mean value

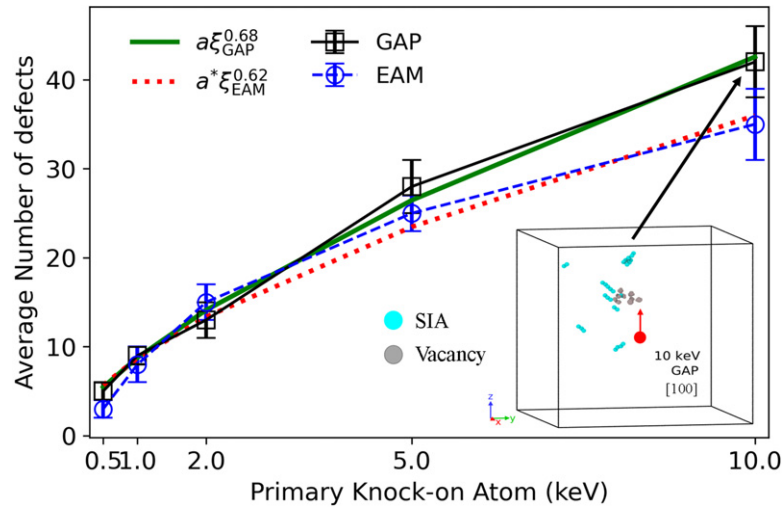


Figure 4. Total number of defects calculated as: Frenkel pairs + $3 \times$ crowdion + dumbbell + SIA vicinity are presented as a function of the PKA. A fitting curve is included to the total number of defects as: $a\xi^b$ with $\xi = E_{\text{PKA}}/E_0$ and $E_0 = 1$ keV; for GAP $a = 8.77$ and $b = 0.68$, while for EAM $a = 8.71$ and $b = 0.62$. A correlation factor of 0.99 is associated to the fitting curve for both methods. In the inset, we show the identified SIA and Mo atoms in the local vicinity (cyan spheres), as well as vacancies (gray spheres) at the final step of the MD simulation at 10 keV PKA with GAP, showing the formation of crowdion defects. The initial position of the projectile and its [100] velocity direction are shown by a red sphere and a red arrow, respectively.

of the GAP data is found by applying the damped least-square method implemented in Python library SciPy with fitting parameters $a = 2.88$ and $b = 0.54$, where the associated correlation factor is 0.99. We also fit the mean value of the EAM results with $a = 101.65$ and $b = 0.667$ and a correlation factor of 0.99.

In figure 3(b) we show results for the average number of dumbbells and crowdions as a function of the PKA energy. Here, the number of crowdions are quantified by identifying four Mo atoms sharing three lattice positions, while dumbbells are detected when two Mo atoms share one lattice position by FaVaD [11], with a $d^M(T) > 0.2$. We notice that the formation of crowdions is more favorable for collision cascades simulated by the GAP potential, while dumbbells are formed for all the PKA energy values with the EAM potential. This is not a surprise, since the EAM potential incorrectly predicts $\langle 110 \rangle$ dumbbells to be lower in energy than $\langle 111 \rangle$ crowdions or dumbbells [9] (see table 2). The GAP describes the relative formation energies of all SIA defects in a good agreement with DFT results [15], and correctly reproduces the $\langle 11\xi \rangle$ dumbbell as the most stable SIA [34]. The $\langle 11\xi \rangle$ dumbbell is a tilted $\langle 111 \rangle$ SIA that easily migrates along the $\langle 111 \rangle$ chain of atoms in a zigzag-like motion [34], fluctuating between dumbbell and crowdion configurations, both of which FaVaD detects and identifies.

In table 3, we present the number of crystal defects as a function of the PKA energy for GAP and EAM MD simulations, as a reference. Frenkel pairs (SIA + single vacancy) are counted for Mo atoms with a maximum DV distance difference. Single vacancies are quantified and identified by the KD-tree algorithm included in FaVaD. Identified crowdions and dumbbells are also tabulated. The total number of defects is calculated as: Frenkel pairs + $3 \times$ crowdion + dumbbell + SIA vicinity and are shown in figure 4. There is a fair agreement between GAP

Table 3. Average number of point defects and vacancies as a function of the PKA, which are identified by our DV based method. SIA are identified as W atoms with the largest $d^M(T)$ to be in an interstitial site, reported into parentheses. SIA vicinity are affected Mo atoms due to the presence of an SIA and thermal motion. Standard deviation values are included in the data.

| Defect | 0.5 | 1 | 2 | 5 | 10 |
|---------------------|--------------|--------------|---------------|---------------|---------------|
| GAP potential | | | | | |
| PKA (keV) | | | | | |
| Frenkel pairs | 2 ± 0.33 | 3 ± 0.45 | 4 ± 0.6 | 7 ± 1.05 | 10 ± 1.53 |
| Crowdion | 1 ± 0.15 | 2 ± 0.33 | 2 ± 0.33 | 3 ± 0.45 | 6 ± 0.9 |
| Dumbbell | 0 | 0 | 1 ± 0.15 | 2 ± 0.33 | 3 ± 0.45 |
| SIA vicinity | 0 | 0 | 2 ± 0.33 | 10 ± 1.53 | 11 ± 1.67 |
| Total | 5 ± 0.20 | 9 ± 0.54 | 13 ± 0.77 | 28 ± 1.94 | 42 ± 2.55 |
| AT-EAM-FS potential | | | | | |
| Frenkel pairs | 1 ± 0.15 | 2 ± 0.33 | 3 ± 0.45 | 4 ± 0.6 | 8 ± 1.33 |
| Crowdion | 0 | 1 ± 0.15 | 2 ± 0.33 | 4 ± 0.6 | 5 ± 0.77 |
| Dumbbell | 1 ± 0.15 | 2 ± 0.33 | 2 ± 0.33 | 4 ± 0.6 | 6 ± 0.9 |
| SIA vicinity | 1 ± 0.15 | 1 ± 0.15 | 4 ± 0.6 | 5 ± 0.77 | 6 ± 0.9 |
| Total | 3 ± 0.26 | 8 ± 0.51 | 15 ± 0.88 | 25 ± 1.29 | 35 ± 2.03 |

and EAM MD results at PKA energies lower than 5 keV. A power-law regression curve aE_{PKA}^b (energy in keV) can be fitted to the total number of defects with $a = 8.77$ and $b = 0.68$ for GAP and $a = 8.71$ and $b = 0.62$ for EAM, resulting in a correlation factor of 0.99 for both MD potentials. The geometry of the identified point defects for the GAP MD simulations is shown in the inset of figure 4, where SIAs and crowdions are represented by spheres in cyan and vacancies are depicted as gray spheres.

3.3. Ion beam mixing

We also analyzed the ion beam mixing from the cascades in Mo, as a means to compare the results with experiments. Ion beam mixing means the effect where the ion irradiation induces ballistic and thermal spike displacements of atoms from their initial sites into other positions, in effect mixing atom positions [5, 6]. The ion beam mixing in bulk material can be measured using marker layer experiments [35, 36]. The ion beam mixing coefficient obtained from the experiments can be compared directly with atom relocation obtained from MD cascade calculations via diffusion theory [37–39].

In the current case, we followed the approach introduced in reference [39] that allows comparing marker layer mixing experiments with simulations for a specific ion beam irradiation. For mixing in Mo, there are experimental data available for several different marker layers materials mixed by a 300 keV Kr ion beam [36]. To compare with these, we first determined the atom relocation as the square of the total atom displacements $R_{\text{sim}}^2(E_p) = \sum_i [r'_i - r_i^0]^2$ in the MD cascade simulations for self-recoils between energies of 500 eV and 10 keV, where the index i runs for all the Mo atom in the samples; and \vec{r}' and \vec{r}^0 are the positions of each atom in the Mo sample at the final and first steps of the MD simulation, respectively. We verified that

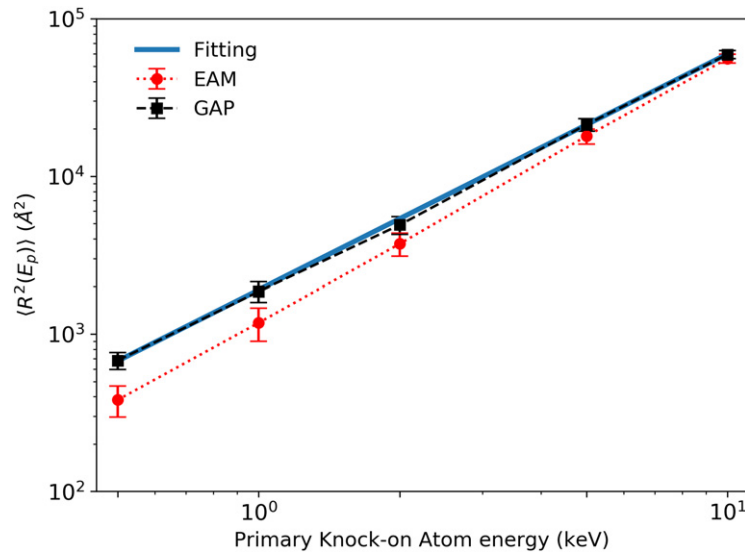


Figure 5. The average of atom displacement $\langle R^2_{\text{sim}}(E_p) \rangle$ over all the projectile velocity directions as a function of cascade energy, E_p . Also shown is a functional fit to the GAP data points as $R^2(E_p) = a \frac{E_p^{3/2}}{E_{\text{sub}}^{1/2} + E_p^{1/2}}$ with $a = 9.5 \text{ \AA}^2 \text{ eV}^{-1}$ is the fitting parameter, see text for details.

the center of mass of the cell is not displaced at the end of the MD simulations. This enables integrating the R^2 data, we then fit the physically motivated [5, 39] functional form

$$R^2(E_p) = a \frac{E_p^{3/2}}{E_{\text{sub}}^{1/2} + E_p^{1/2}}, \quad (9)$$

to the data. Here a is a fitting constant and E_{sub} the subcascade threshold energy. In the current case, we estimated from SRIM simulations [24] that $E_{\text{sub}} \approx 10 \text{ keV}$ for Mo. Using this value the fit to the MD simulations gave $a = 9.5 \text{ \AA}^2 \text{ eV}^{-1}$. In figure 5 we report the average over all the projectile velocity directions of the MD results, $\langle R^2_{\text{sim}}(E_p) \rangle$ as a function of the PKA energy.

Our MD calculations are compared to experimental measurements of mixing efficiency Q_{exp} [36, 40]. The mixing for a 300 keV Kr beam, used in experiments of Kim *et al* [36], can then be calculated by first obtaining the primary recoil spectrum $n(E)dE$ from ion range calculations with the MDRANGE code [41], followed by the computation of the total mixing efficiency caused by the 300 keV Kr beam in Mo as [39]

$$Q_{\text{sim}} = \frac{\int_0^{E_0} R^2(E) n(E) dE}{6n_0 E_{D_n}}, \quad (10)$$

where E_0 is the initial implantation energy (300 keV in the current case), the atomic density $n_0 = 3/a_0^3$ with $a_0 = 3.16 \text{ \AA}$ as the lattice constant of Mo, and E_{D_n} the deposited nuclear energy of the Kr beam around the experimental marker layer depth (400 \AA for the experiments in reference [36]).

Using this procedure, we obtained the value of $Q_{\text{sim}} = 14 \pm 2 \text{ \AA}^5 \text{ eV}^{-1}$ which compares well with the experimental values that are in the range of $Q_{\text{exp}} = 4\text{--}13 \text{ \AA}^5 \text{ eV}^{-1}$. Although there is considerable uncertainty in both simulation and experiments, nevertheless this comparison shows that the GAP potential for Mo describes atom relocation in cascades reasonably well compared to experiments.

4. Concluding remarks

In this paper, we performed classical and machine learning molecular dynamics simulations to emulate neutron bombardment on molybdenum samples in an impact energy range of 0.5–10 keV, and a sample temperature of 300 K. For this, we use a new ML interatomic potential based on the GAP framework. The results are compared to those obtained by using a traditional EAM potential. Formation of Frenkel pairs and more complex defects like crowdions and dumbbells are identified and quantified by using the recently developed software workflow for fingerprinting and visualizing defects in damaged crystal structures (FaVAD). Here, the local environment of each atom of the sample is represented by a DV. The difference between a pristine Mo sample and a damaged one is computed taking into account the magnitude of the thermal motion also.

Supersonic, sonic, and thermalization phases are identified by analyzing the average KE of the identified ‘defective’ Mo atoms. The information of the liquid phase included in the fitting of the GAP potential leads to better modeling of the transition between supersonic and sonic phases, where complex defects start to form. The formation of crowdions is more favorable for the GAP potential. The number of Frenkel pairs formed by using the EAM and GAP MD potentials for collision cascades simulations can be fitted to a $\sim \xi^b$ scaling law, with $\xi = E_{\text{PKA}}/E_0$ where E_{PKA} is the PKA energy and $E_0 = 1 \text{ keV}$. Here $b = 0.54$ for the GAP and $b = 0.667$ for the EAM. Finally, the damage in Mo sample is also analyzed by calculating the ion beam mixing for the GAP results with a value of $Q_{\text{sim}} = 14 \pm 2 \text{ \AA}^5 \text{ eV}^{-1}$ which is in good agreement to experimental measurements of mixing efficiency for a Kr beam with a range of $Q_{\text{exp}} = 4\text{--}13 \text{ \AA}^5 \text{ eV}^{-1}$. This demonstrates the improvement of the numerical modeling of atom relocation in cascades by the GAP potential.

ML interatomic potentials have been developed for further transitional metals like V, Nb, and Ta. With the expected improved accuracy of these ML-based MD potentials subsequent refined investigations of damage processes in these materials are foreseen.

Acknowledgments

FJDG gratefully acknowledges funding from A von Humboldt Foundation and C F von Siemens Foundation for research fellowship. Simulations were performed using the SeaWulf cluster at the Stony Brook University. KN, FD and JB acknowledge that their part of this work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom Research and Training programme 2014–2018 under Grant Agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

F J Domínguez-Gutiérrez  <https://orcid.org/0000-0002-1429-0083>

J Byggmästar  <https://orcid.org/0000-0002-4898-6150>

F Djurabekova  <https://orcid.org/0000-0002-5828-200X>

References

- [1] Rudakov D L *et al* 2006 *Rev. Sci. Instrum.* **77** 10F126
- [2] Wirth B D, Hu X, Kohnert A and Xu D 2015 *J. Mater. Res.* **30** 1440
- [3] Eren B, Marot L, Langer M, Steiner R, Wisse M, Mathys D and Meyer E 2011 *Nucl. Fusion* **51** 103025
- [4] Litnovsky A *et al* 2007 *J. Nucl. Mater.* **363–365** 1395–402
- [5] Nordlund K *et al* 2018 *Nat. Commun.* **9** 1084
- [6] Nordlund K *et al* 2018 *J. Nucl. Mater.* **512** 450–79
- [7] Bolt H, Barabash V, Federici G, Linke J, Loarte A, Roth J and Sato K 2002 *J. Nucl. Mater.* **307–311** 43–52
- [8] Daw M S and Baskes M I 1984 *Phys. Rev. B* **29** 6443–53
- [9] Ackland G J and Thetford R 1987 *Phil. Mag. A* **56** 15–30
- [10] Salonen E, Järvi T, Nordlund K and Keinonen J 2003 *J. Phys.: Condens. Matter* **15** 5845–55
- [11] Domínguez-Gutiérrez F J, Byggmästar J, Nordlund K, Djurabekova F and von Toussaint U 2020 *Nucl. Mater. Energy* **22** 100724
- [12] Byggmästar J, Hamedani A, Nordlund K and Djurabekova F 2019 *Phys. Rev. B* **100** 144105
- [13] Bartók A P, Payne M C, Kondor R and Csányi G 2010 *Phys. Rev. Lett.* **104** 136403
- [14] Hamedani A, Byggmästar J, Djurabekova F, Alahyarizadeh G, Ghaderi R, Minuchehr A and Nordlund K 2020 *Mater. Res. Lett.* **8** 364–72
- [15] Byggmästar J, Nordlund K and Djurabekova F 2020 *Phys. Rev. Mater.* **4** 093802
- [16] von Toussaint U, Domínguez-Gutiérrez F, Compostella M and Rampp M 2020 *Comput. Phys. Commun.* **262** 107816
- [17] von Toussaint Udo 2020 FaVAD: A software workflow for characterization and visualizing of defects in crystalline structures <https://gitlab.mpcdf.mpg.de/NMPP/favad/-/tree/master/>
- [18] Domínguez-Gutiérrez F and von Toussaint U 2019 *J. Nucl. Mater.* **528** 151833
- [19] Bartók A P, Kondor R and Csányi G 2013 *Phys. Rev. B* **87** 184115
- [20] Park H, Fellingner M R, Lenosky T J, Tipton W W, Trinkle D R, Rudin S P, Woodward C, Wilkins J W and Hennig R G 2012 *Phys. Rev. B* **85** 214121
- [21] Pearson W B 1967 *Handbook of Lattice Spacing and Structures of Metals* (Oxford: Pergamon)
- [22] Domínguez-Gutiérrez F J and Krstić P S 2017 *J. Nucl. Mater.* **492** 56–61
- [23] Ziegler J F, Biersack J P and Littmark U 1985 *The Stopping and Range of Ions in Matter* (Oxford: Pergamon)
- [24] Ziegler J F 2013 SRIM-2013 software package <http://srim.org>
- [25] Rumble J 2019 *CRC Handbook of Chemistry and Physics* (Boca Raton, FL: CRC Press)
- [26] Plimpton S 1995 *J. Comput. Phys.* **117** 1–19
- [27] Bernstein Noam, Csanyi Gabor and Kermode James 2018 QUIP - QUantum mechanics and Interatomic Potentials <http://libatoms.github.io/QUIP/>
- [28] Szlachta W J, Bartók A P and Csányi G 2014 *Phys. Rev. B* **90** 104108
- [29] Mahalanobis P 1930 *J. Proc. Asiat. Soc. Bengal* **26** 541
- [30] Béland L K, Osetsky Y N and Stoller R E 2016 *Acta Mater.* **116** 136–42
- [31] Dickinson J M and Armstrong P E 1967 *J. Appl. Phys.* **38** 602–6
- [32] Calder A F, Bacon D J, Barashev A V and Osetsky Y N 2010 *Phil. Mag.* **90** 863–84
- [33] Stoller R E, Odette G R and Wirth B D 1997 *J. Nucl. Mater.* **251** 49–60
- [34] Ma P W and Dudarev S L 2019 *Phys. Rev. Mater.* **3** 043606
- [35] Paine B M and Averback R S 1985 *Nucl. Instrum. Methods Phys. Res. B* **7–8** 666
- [36] Kim S-J, Nicolet M-A, Averback R S and Peak D 1988 *Phys. Rev. B* **37** 38
- [37] Diaz de la Rubia T, Averback R S, Benedek R and King W E 1987 *Phys. Rev. Lett.* **59** 1930–3

Diaz de la Rubia T, Averback R S, Benedek R and King W E 1988 *Phys. Rev. Lett.* **60** 76 (erratum)

- [38] Nordlund K, Ghaly M and Averback R S 1998 *J. Appl. Phys.* **83** 1238–46
- [39] Nordlund K, Wei L, Zhong Y and Averback R S 1998 *Phys. Rev. B* **57** 13965–8
- [40] Björkas C and Nordlund K 2009 *Nucl. Instrum. Methods Phys. Res.* **267** 1830–6
- [41] Nordlund K 1995 *Comput. Mater. Sci.* **3** 448–56