Database-driven semigrand canonical Monte Carlo method: Application to segregation isotherm on defects in alloys

Rodrigo Pinheiro Campos,* Shuhei Shinzato, and Akio Ishii
Department of Mechanical Science and Bioengineering, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

Shuichi Nakamura
Steel Research Laboratories, Nippon Steel Corporation, 20-1 Shintomi, Futtsu, Chiba 293-8511, Japan

Shigenobu Ogata†
Department of Mechanical Science and Bioengineering, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan
and Elements Strategy Initiative for Structural Materials, Kyoto University, Kyoto 606-8501, Japan

(Received 5 February 2021; accepted 30 July 2021; published 25 August 2021)

The application of existing semigrand canonical ensemble Monte Carlo algorithms to alloys requires the chemical potential difference values between pairs of atomic species in the alloys as inputs. However, finding the appropriate values for a target system at a desired temperature and bulk composition is a time-consuming task consisting of multiple test runs to determine the chemical potential differences. This problem becomes more serious when dealing with systems containing three or more atomic species, such as medium- and high-entropy alloys, due to the increase of the number of chemical potential differences that need to be calculated. Here we propose a method for sampling from the semigrand canonical ensemble that relies on energy databases acting as an external atomic reservoir at the desired temperature and composition. Given these energy databases, the desired bulk composition and corresponding chemical potential differences can be satisfied in a "single" Monte Carlo simulation. Moreover, the energy databases shed light on the underlying energetics of alloys, reflecting their local chemical ordering. We demonstrate the validity of this method using analyses of segregation isotherms at grain boundaries and dislocations in two alloy systems: Fe–1-at.-%-Si and NiCoCr medium-entropy alloy. We also discuss the possibly relevant information contained in such energy databases.

DOI: 10.1103/PhysRevE.104.025310

I. INTRODUCTION

Markov chain Monte Carlo (MCMC) atomistic simulations are often used to calculate thermodynamic quantities in metallic systems [1]. In the study of segregation and distribution of atomic species in substitutional metallic alloys, there are widely employed MCMC algorithms to sample thermodynamic quantities while preserving the lattice structure, such as "on-lattice" canonical ensemble Monte Carlo and semigrand canonical ensemble Monte Carlo (SGCMC) algorithms [1–5]. These algorithms have been extensively used to determine phase transformations and equilibrium properties of alloys [1,6–12]. Although these algorithms are general in purpose, they are the most commonly used approach to study the thermodynamic properties of alloy systems containing defects [13–16]. In this paper, we apply such algorithms to analyze the segregation and distribution of atoms and thermodynamic quantities for metallic alloy systems of known chemical composition in which a structural defect is present. Under this approach, the target system exchanges particles with a heat and atomic reservoir that constrains the temperature and chemical potential differences between pairs of atomic species while keeping the volume and total number of atoms in the target system constant, constituting the semigrand canonical (SGC) ensemble [1,2,10]. An SGCMC simulation requires as input the differences in chemical potential between pairs of atomic species; these differences are used to compute the acceptance probability of a particle identity change trial move, as is discussed in the Methods section. Therefore, before starting an SGCMC simulation of an alloy system with a given chemical composition, these input values must be precalculated. However, in situations where the target system has multiple components, this task proves to be quite cumbersome, requiring many trial-and-error SGCMC runs to minimize the error between the desired chemical composition and the chemical composition computed from the SGCMC simulation. Adjustment of the trial chemical potential differences between alloy components is also necessary after each run. For instance, in high-entropy alloys with five elements, four chemical potential differences must be calculated by continuously adjusting each value until the equilibrium composition matches the desired bulk composition. As a consequence of this cumbersomeness, SGCMC algorithms have rarely been applied to medium- or high-entropy alloys [11,12].

*rdgpcampos@tsme.me.es.osaka-u.ac.jp
†ogata@me.es.osaka-u.ac.jp

In this paper, we describe three possible approaches to SGCMC simulations: (1) The traditional calculation of chemical potential differences through a series of trial-and-error SGCMC simulations with adjustment of the chemical potential difference values (Sec. II A); (2) evaluation of the chemical potential differences using energy data obtained from particle exchanges in an NVT simulation of a pair of replicas of the given material at a given temperature (Sec. II B 1); and (3) construction of a set of energy probability distributions of the particle exchanges that occur during the simulation in approach 2, noting that these energy probability distributions can replace chemical potential differences in an SGCMC simulation to realize the SGC ensemble (Secs. II B 2 and II B 3).

Approaches 2 and 3 bypass the trial-and-error calculations required by the traditional SGCMC approach 1. The energy sampling required by approaches 2 and 3 can be done with a "single" MC simulation, after which the SGCMC simulation on the target system can be executed. Moreover, by following approach 3 and building the energy probability distributions, hereinafter referred to as "energy databases," we can shed light on the underlying energetics of alloys, reflecting their local chemical ordering. We applied approaches 2 and 3 to a Fe–1-at.-%-Si alloy system (Sec. III A) and a NiCoCr medium-entropy alloy (MEA) system (Sec. III B), each with defects, such as a grain boundary (GB) or a stacking fault, to demonstrate the methods' general applicability. We then analyzed the temperature-dependent segregation behavior at the defect sites. These analyses confirmed that the proposed approach 3 using energy databases reproduces the results of conventional SGCMC, which uses chemical potential differences as inputs. In each situation, we also made inferences on the chemical interactions in each system based on the profile of the energy databases to showcase the advantages of approach 3 compared to approaches 1 and 2.

II. METHODS

A. Semigrand canonical Monte Carlo

The semigrand canonical ensemble describes the thermodynamics of a target system interacting with a large atomic and heat reservoir, under the restriction that the total number of particles inside the target system is constant. In the constant-volume SGCMC algorithm, the trial moves taken at each MC step traditionally attempt particle displacements and particle identity changes [1,3–5,8,10–16].

Consider an n-nary system at constant temperature T and volume V, where the number of particles of identity m is denoted by N_m and its corresponding chemical potential is denoted by μ_m. Additionally, the total number of particles in the system is denoted by N; the number of particles of each identity is represented by $\boldsymbol{N} = \sum_m N_m \boldsymbol{e}_m \in \mathbb{R}^n$, where the set of $\boldsymbol{e}_m$ vectors forms the standard basis of $\mathbb{R}^n$; and the coordinates of the system's particles are expressed by the vector $\boldsymbol{r}^N \in \mathbb{R}^{3N}$. The particle displacement trial moves sample microstates from the configuration phase space of the system by attempting to randomly displace the positions of the particles ($\boldsymbol{r}^N \rightarrow \boldsymbol{r}'^N$), while the particle identity change trial moves sample microstates from the composition phase space of the system by attempting to change the atomic species of a randomly chosen particle of identity i to a new identity j selected with uniform probability from $\{1, \ldots, n\}$ ($\boldsymbol{N} \rightarrow \boldsymbol{N}'$). The acceptance probability of these trial moves is determined by Eq. (1) [1]:

$$p^{\mathrm{Acc}} = \min\{1, \exp\{-\beta[\Delta U - (\mu_j - \mu_i)]\}\}, \qquad (1)$$

where ΔU denotes the energy change of the system due to either a particle identity change trial move ($\boldsymbol{N} \rightarrow \boldsymbol{N}'$) or a particle displacement trial move ($\boldsymbol{r}^N \rightarrow \boldsymbol{r}'^N$), and β denotes the inverse temperature $[(k_B T)^{-1}]$. Note that, for a particle displacement trial move, the chemical potential difference factor $\mu_j - \mu_i$ is equal to zero, considering that there is no particle identity change: $j = i$. Since we need the differences between chemical potentials to calculate the particle identity change acceptance probability, we can arbitrarily set one of the chemical potential values to zero, for example, $\mu_1 = 0$, and use as inputs to the simulation a set of $n - 1$ values $\{\mu_2, \ldots, \mu_n\}$. If these inputs are not known beforehand, multiple trial-and-error SGCMC simulations would be necessary in order to find the appropriate chemical potential difference values that minimize the error in the composition of the target system. This task can be challenging for systems that take a long time to reach equilibrium and/or are highly multicomponent, since more chemical potential difference values need to be calculated.

It is worth noting that, instead of the particle displacement trial moves in the MC steps, molecular dynamics (MD) steps subject to a Nosé-Hoover thermostat can be used, which ensures faster convergence [10].

B. Database-driven semigrand canonical Monte Carlo simulation

1. Direct chemical potential difference estimation by energy change sampling

To avoid the multiple trial-and-error SGCMC simulations in the conventional method, we propose an approach to evaluate the chemical potential differences by performing an MC simulation of a pair of replicas of the material in consideration. This approach is inspired by the Gibbs ensemble method, which was introduced by Panagiotopoulos to analyze phase coexistence in fluids [17]. In the original implementation of the Gibbs ensemble method, we have an NVT simulation consisting of two interacting systems, with each system being described by a generalized ensemble (or Gibbs ensemble). Normally, the total volume is kept constant throughout the simulation, while the volume of each system can vary. If, instead, the volume of each system is kept constant, they each become denoted by the grand canonical (μVT) ensemble [17].

We devised an algorithm that samples the energy change values from a pair of replicas of the bulk system at constant volume [see Fig. 1(a)]. Note that, similar to the Gibbs ensemble approach, despite the pair of replicas constituting the NVT ensemble, each replica individually can move along the composition phase space by exchanging particles with its counterpart. However, here, both the volume and the aggregate number of particles in each replica are constrained, which results in each replica being denoted by the SGC ensemble

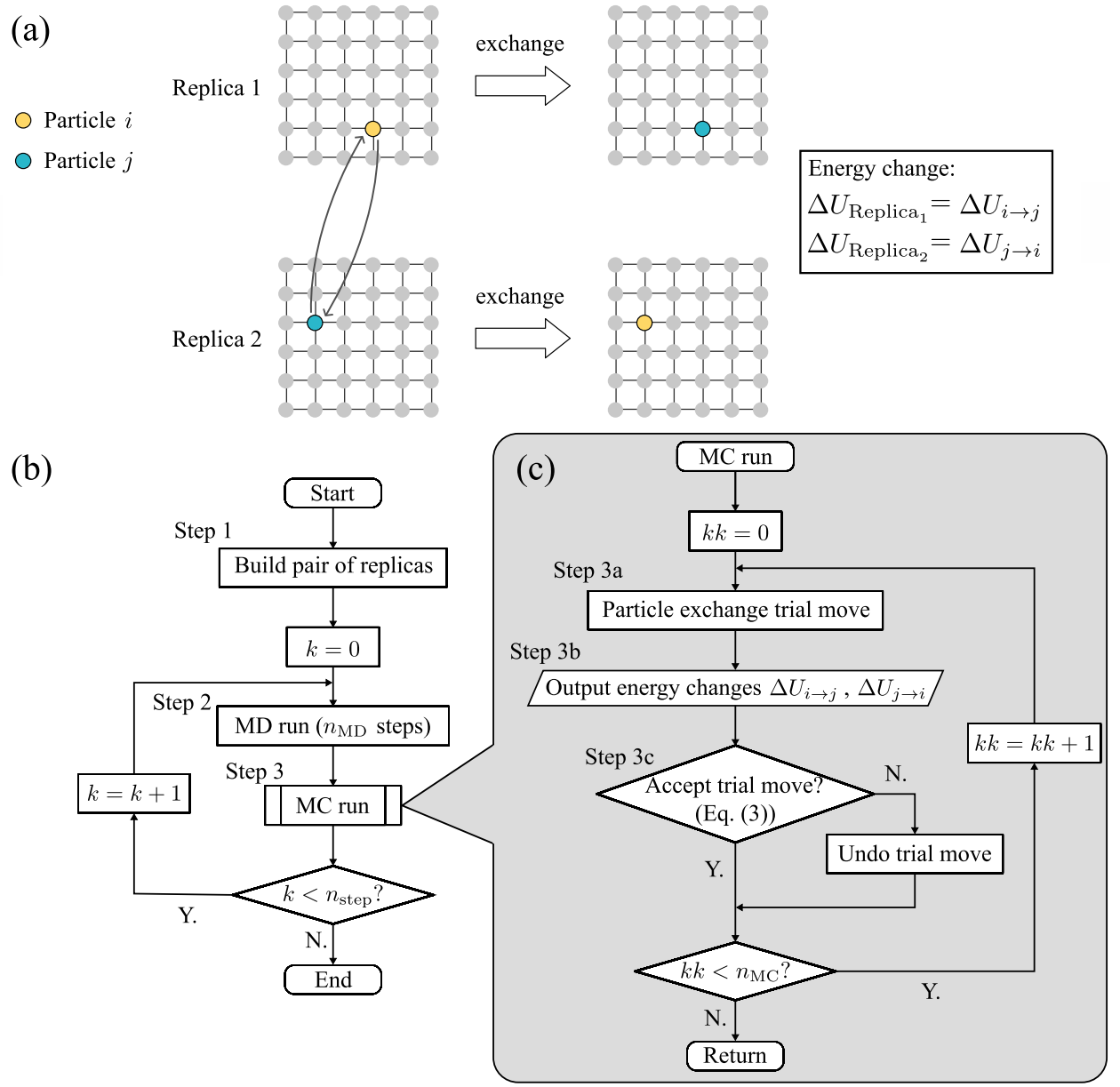

FIG. 1. (a) Particle exchange trial move. (b) Flowchart of the energy change sampling algorithm. (c) Flowchart of the MC run in the energy change sampling algorithm.

in the thermodynamic limit. Throughout this paper we use the term "energy change sampling" to refer to this algorithm. The energy change values sampled from the SGC ensemble in each replica can then be used to estimate the chemical potential differences between the system's components according to Eq. (2) [1] (see also Appendix A):

$$\mu_j - \mu_i = -\frac{1}{\beta} \ln \left\langle \frac{N_i}{N_j + 1} \exp\left(-\beta \Delta U_{i\to j}\right) \right\rangle. \qquad (2)$$

Here N_i denotes the number of particles of identity i, $\langle \dots \rangle$ denotes the SGC ensemble average, and $\Delta U_{i\to j}$ denotes the energy change of the system caused by an exchange of a particle of identity i for a particle of identity j. $\Delta U_{i\to j}$ is a function of both $\boldsymbol{r}^N$ and $\boldsymbol{N}$, so we must sample $\Delta U_{i\to j}$ values from the configuration and composition phase spaces in order to calculate the SGC ensemble average in Eq. (2).

A flowchart of the energy change sampling algorithm is shown in Fig. 1(b). First, we build a pair of replicas of the bulk system with a random chemical distribution (step 1). Next, we perform n_{step} MD or MC cycles on the replicas, each cycle consisting of n_{MD} MD steps (step 2) followed by an MC run (step 3) [see Fig. 1(c)]. The MD run is subject to the Nosé-Hoover thermostat in order to explore the configuration phase space of the replicas [10]. All MD runs reported in this paper were performed in LAMMPS [18,19]. We begin each step of the MC run by performing a particle exchange trial move (step 3a), as shown in Fig. 1(a), where a particle i in one replica is exchanged by a particle j in the other replica, and vice-versa [see Fig. 1(a)]. Note that this process actually exchanges the particles between the replicas, which differs from the particle identity change trial move in a single system. Next, we calculate and output the energy

change values $\Delta U_{\text{Replica}_1} = \Delta U_{i\to j}$ and $\Delta U_{\text{Replica}_2} = \Delta U_{j\to i}$ pertaining to each respective replica, where i and j denote the identities of particles exchanged during the trial move (step 3b). This step corresponds to the sampling of the energy change values and, as we will describe later, it is valuable to store the $\Delta U_{i\to j}$ and $\Delta U_{j\to i}$ values to build a set of databases from their probability distributions. The databases allow us to obtain useful insights into the chemical interactions between different components of the materials, which is a key information to figure out the energetical origin of the chemical ordering in solid solutions. This is the most important advantage of the database approach over the approaches that use only the chemical potential differences. The acceptance probability for the particle exchange trial move follows the standard METROPOLIS algorithm [Eq. (3)] [1] (step 3c):

$$p^{\text{Acc}} = \min\{1, \exp[-\beta(\Delta U_{\text{Replica}_1} + \Delta U_{\text{Replica}_2})]\}. \quad (3)$$

This trial move results in each replica constituting the SGC ensemble, analogous to the Gibbs ensemble method. Therefore, the energy change values $\Delta U_{i\to j}$ sampled during step 3b follow the probability distribution of the SGC ensemble and can be used to estimate the chemical potential differences $\mu_j - \mu_i$ in Eq. (2).

2. Energy database construction

The output data from the energy change sampling can be used to build energy databases for the database-driven semigrand canonical Monte Carlo (DD-SGCMC) simulation (Sec. II B 3). Let $\{\Delta U_{i\to j}\}$ be the set of all the energy change values caused by exchanging a particle of identity i for a particle of identity j in the energy change sampling [for an n-nary system, there are $n \times (n-1)$ such sets]. We can build the histogram of these sets, setting the bin size to $\Delta\varepsilon$. Normalizing the histogram yields the probability distribution function (PDF) of $\{\Delta U_{i\to j}\}$, and integrating the PDF produces the cumulative distribution function (CDF) of $\{\Delta U_{i\to j}\}$. As shown in Fig. 2, each energy database consists of the PDF and the CDF of $\{\Delta U_{i\to j}\}$ for each possible particle identity of i and j.

While building the histogram of the energy change values observed in the energy change sampling, we assumed that the probability distribution of the energy change values is continuous. This assumption contrasts with the fact that we could, in theory, calculate the energy change for each possible local environment around an exchange site and build the histogram of this discrete set of energy changes. However, in a large reservoir, we can reasonably expect some fluctuation in its structure, which causes the distribution of energy change values to become continuous.

It is relevant to note that, if the replicas used during the energy change sampling are too small, the PDFs of the energy change values will not be aligned with the corresponding PDFs of the bulk system. The size necessary for the PDFs to converge varies greatly with the material in question, but generally we can consider that systems with a long-range chemical ordering or large separated phases will require large replicas to ensure a sufficient sampling of the energy changes of particles (atoms) located at sites with a rare local chemical environment. Thus, in such systems the size of the replicas has to be much larger than the characteristic length of chemical ordering and separated phases.

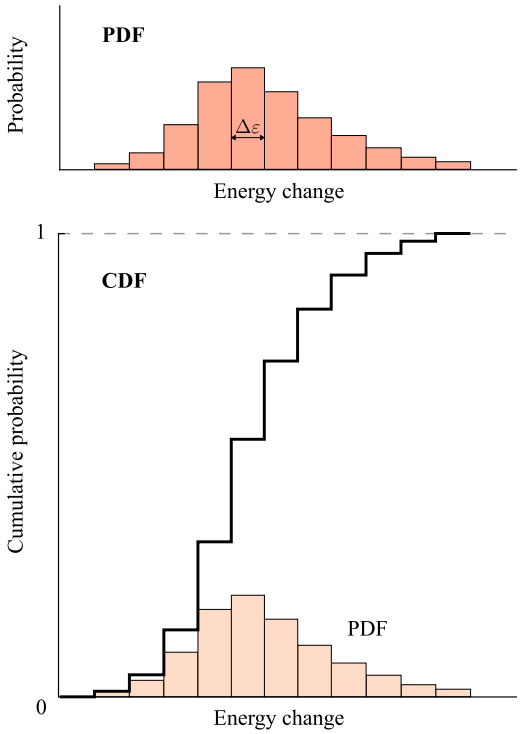

FIG. 2. Energy database information. Each energy database is built from the set of energy change values $\{\Delta U_{i\to j}\}$ output from the energy change sampling; these energy databases contain the probability distribution functions (PDFs) and the cumulative distribution functions (CDFs) of $\{\Delta U_{i\to j}\}$.

3. Database-driven SGCMC simulation

Figure 3(a) shows a flowchart of the DD-SGCMC algorithm. First, we build the target system (step 1) and subject it to an MD simulation with temperature and volume controlled by using a Nosé-Hoover thermostat (step 2). As in the energy change sampling simulation, the MD run is interrupted at regular intervals by an MC run (step 3). During the MC run, the target system undergoes particle identity change trial moves [see Fig. 3(b)].

We designed the MC run in such a way that the target system is described by the SGC ensemble. To demonstrate this, first consider a hypothetical MC simulation of a large atomic reservoir interacting with the target system in the SGC ensemble. If the transition probabilities between microstates of the target systems in this hypothetical MC simulation and in the DD-SGCMC simulation are equivalent, then the target system in the DD-SGCMC will also be described by the SGC ensemble.

In the hypothetical MC simulation, at each MC step we have a particle exchange trial move where a particle of identity i in the target system is exchanged with a particle of identity j in the reservoir. The transition probability $P(i \to j)$ of this MC simulation is shown in Eq. (4), where $c_{i,\text{tar}}$ is the concentration of i in the target system, $c_{j,\text{rsv}}$ is the concentration of j in the reservoir, and $p^{\text{Acc}}_{i\to j}$ is the acceptance probability of this

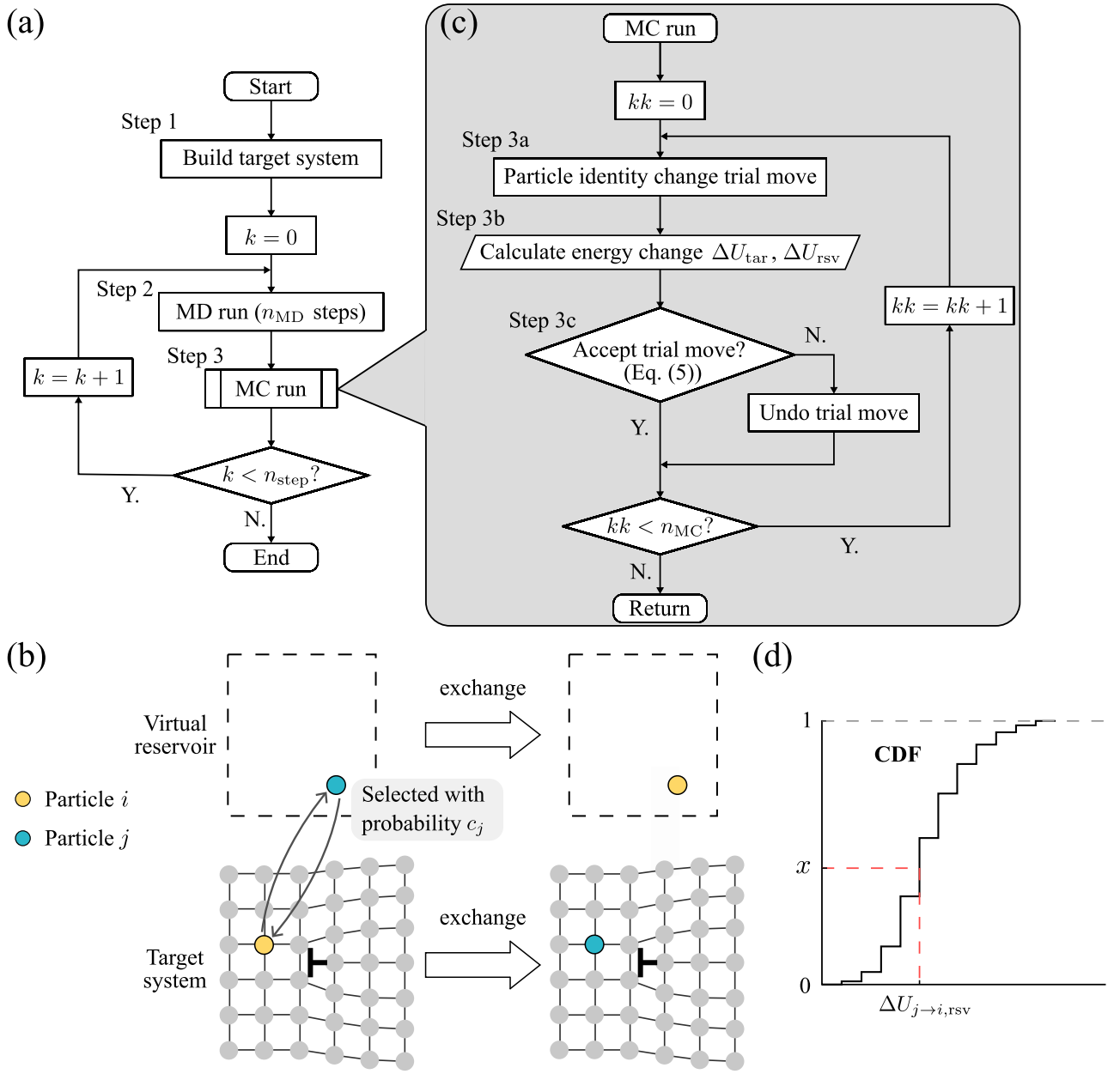

FIG. 3. (a) Flowchart of the DD-SGCMC algorithm. (b) Schematics of the particle identity change trial move. (c) Flowchart of the MC run in the DD-SGCMC algorithm. (d) Calculation of $\Delta U_{\rm rsv}$ from the energy database using a random number x.

trial move:

$$P(i \to j) = c_{i,\rm tar} \times c_{j,\rm rsv} \times p_{i\to j}^{\rm Acc}. \tag{4}$$

The acceptance probability $p_{i\to j}^{\rm Acc}$ is shown in Eq. (5), which follows from the METROPOLIS algorithm [1]:

$$p_{i\to j}^{\rm Acc} = \min\{1, \exp\big[-\beta\big(\Delta U_{i\to j,\rm tar} + \Delta U_{j\to i,\rm rsv}\big)\big]\}, \tag{5}$$

where $\Delta U_{i\to j,\rm tar}$ and $\Delta U_{j\to i,\rm rsv}$, respectively, denote the energy changes of the target system and the reservoir due to the trial move.

In order to build an MC run for the DD-SGCMC algorithm where the target system is described by the SGC ensemble, we impose that the transition probability $P(i \to j)$ must be given by Eq. (4). Figure 3(c) shows a flowchart of this MC run. First, we perform a particle identity change trial move [see Fig. 3(b)] where a particle of identity i is randomly selected from the target system (with probability $c_{i,\rm tar}$) and its identity is changed to j (step 3a). Following Eq. (4), the probability of choosing a particular identity j is set to $c_{j,\rm rsv}$, which is known from the composition of the bulk system. Next, the energy changes of the target system $\Delta U_{i\to j,\rm tar}$ and of the virtual reservoir $\Delta U_{j\to i,\rm rsv}$ are calculated (step 3b). As shown in Fig. 3(d), the energy change of the virtual reservoir $\Delta U_{j\to i,\rm rsv}$ can be obtained from the energy database of $\{\Delta U_{j\to i}\}$ by selecting a uniformly distributed random number x and finding the corresponding $\Delta U_{j\to i,\rm rsv}$ value in the CDF. Energy change values generated this way have a probability distribution identical to energy change values $\Delta U_{j\to i,\rm rsv}$ calculated from a reservoir in the hypothetical

MC simulation. Lastly, we calculate the acceptance probability for this particle identity change trial move (step 3c), which is given by Eq. (5), noting that $\Delta U_{j\to i,\mathrm{rsv}}$ was not calculated directly, but instead extracted from the energy database of $\{\Delta U_{j\to i}\}$.

III. APPLICATIONS AND METHOD VERIFICATION

A. Fe–1-at.-%-Si system

1. Database construction

We built the energy databases of a Fe–1-at.-%-Si system at 300, 500, and 1100 K. In both the energy change sampling and the DD-SGCMC simulation we used an MC and MD algorithm in which an MD run was interrupted every n_{MD} steps to run a cycle of n_{MC} MC steps. It should be noted that, after each particle exchange trial move, we additionally performed a structural relaxation, which consists of an energy minimization at 0 K using a conjugate gradient algorithm. In doing so, we sample from the configuration phase space only the "inherent structures" of the replicas, i.e., the local minima of the potential basin [20]. Such a calculation is not entirely accurate because it excludes the vibrational entropy contribution from the configuration phase space of the "true" SGC ensemble. However, to reduce computational costs, we perform this structural relaxation in the Fe–1-at.-%-Si and the NiCoCr MEA systems shown in this paper in Sec. III, and approximate the chemical potential difference values. The MD run is still used to explore the metastable configurations of the system from the configuration phase space of the SGC ensemble, which also prevents the system from getting trapped in very shallow metastable states that may appear in multicomponent systems containing defects. This approach greatly reduces the rejection rate of the particle identity change trial moves, leading to faster convergence to the equilibrium state. However, in computations for vibrational entropy dominant systems, which are systems stabilized by the vibrational entropy contribution [21,22], one should not perform this energy minimization. Due to the structural relaxation, the atomic sites move during the MC run. To offset this displacement, at the end of each MC run, we return the atomic sites to their positions at the end of the previous MD run.

The MD run had a step size of 1 fs, and the volume and temperature were maintained with a Nosé-Hoover thermostat. The chemical potential differences for the Fe–1-at.-%-Si system were obtained from a pair of bcc lattices with $15 \times 15 \times 12$ two-atom conventional unit cells at 300, 500, and 1100 K using an embedded atom method (EAM) potential to describe the Fe-Fe and Fe-Si interactions [23]. Si-Si interactions were described by an addition of Mie-type and Gauss-type potentials (see Appendix B). We set n_{MC} to 5400 steps and n_{MD} to 400 steps, with the total run lasting 7×10^4 MD steps. The obtained energy databases are shown in Fig. 4. The bin width $\Delta\varepsilon$ was set to 0.001 eV.

The replicas used to build the energy databases were considered to be in thermodynamic equilibrium when the running mean of the total energy every 1600 MD steps had stabilized with fluctuations of around 0.1% of the total energy. The total number of energy change values obtained was considered sufficiently large when the cumulative average of the total energy

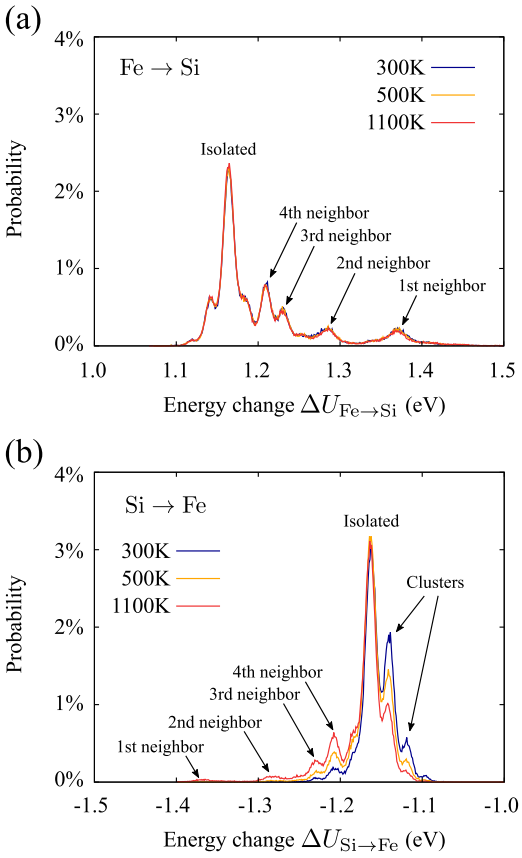

FIG. 4. Energy databases for the Fe–1-at.-%-Si system. (a) Fe to Si energy databases. (b) Si to Fe energy databases.

in equilibrium showed changes of less than 0.01%. We also tested various replica sizes (5400 and 16 000 atoms) to verify that the obtained energy databases had converged. Since the first moment and the standard deviation of the probability distributions differed by less than 0.003 eV, we considered the distributions to have converged.

We note that the energy databases contain relevant information on chemical ordering and its energetics that provides deeper insight into the equilibrium state of multicomponent alloy systems. For example, as seen in Fig. 4, the absence of strong temperature dependence in the $\{\Delta U_{\mathrm{Fe}\to\mathrm{Si}}\}$ energy databases suggests that the composition of neighbor shells around Fe atoms remains mostly unchanged with temperature, since a different composition of these neighbor shells would give a different energy change profile. We can verify this finding by calculating the probability $P_{i\text{-}j}$ of finding a particle with identity j near a particle with identity i, up to a certain cutoff distance. Figure 5 shows $P_{\mathrm{Fe\text{-}Si}}$ and $P_{\mathrm{Si\text{-}Si}}$ up to the third neighbor shell for each temperature analyzed in this paper, where $P_{\mathrm{Fe\text{-}Fe}}$ and $P_{\mathrm{Si\text{-}Fe}}$ are simply the respective complements of the values shown. This result confirms that the average local composition near Fe atoms does not change noticeably with temperature, unlike the case for Si atoms, which show a noticeable temperature dependence in neighbor shell average composition, suggesting a temperature-dependent chemical structure deviation (chemical ordering) from the fully random Si distribution. This analysis also demonstrates the tendency

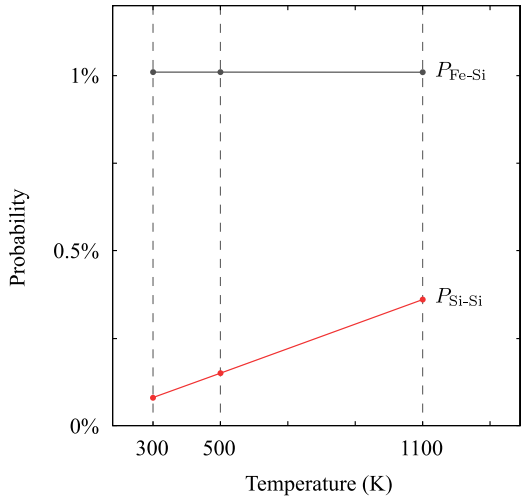

FIG. 5. Temperature-dependent probabilities of finding Si atoms out to the third neighbor shell around Fe and Si atoms.

for Si atoms to repel each other, as evidenced by the low $P_{\text{Si-Si}}$ values (noticeably below 1%) at each evaluated temperature. This tendency also seems to decrease linearly with temperature.

Another interesting observation drawn from the energy databases is the multiplicity of peaks, each appearing at a specific energy change value. These peaks imply that certain energy change values occur with high probability. Since the energy change due to a particle identity change trial move is closely related to the characteristics of the neighbor shells surrounding the exchange site, we can attribute each peak to a neighbor shell profile. In this case, the highest peak in each energy database corresponds to the situation in which there are no Si atoms around the exchange site (i.e., the Si atom is isolated). Figure 4 lists the neighbor shell profiles corresponding to each energy peak, according to the location of the nearest Si atom to the exchange site. Figure 4 also includes the situations where three or more Si atoms form a cluster; as we can observe by comparing the height of the peaks at different temperatures, this situation occurred more often at lower temperatures.

2. *Temperature-dependent grain boundary segregation in Fe–1-at.-% Si*

The constructed energy databases for the Fe–1-at.-%-Si system were used in DD-SGCMC-based analysis of grain boundary segregation at a $\Sigma3(112)[1\,\bar{1}\,0]$ symmetric tilt GB at 300, 500, and 1100 K [see Fig. 6(a), where the position of the GB is highlighted using a common neighbor analysis to identify the local crystal structures]. Initially, the target system contained only Fe atoms. To calculate the ensemble-averaged Si distribution, the average concentration of Si in each atomic layer throughout the MC and MD run within each layer was extracted and plotted in Fig. 6(b), where the GB position is highlighted in gray. This result confirms that the Si concentration in the region far from the GB, i.e., the bulk Si concentration, is close to 1 at. %. The exact values are included in Table I.

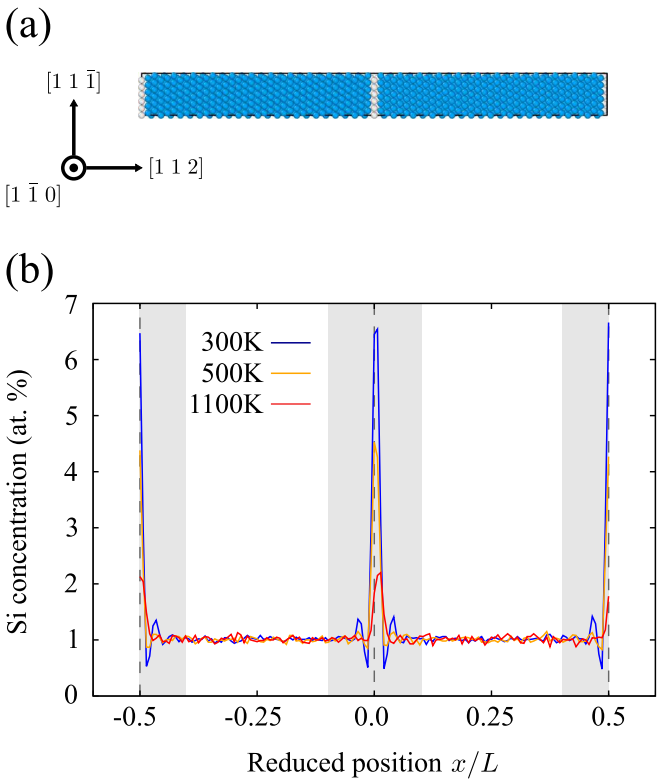

FIG. 6. (a) A $\Sigma3(112)[1\,\bar{1}\,0]$ symmetric tilt grain boundary system containing 4320 atoms. Atoms ordered in bcc structure are colored in blue, while atoms in another type of structure, namely, the atoms on the grain boundary, are colored in white. The atomic system was created by using ATOMSK [24], and the figure was obtained by using OVITO [25]. (b) Si concentration along the direction perpendicular to the GB, as obtained using a DD-SGCMC method (starting position of the center of the GB shown by dashed lines). The dimensions of the system are 165.430 × 31.837 × 19.496Å, and L is equal to 165.43 Å.

It is important to compare the results obtained from SGCMC simulations and DD-SGCMC simulations to verify that they are identical. We compared segregation and atomic distribution results in the presence of defects in a $\Sigma3(112)[1\,\bar{1}\,0]$ symmetric tilt GB in a Fe–1-at.-%-Si system at 300, 500, and 1100 K. We estimated the chemical potential difference input values according to Eq. (2).

The calculated SGCMC chemical potential difference input values at each temperature are shown in Table II, and the Si distribution obtained using these inputs is shown in Fig. 7(a). Figure 7(b) compares the concentrations near the GB obtained by using DD-SGCMC and SGCMC methods. The two approaches produced average concentrations of Si

TABLE I. Bulk Si concentration.

Temperature (K)	Concentration (at. %)
300	1.02
500	1.01
1100	1.01

TABLE II. Chemical potential differences for the Fe–1-at.-%-Si system.

Temperature (K)	$\mu_{Si} - \mu_{Fe}$ (eV)
300	1.051
500	0.981
1100	0.751

atoms near the GB differing by less than 0.5 at. % of the atoms in the atomic layer containing the GB.

It is also possible to obtain the ensemble-averaged grain boundary energy γ_{GB} directly from the data obtained during the DD-SGCMC simulation and the energy change sampling simulation as

$$\gamma_{GB} = \left[\langle U_{tar}\rangle - (\langle N_{Si}\rangle - N_{tar}c_{Si})(\mu_{Si}-\mu_{Fe}) - \frac{N_{tar}}{N_{rep}}\langle U_{rep}\rangle\right] \Big/ A_{GB}, \tag{6}$$

where the lower indices "tar" and "rep," respectively, denote quantities obtained from the target system in the DD-SGCMC simulation and from the replicas used in the energy change sampling simulation. U denotes the system energy (eV), N denotes the total number of particles in the system, N_{Si} denotes the number of Si particles in the target system, c_{Si} denotes the desired Si concentration, and A_{GB} denotes the GB area (Å^2). We calculated γ_{GB} for each temperature and plotted the results in Fig. 8. To analyze the difference in energy due to Si segregation, we also calculated the GB energy in pure Fe using the same potential, which can be obtained from Eq. (7). In Eq. (7), the indices tar and "perf," respectively, denote quantities obtained from a pure Fe target system containing a GB and a perfect crystal system; U denotes the system energy; and N denotes the total number of particles in the system. Since solute segregation is expected to decrease the energy of a defect, the decrease of solute segregation with increasing temperature, as shown in Fig. 7, implies that the grain boundary energy should increase with increasing temperature. Figure 8 confirms this implication:

$$\gamma_{GB,\,pure} = \left[U_{tar} - \frac{N_{tar}}{N_{perf}}U_{perf}\right] \Big/ A_{GB}. \tag{7}$$

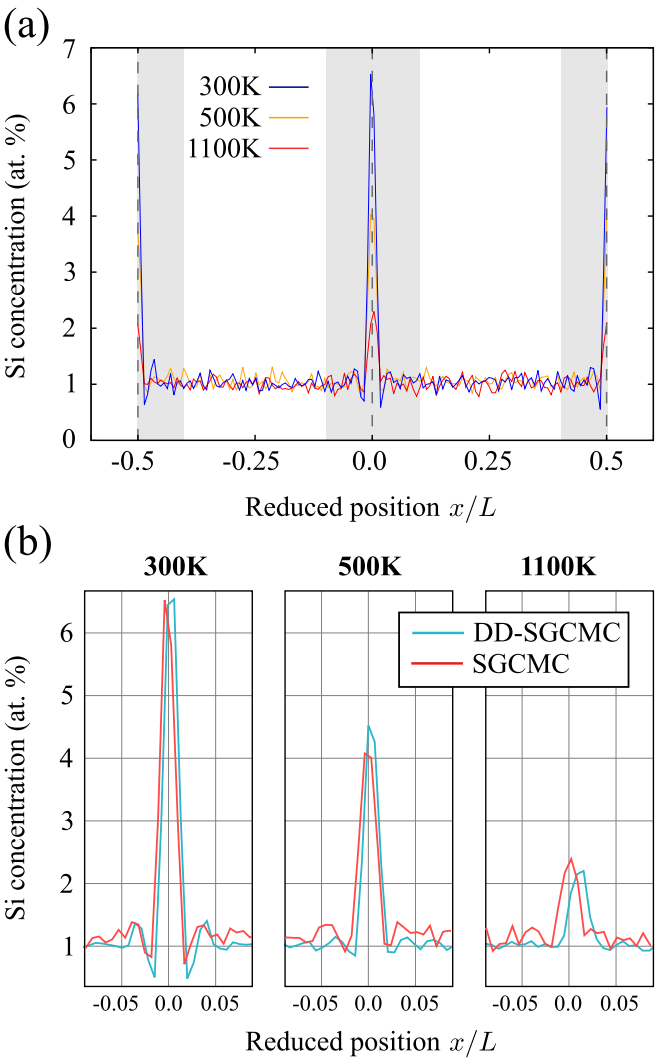

FIG. 7. (a) Si concentration along the direction perpendicular to the GB, as obtained using an SGCMC method. (b) Comparison between the Si concentrations obtained with database-driven SGCMC and SGCMC algorithms near the GB. The total length of the system L is equal to 165.43 Å

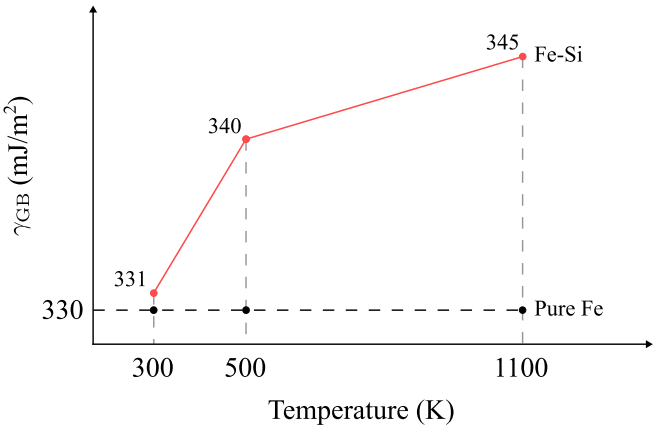

FIG. 8. GB energy of the Fe–1-at.-%-Si alloy at 300, 500, and 1100 K. The calculated GB energy of pure Fe is also included for comparison.

B. NiCoCr MEA system

1. Energy database construction

We also obtained the energy databases for a NiCoCr MEA system at 850 and 1050 K. As above, we used a step size of 1 fs, and the volume and temperature were controlled with a Nosé-Hoover thermostat. After each particle identity change trial move we performed a structural relaxation using a conjugate gradient algorithm. The replicas were composed of fcc lattices with $6 \times 6 \times 6$ conventional unit cells of four atoms each, at 850 and 1050 K, with interatomic interactions expressed by the NiCoCr EAM potential developed by Li *et al.*

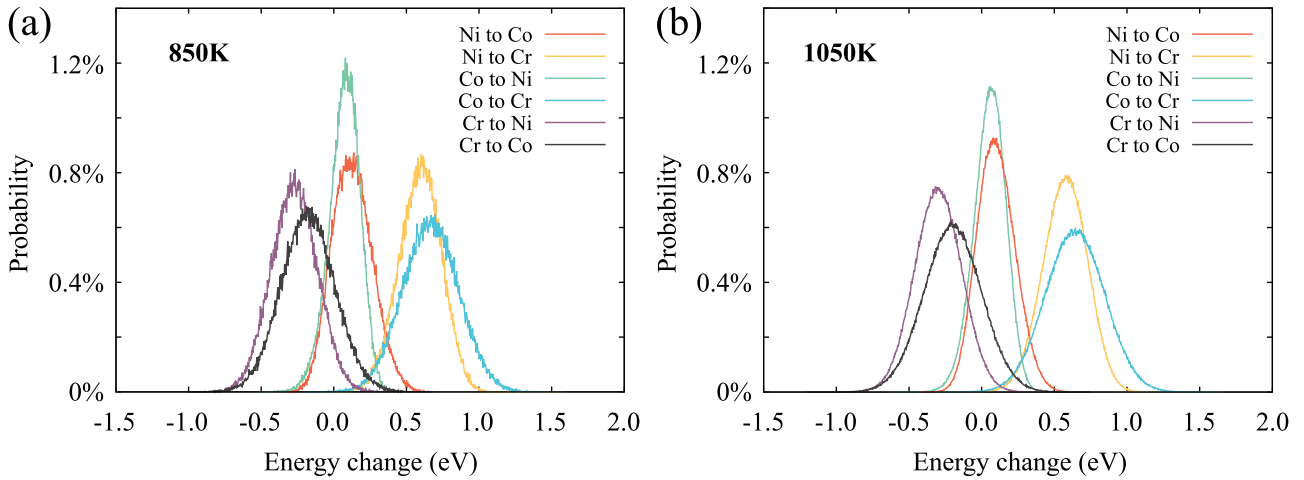

FIG. 9. Energy databases for the NiCoCr MEA system at (a) 850 K and (b) 1050 K.

[12]. Here, n_{MD} was set to 400 steps and n_{MC} was set to 864 steps with a total length of 2.0×10^6 MD steps. The obtained energy databases are shown in Fig. 9. The bin width $\Delta\varepsilon$ was set to 0.003 eV.

As in the Fe–1-at.-%-Si systems, here the replicas used in the energy change sampling were considered to be in thermodynamic equilibrium when the running mean of the total energy every 8000 MD steps had stabilized with fluctuations of around 0.1% the total energy. The total number of energy change values obtained was considered sufficiently large when the cumulative average of the total energy in equilibrium showed changes of less than 0.01%. We also tested various replica sizes (108, 216, 432, and 1536 atoms) to verify that the obtained energy databases had converged. Once the first moment and the standard deviation of the probability distributions differed by less than 0.003 eV, we considered the distributions to have converged.

Figure 9 shows that each of the energy databases exhibits a distribution close to normal. The standard deviation of each of these distributions increased with temperature, with the exception of the energy database of $\{U_{\mathrm{Ni}\to\mathrm{Co}}\}$ (see Table III). This change suggests a decrease in the short-range order (SRO) of the system. We confirm this suggestion more explicitly in Sec. III B 2.

The energy databases were used to calculate the chemical potential differences according to Eq. (2). The results are shown in Table IV. Using these chemical potential differences as inputs for SGCMC simulations of the bulk system, the average concentration of each element was found to be $33.3 \pm 0.3\%$.

TABLE III. Standard deviations for the energy databases for the NiCoCr MEA system.

	Standard deviation (eV)	
Energy database	850 K	1050 K
Ni to Co	0.138	0.130
Ni to Cr	0.146	0.152
Co to Ni	0.105	0.110
Co to Cr	0.192	0.202
Cr to Ni	0.158	0.161
Cr to Co	0.184	0.197

2. *Short-range order change with equilibrium temperature in NiCoCr MEA*

MEA structures present some degree of SRO that decreases with increasing temperature for most atomic pairs, as investigated by Li *et al.* [12]. The SRO of a system can be measured using the SRO parameters defined by de Fontaine [26], shown in Eq. (8). We used a DD-SGCMC simulation to calculate the SRO parameters in a $16 \times 16 \times 12$ fcc unit cell NiCoCr MEA system at 850 and 1050 K to showcase the application of the method:

$$\alpha_{ij} = 1 - \frac{P_{i-j}}{\overline{c_j}} \ (i \neq j),$$
$$\alpha_{ii} = \frac{P_{i-i} - \overline{c_i}}{1 - \overline{c_i}}. \tag{8}$$

Here α_{ij} denotes the SRO parameter between particles with identity i and j ($\alpha_{ij} = \alpha_{ji}$), P_{i-j} denotes the probability of finding a particle with identity j near a particle with identity i up to a given cutoff radius (in this case up to the first neighbor shell), and $\overline{c_i}$ denotes the average concentration of particles with identity i in the system.

Using the energy databases shown in Fig. 9 in a DD-SGCMC simulation of bulk NiCoCr, the average concentration of each element was found to be $33.3 \pm 0.1\%$. The ensemble-averaged SRO was calculated according to the method presented by de Fontaine [26]; Fig. 10 compares SRO parameter values α_{ij} as obtained here with those calculated by

TABLE IV. Chemical potential differences for the NiCoCr MEA system.

Temperature (K)	$\mu_{\mathrm{Co}} - \mu_{\mathrm{Ni}}$ (eV)	$\mu_{\mathrm{Cr}} - \mu_{\mathrm{Ni}}$ (eV)
850	0.437	0.019
1050	0.438	0.019

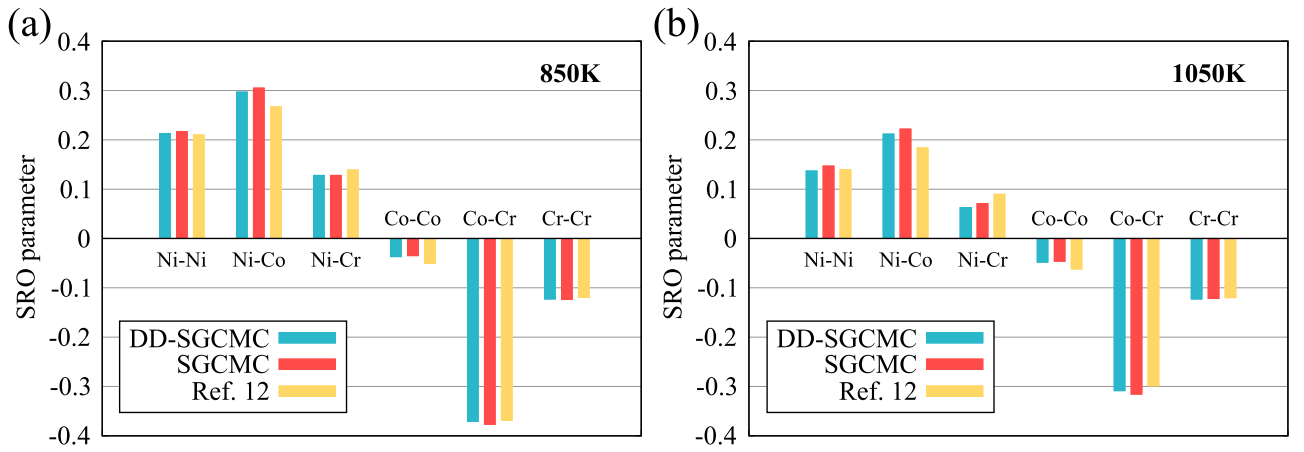

FIG. 10. Comparison of the short-range order parameters for the NiCoCr MEA system at (a) 850 K and (b) 1050 K, as obtained by using database-driven SGCMC and SGCMC methods and the short-range order parameters calculated by Li *et al.* [12].

Li *et al.* [12]. These comparisons clearly demonstrate that the DD-SGCMC method can reproduce conventional SGCMC results.

3. Element segregation near a stacking fault in NiCoCr MEA

Segregation near stacking faults in medium- and high-entropy alloys has also gained attention in recent years. Since there is a close relation between stacking faults and the deformation mechanisms of these alloys, changes in the local chemical composition near the stacking fault affect the mechanical properties of the material [27–33].

Therefore, we also used the energy databases shown in Fig. 9 to examine the distribution of elements in NiCoCr MEA near an extrinsic stacking fault (ESF) at 850 and 1050 K. The target system of the DD-SGCMC simulation is shown in Fig. 11. Figure 11 also highlights the position of the ESF by coloring the particles of the system according to their structure, as identified using a common neighbor analysis. The obtained element distributions along the direction perpendicular to the ESF at 850 and 1050 K are shown in Fig. 12. The segregation of Ni that we observed near the ESF potentially changes the extrinsic stacking fault energy (ESFE), a property closely related to the twinning behavior of the material. Figure 12 shows that segregation decreases with increasing temperature as the configuration entropic effect becomes more dominant.

For purposes of comparison, we also included the results of SGCMC simulations in Fig. 12 to confirm the general trend of Ni segregation to the ESF. The chemical potential differences used for the SGCMC simulations are shown in Table IV.

Expanding Eq. (6) for an n-nary alloy with element types denoted by $S = \{1, \ldots, n\}$, we obtain Eq. (9), which can be used to calculate the ESFE γ_{ESF}:

$$\gamma_{\mathrm{ESF}} = \left[\langle U_{\mathrm{tar}} \rangle - \sum_{i \in S \setminus \{1\}} (\langle N_i \rangle - N_{\mathrm{tar}} c_i)(\mu_i - \mu_1) - \frac{N_{\mathrm{tar}}}{N_{\mathrm{rep}}} \langle U_{\mathrm{rep}} \rangle \right] \Big/ A_{\mathrm{SF}}. \tag{9}$$

Here the lower indices tar and rep, respectively, denote quantities obtained from the target system in the DD-SGCMC simulation and from the replicas used in the energy change sampling simulation. U denotes the system energy (eV), N denotes the total number of particles in the system, N_i denotes the number of particles of identity i in the target system, c_i denotes the desired concentration of particles of identity i, and A_{SF} denotes the area of the stacking fault (Å^2). During the calculation, we arbitrarily assigned Ni to particle identity 1 in the context of the equation above. Figure 13(a) shows the obtained ESFE values.

Since there was a thermodynamic driving force that caused the Ni atoms to segregate toward the ESF, we also calcu-

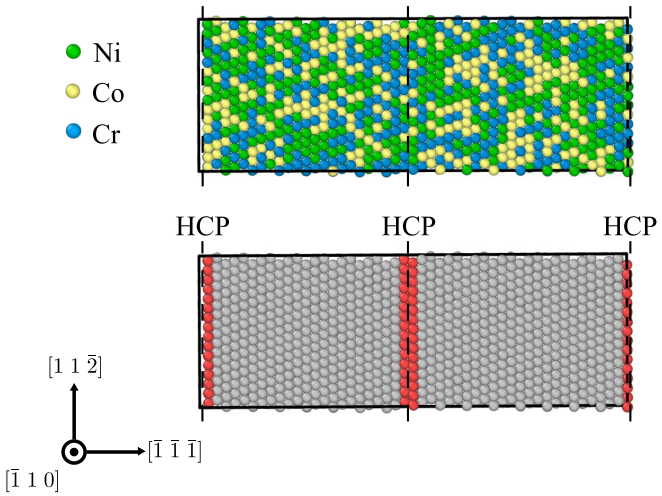

FIG. 11. NiCoCr MEA system with stacking fault containing 9216 atoms. The atoms in HCP structure are highlighted in red, and the atoms in fcc structure are shown in gray. The atomic system was created by using ATOMSK [24], and the figures were obtained by using OVITO [25]. The dimensions of the system are 30.176 × 98.554 × 34.844Å.

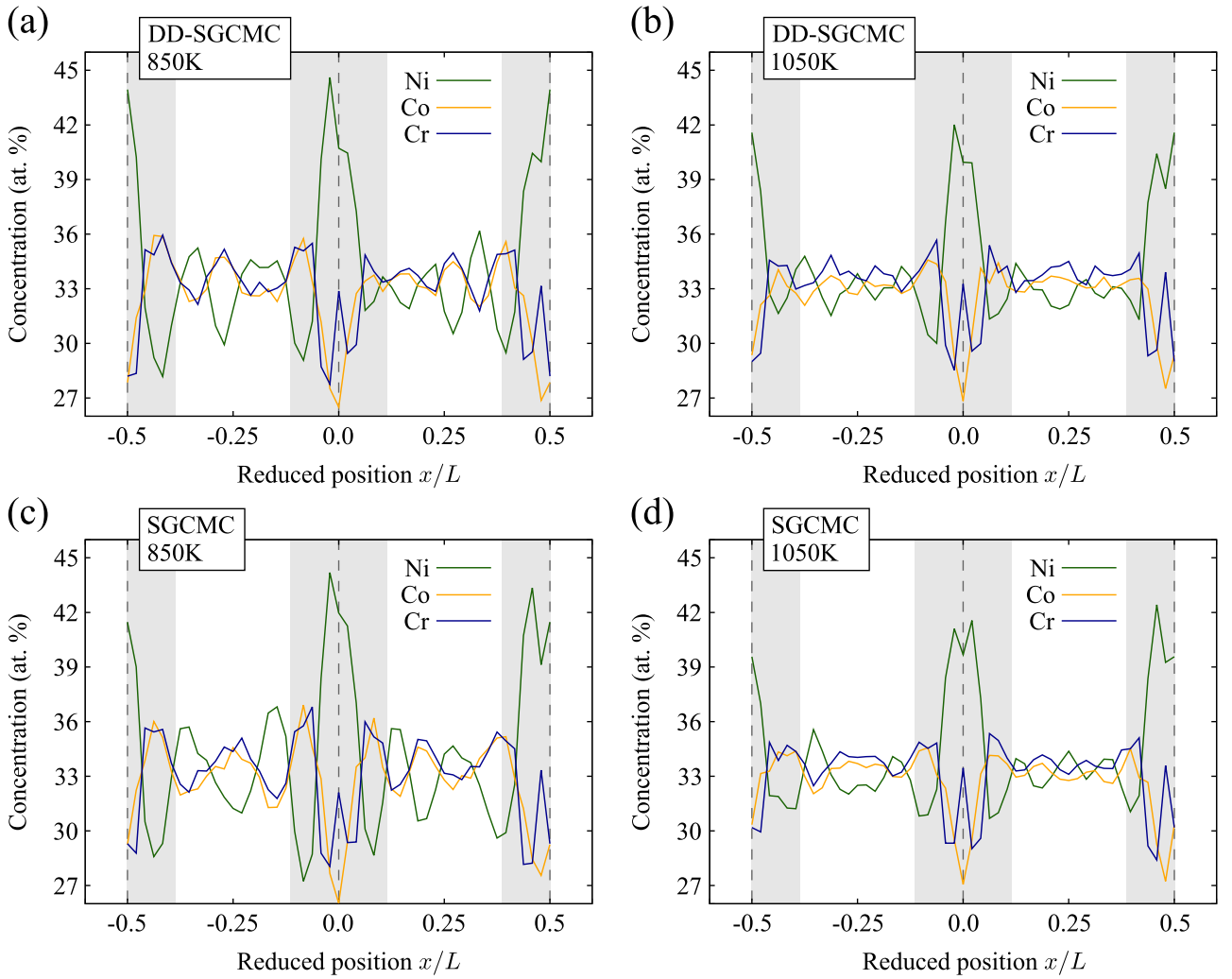

FIG. 12. Elemental distribution in a NiCoCr MEA stacking fault system according to database-driven SGCMC simulations at (a) 850 K and (b) 1050 K. Elemental distribution in a NiCoCr MEA stacking fault system according to SGCMC simulations at (c) 850 K and (d) 1050 K. The total length of the system L is equal to 98.55 Å.

lated the energy difference due to segregation. The energy of the ESF system after segregation was obtained by relaxing the atomic structures of the system obtained from the DD-SGCMC simulation of the ESF target system. To obtain the ESF system energy before segregation, we first selected atomic structures obtained from DD-SGCMC simulations with a perfect crystal as the target system, and then created two ESFs by shifting half of the (111) atomic planes in the $\langle 11\bar{2} \rangle$ direction by 1 Å. We subsequently relaxed these structures, thereby generating systems similar to the one shown in Fig. 11. The ESF system energy before segregation was the average energy obtained from these systems after relaxation. We point out that these systems do not rigorously have the same SRO degree as the bulk system due to the change in distance between atoms close to the stacking fault. The ESFE values before and after segregation were calculated from Eq. (9), and their difference, here referred to as the segregation energy, is plotted in Fig. 13(b). This result confirms that the decrease in entropy expected due to segregation near the stacking fault was accompanied by a decrease in energy. As expected, this decrease was more significant at 850 K than at 1050 K, since segregation was more pronounced at 850 K.

IV. CONCLUSION

We have presented a method for running semigrand canonical ensemble Monte Carlo simulations starting from a nonequilibrium state using a set of energy databases that act as an external atomic reservoir. In doing so, we constrained the bulk concentrations to specified values without the need to precalculate any chemical potential-related parameters in the acceptance probability calculation. This method thus simplifies the process of performing semigrand canonical ensemble simulations. This method aids in metallic alloy research, since it can be promptly applied to systems with two or more components while requiring only a single Monte Carlo run to obtain the energy databases. These energy databases can then be used as input in database-driven SGCMC algo-

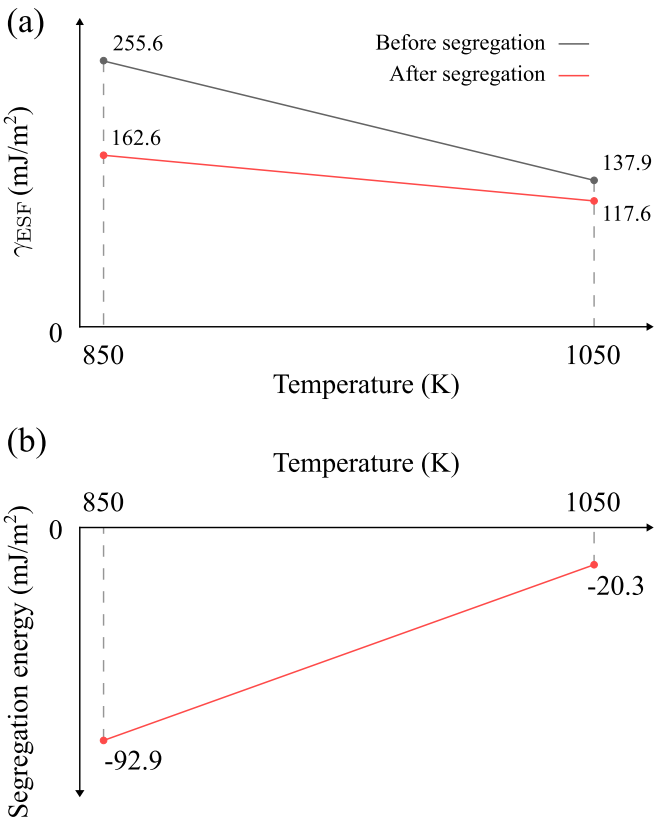

FIG. 13. (a) Extrinsic stacking fault energy of NiCoCr MEA at 850 and 1050 K before segregation and after segregation. (b) Segregation energy in NiCoCr MEA systems at 850 and 1050 K containing an extrinsic stacking fault.

rithms for a system of given composition. Notably, the energy database also offers a glimpse into the energetic and structural information regarding chemical ordering in alloys. Finally, we applied this method to a Fe–1-at.-%-Si system containing a GB and to NiCoCr MEA systems with and without a stacking fault defect. In each case, we obtained results equivalent to those calculated from an SGCMC approach. We expect this method to facilitate the calculation of thermodynamic quantities and the observation of phenomena in systems in thermodynamic equilibrium. Furthermore, we showed that the energy databases can contain relevant information regarding the chemical interactions between components of a material system using the Fe–1-at.-%-Si and the NiCoCr MEA systems.

ACKNOWLEDGMENTS

S.O. acknowledges support by Japan Society for the Promotion of Science (JSPS) KAKENHI Grants No. JP18H05453, No. JP17H01238, and No. JP17K18827, through the Elements Strategy Initiative for Structural Materials. S.S. acknowledges support by JSPS KAKENHI Grant No. JP19K23487. This work was partly achieved through the use of large-scale computer systems OCTOPUS at the Cybermedia Center, Osaka University and MASAMUNE-IMR at the Center for Computational Materials Science, Tohoku University.

APPENDIX A: CHEMICAL POTENTIAL DIFFERENCE ESTIMATION

Below we demonstrate Eq. (2), which establishes a connection between the energy change values calculated from the energy change sampling and the chemical potential inputs used in the traditional SGCMC approach. This demonstration is similar to the calculation of the chemical potential in the Gibbs ensemble, with the exception that the volume of each replica is constant [1].

The energy change sampling simulation (Sec. II B 1) consists of two n-nary replicas, each with a constant volume V and total number of particles N, where the number of particles of identity m in a replica p ($p = 1$ or 2) is denoted by $N_{m,p}$. These replicas are coupled with an infinitely large heat reservoir at a temperature T. In a replica p, the number of particles of each identity $N_{m,p}$ can be described by a n-dimensional vector $\boldsymbol{N}_p = \sum_m N_{m,p}\boldsymbol{e}_m$, where the set of $\boldsymbol{e}_m$ vectors forms the standard basis of $\mathbb{R}^n$, and the generalized coordinates of the particles are expressed by the vector $\boldsymbol{r}_p^N \in \mathbb{R}^{3N}$. The canonical partition function $Q(V, T, \boldsymbol{N}_p)$ of a replica p is then written as

$$Q(V, T, \boldsymbol{N}_p) = \prod_{m=1}^{n} \frac{q_m^{N_{m,p}}}{N_{m,p}!} \int d\boldsymbol{r}_p^N \exp\left[-\beta U\left(\boldsymbol{r}_p^N, \boldsymbol{N}_p\right)\right], \tag{A1}$$

where q_m denotes the translational contribution of the particles of identity m to the partition function, and $U(\boldsymbol{r}_p^N, \boldsymbol{N}_p)$ denotes the potential energy of the replica p. The two replicas interact with each other through particle exchanges, keeping the total number of particles of each identity in the two replicas constant, i.e., $\boldsymbol{N}_1 + \boldsymbol{N}_2 = \boldsymbol{N}$, where $\boldsymbol{N}(=\text{constant vector}) \in \mathbb{R}^n$. The partition function of both replicas $\bar{Q}(V, T, \boldsymbol{N})$ in the energy change sampling simulation is defined in Eq. (A2):

$$\bar{Q}(V, T, \boldsymbol{N}) = \sum_{\boldsymbol{N}_1=\boldsymbol{0}^n} Q(V, T, \boldsymbol{N}_1) Q(V, T, \boldsymbol{N} - \boldsymbol{N}_1). \tag{A2}$$

As shown in Eq. (A3), for a sufficiently large N, we can define the chemical potentials μ_i and μ_j pertaining to particles of identity i and j at replica 1, respectively:

$$\mu_i = -\frac{1}{\beta} \ln\left(\sum_{\boldsymbol{N}_1=\boldsymbol{0}^n} \frac{Q(V, T, \boldsymbol{N}_1)}{Q(V, T, \boldsymbol{N}_1 - \boldsymbol{e}_i)} Q(V, T, \boldsymbol{N} - \boldsymbol{N}_1)\right),$$
$$\mu_j = -\frac{1}{\beta} \ln\left(\sum_{\boldsymbol{N}_1=\boldsymbol{0}^n} \frac{Q(V, T, \boldsymbol{N}_1 + \boldsymbol{e}_j)}{Q(V, T, \boldsymbol{N}_1)} Q(V, T, \boldsymbol{N} - \boldsymbol{N}_1)\right). \tag{A3}$$

Note that the ratio between partition functions in Eq. (A3) can be expressed as

$$\frac{Q(V,T,\boldsymbol{N}_1)}{Q(V,T,\boldsymbol{N}_1-\boldsymbol{e}_i)} = \frac{\prod_{m=1}^{n}\frac{q_m^{N_{m,1}}}{N_{m,1}!}\int d\boldsymbol{r}_1^N \exp\left[-\beta U\left(\boldsymbol{r}_1^N,\boldsymbol{N}_1\right)\right]}{\frac{N_{i,1}}{q_i}\prod_{m=1}^{n}\frac{q_m^{N_{m,1}}}{N_{m,1}!}\int d\boldsymbol{r}_1^{N-1} \exp\left[-\beta U\left(\boldsymbol{r}_1^{N-1},\boldsymbol{N}_1-\boldsymbol{e}_i\right)\right]}$$
$$= \frac{q_i}{N_{i,1}}\frac{\int d\boldsymbol{r}_1^N \exp\left(-\beta\Delta U_i\right)\exp\left[-\beta U\left(\boldsymbol{r}_1^{N-1},\boldsymbol{N}_1-\boldsymbol{e}_i\right)\right]}{\int d\boldsymbol{r}_1^{N-1}\exp\left[-\beta U\left(\boldsymbol{r}_1^{N-1},\boldsymbol{N}_1-\boldsymbol{e}_i\right)\right]},$$
$$\frac{Q(V,T,\boldsymbol{N}_1+\boldsymbol{e}_j)}{Q(V,T,\boldsymbol{N}_1)} = \frac{\frac{q_j}{N_{j,1}+1}\prod_{m=1}^{n}\frac{q_m^{N_{m,1}}}{N_{m,1}!}\int d\boldsymbol{r}_1^{N+1} \exp\left[-\beta U\left(\boldsymbol{r}_1^{N+1},\boldsymbol{N}_1+\boldsymbol{e}_j\right)\right]}{\prod_{m=1}^{n}\frac{q_m^{N_{m,1}}}{N_{m,1}!}\int d\boldsymbol{r}_1^{N} \exp\left[-\beta U\left(\boldsymbol{r}_1^{N},\boldsymbol{N}_1\right)\right]}$$
$$= \frac{q_j}{N_{j,1}+1}\frac{\int d\boldsymbol{r}_1^{N+1}\exp\left(-\beta\Delta U_j\right)\exp\left[-\beta U\left(\boldsymbol{r}_1^{N+1},\boldsymbol{N}_1\right)\right]}{\int d\boldsymbol{r}_1^{N}\exp\left[-\beta U\left(\boldsymbol{r}_1^{N},\boldsymbol{N}_1\right)\right]}, \tag{A4}$$

where $\Delta U_i = U(\boldsymbol{r}_1^N,\boldsymbol{N}_1) - U(\boldsymbol{r}_1^{N-1},\boldsymbol{N}_1-\boldsymbol{e}_i)$ and $\Delta U_j = U(\boldsymbol{r}_1^{N+1},\boldsymbol{N}_1+\boldsymbol{e}_j) - U(\boldsymbol{r}_1^{N},\boldsymbol{N}_1)$. Thus, following Eqs. (A3) and (A4), the chemical potential difference $\mu_j - \mu_i$ can be expressed as

$$\mu_j - \mu_i = -\frac{1}{\beta}\ln\left(\frac{\sum_{\boldsymbol{N}_1=\boldsymbol{0}^n}\frac{Q(V,T,\boldsymbol{N}_1+\boldsymbol{e}_j)}{Q(V,T,\boldsymbol{N}_1)}Q(V,T,\boldsymbol{N}-\boldsymbol{N}_1)}{\sum_{\boldsymbol{N}_1=\boldsymbol{0}^n}\frac{Q(V,T,\boldsymbol{N}_1)}{Q(V,T,\boldsymbol{N}_1-\boldsymbol{e}_i)}Q(V,T,\boldsymbol{N}-\boldsymbol{N}_1)}\right)$$
$$= -\frac{1}{\beta}\ln\left(\frac{\sum_{\boldsymbol{N}_1=\boldsymbol{0}^n}\frac{q_j}{N_{j,1}+1}\frac{\int d\boldsymbol{r}_1^{N+1}\exp\left(-\beta\Delta U_j\right)\exp\left[-\beta U\left(\boldsymbol{r}_1^{N+1},\boldsymbol{N}_1\right)\right]}{\int d\boldsymbol{r}_1^{N}\exp\left[-\beta U\left(\boldsymbol{r}_1^{N},\boldsymbol{N}_1\right)\right]}\times Q(V,T,\boldsymbol{N}-\boldsymbol{N}_1)}{\sum_{\boldsymbol{N}_1=\boldsymbol{0}^n}\frac{q_i}{N_{i,1}}\frac{\int d\boldsymbol{r}_1^{N}\exp\left(-\beta\Delta U_i\right)\exp\left[-\beta U\left(\boldsymbol{r}_1^{N-1},\boldsymbol{N}_1-\boldsymbol{e}_i\right)\right]}{\int d\boldsymbol{r}_1^{N-1}\exp\left[-\beta U\left(\boldsymbol{r}_1^{N-1},\boldsymbol{N}_1-\boldsymbol{e}_i\right)\right]}\times Q(V,T,\boldsymbol{N}-\boldsymbol{N}_1)}\right)$$
$$= -\frac{1}{\beta}\ln\left(\frac{q_j}{q_i}\right) - \frac{1}{\beta}\ln\left(\frac{\left\langle\frac{1}{N_{j,1}+1}\exp\left(-\beta\Delta U_j\right)\right\rangle}{\left\langle\frac{1}{N_{i,1}}\exp\left(-\beta\Delta U_i\right)\right\rangle}\right)$$
$$= -\frac{1}{\beta}\ln\left(\frac{q_j}{q_i}\right) - \frac{1}{\beta}\ln\left\langle\frac{N_{i,1}}{N_{j,1}+1}\exp\left(-\beta\Delta U_{i\to j}\right)\right\rangle, \tag{A5}$$

where $\Delta U_{i\to j} = \Delta U_j - \Delta U_i$ is equivalent to the energy change of the system due to the change of identity of one particle from i to j at the thermodynamic limit, $-\frac{1}{\beta}\ln\left(\frac{q_j}{q_i}\right)$ is the ideal gas contribution, and $\langle\ldots\rangle$ denotes the average over the ensemble of replica 1. An analogous result can be obtained for replica 2. The ensemble of both replicas corresponds to the NVT ensemble; however, the ensemble of a single replica in the thermodynamic limit is equivalent to the semigrand canonical ensemble, in which case the average $\langle\ldots\rangle$ in Eq. (A5) [and in Eq. (2)] corresponds to the SGC ensemble average. The acceptance probability of an SGCMC particle identity change trial move is determined by the excess chemical potential difference, therefore we consider only the second term on the right-hand side of Eq. (A5). For simplicity, we refer to the excess chemical potential difference as the chemical potential difference. Additionally, since both replicas reproduce the same SGC ensemble in the thermodynamic limit, we do not make a distinction between the chemical potential differences obtained from each replica. Thus, we denote each $N_{m,p}$ in Eq. (A5) as N_m in the main text.

APPENDIX B: INTERATOMIC POTENTIAL FOR THE FE–1-AT.-%-SI SYSTEM

In this research, the interaction between atoms in the Fe–1-at.-%-Si system was defined using a combination of EAM potential and a hybrid potential composed of the addition of Mie-type and Gauss-type potentials [23]. The hybrid potential was used to describe Si-Si interactions, and its parameters were set to fit first-principle density-functional theory (DFT) calculations of the two-body interactions.

To calculate the Si-Si two-body interactions, we built a $3\times3\times3$ bcc cell model of pure Fe and relaxed it. We then substitute two Fe atoms in the system for Si atoms and calculate the energy according to Eq. (B1). We considered five cases divided according to the neighboring relationship between the Si atoms, ranging from first to fifth neighbor

TABLE V. Si-Si interaction energy calculated by DFT.

Neighboring relation	Distance r (Å)	U_{int} (eV)
First neighbor	2.473	0.5046
Second neighbor	2.855	0.1565
Third neighbor	4.038	−0.04238
Fourth neighbor	4.735	−0.004616
Fifth neighbor	4.946	0.03930

TABLE VI. Si-Si interaction hybrid potential parameters.

Parameter	Value
ϵ	0.000 053 985 7 (eV)
σ	6.133 74 (Å)
γ_{rep}	6.748 56
γ_{att}	7.1167
a	−0.038 174 5 (eV Å)
b	0.279 857 (Å)
c	4.151 77 (Å)

positioning:

$$U_{\text{int}} = U_{2\text{Si}} + U_{0\text{Si}} - 2U_{1\text{Si}}, \tag{B1}$$

where U_{int} denotes the energy of the two-body interaction for each of the five cases considered; $U_{0\text{Si}}$ and $U_{2\text{Si}}$, respectively, denote the energy of the system before and after substituting two Fe atoms for Si atoms; and $U_{1\text{Si}}$ denotes the energy of the system after substituting a single Fe atom for a Si atom. The values of U_{int} obtained for each case are shown in Table V.

The energy values from first to fourth neighbor interactions were fit to the following function $f(r)$ of the distance between the Si atoms r shown in Eq. (B2). The fitting parameters are shown in Table VI. Notice that $f(r)$ is a combination of Mie-type and Gauss-type potentials. Additionally, the cutoff radius of this potential was set to 5 Å:

$$f(r) = C\epsilon\left[\left(\frac{\sigma}{r}\right)^{\gamma_{\text{rep}}} - \left(\frac{\sigma}{r}\right)^{\gamma_{\text{att}}}\right] + \frac{a}{b\sqrt{2\pi}}\exp\left[-\frac{(r-c)^2}{2b^2}\right],$$
$$C = \left(\frac{\gamma_{\text{rep}}}{\gamma_{\text{rep}} - \gamma_{\text{att}}}\right)\left(\frac{\gamma_{\text{rep}}}{\gamma_{\text{att}}}\right)^{\frac{\gamma_{\text{att}}}{\gamma_{\text{rep}}-\gamma_{\text{att}}}}. \tag{B2}$$

[1] D. Frenkel and B. Smit, *Understanding Molecular Simulation From Algorithms to Applications* (Academic, New York, 2002).

[2] A. van de Walle, M. Asta, and G. Ceder, The alloy theoretic automated toolkit: A user guide, Calphad **26**, 539 (2002).

[3] P. Erhart, A. Caro, M. S. de Caro, and B. Sadigh, Short-range order and precipitation in Fe-rich Fe-Cr alloys: Atomistic off-lattice Monte Carlo simulations, Phys. Rev. B **77**, 134206 (2008).

[4] A. R. Natarajan, E. L. S. Solomon, B. Puchala, E. A. Marquis, and A. V. der Ven, On the early stages of precipitation in dilute Mg-Nd alloys, Acta Mater. **108**, 367 (2016).

[5] Z. Liu, Y. Lei, C. Gray, and G. Wang, Examination of solid-solution phase formation rules for high entropy alloys from atomistic Monte Carlo simulations, JOM **67**, 2364 (2015).

[6] M. Byshkin and M. Hou, Phase transformation and segregation in Fe-Ni alloys and nanoalloys, J. Mater. Sci. **47**, 5784 (2012).

[7] F. M. Marquez, C. Cienfuegos, B. K. Pongsai, M. Y. Lavrentiev, N. L. Allan, J. A. Purton, and G. D. Barrera, Semigrand-canonical ensemble simulations of the phase diagrams of alloys, Modell. Simul. Mater. Sci. Eng. **11**, 115 (2003).

[8] T. Frolov, S. V. Divinski, M. Asta, and Y. Mishin, Effect of Interface Phase Transformations on Diffusion and Segregation in High-Angle Grain Boundaries, Phys. Rev. Lett. **110**, 255502 (2013).

[9] M. Laradji, D. P. Landau, and B. Dünweg, Structural properties of $Si_{1-x}Ge_x$ alloys: A Monte Carlo simulation with the Stillinger-Weber potential, Phys. Rev. B **51**, 4894 (1995).

[10] B. Sadigh, P. Erhart, A. Stukowski, A. Caro, E. Martinez, and L. Zepeda-Ruiz, Scalable parallel Monte Carlo algorithm for atomistic simulations of precipitation in alloys, Phys. Rev. B **85**, 184203 (2012).

[11] D. Utt, A. Stukowski, and K. Albe, Grain boundary structure and mobility in high-entropy alloys: A comparative molecular dynamics study on a $\sigma 11$ symmetrical tilt grain boundary in face-centered cubic CuNiCoFe, Acta Mater. **186**, 11 (2020).

[12] Q.-J. Li, H. Sheng, and E. Ma, Strengthening in multi-principal element alloys with local-chemical-order roughened dislocation pathways, Nat. Commun. **10**, 3563 (2019).

[13] Z. Pan and T. J. Rupert, Effect of grain boundary character on segregation-induced structural transitions, Phys. Rev. B **93**, 134113 (2016).

[14] J. Schäfer and K. Albe, Competing deformation mechanisms in nanocrystalline metals and alloys: Coupled motion versus grain boundary sliding, Acta Mater. **60**, 6076 (2012).

[15] G. P. P. Pun, V. Yamakov, and Y. Mishin, Interatomic potential for the ternary Ni-Al-Co system and application to atomistic modeling of the B2-$L1_0$ martensitic transformation, Modell. Simul. Mater. Sci. Eng. **23**, 065006 (2015).

[16] T. Frolov, M. Asta, and Y. Mishin, Segregation-induced phase transformations in grain boundaries, Phys. Rev. B **92**, 020103(R) (2015).

[17] A. Z. Panagiotopoulos, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, Mol. Phys. **61**, 813 (1987).

[18] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, J. Comput. Phys. **117**, 1 (1995).

[19] http://lammps.sandia.gov.

[20] F. H. Stillinger and T. A. Weber, Dynamics of structural transitions in liquids, Phys. Rev. A **28**, 2408 (1983).

[21] A. Manzoor, S. Shubham, S.R. Phillpot, and D. S. Aidhy, Entropy contributions to phase stability in binary random solid solutions, npj Comput. Mater. **4**, 47 (2018).

[22] J.W. Doak, C. Wolverton, and V. Ozolins, Vibrational contributions to the phase stability of PbS-PbTe alloys, Phys. Rev. B **92**, 174306 (2015).

[23] S. Shinzato, M. Wakeda, and S. Ogata, An atomistically informed kinetic Monte Carlo model for predicting solid solution strengthening of body-centered cubic alloys, Int. J. Plast. **122**, 319 (2019).

[24] P. Hirel, Atomsk: A tool for manipulating and converting atomic data files, Comput. Phys. Commun. **197**, 212 (2015).

[25] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—The Open Visualization Tool, Modell. Simul. Mater. Sci. Eng. **18**, 015012 (2009).

[26] D. de Fontaine, The number of independent pair-correlation functions in multicomponent systems, J. Appl. Crystallogr. **4**, 15 (1971).

[27] H. Huang, X. Li, Z. Dong, W. Li, S. Huang, D. Meng, X. Lai, T. Liu, S. Zhu, and L. Vitos, Critical stress for twinning nucleation in crconi-based medium and high entropy alloys, Acta Mater. **149**, 388 (2018).

[28] S. F. Liu, Y. Wu, H. T. Wang, W. T. Lin, Y. Shang, J. B. Liu, K. An, X. J. Liu, H. Wang, and Z. P. Lu, Transformation-reinforced high-entropy alloys with superior mechanical properties via tailoring stacking fault energy, J. Alloys Compd. **792**, 444 (2019).

[29] Y. Ikeda, F. Körmann, I. Tanaka, and J. Neugebauer, Impact of chemical fluctuations on stacking fault energies of CrCoNi and CrMnFeCoNi high entropy alloys from first principles, Entropy **20**, 655 (2018).

[30] S. Zhao, G. M. Stocks, and Y. Zhang, Stacking fault energies of face-centered cubic concentrated solid solution alloys, Acta Mater. **134**, 334 (2017).

[31] G. Laplanche, A. Kostka, C. Reinhart, J. Hunfeld, G. Eggeler, and E. P. George, Reasons for the superior mechanical properties of medium-entropy CrCoNi compared to high-entropy CrMnFeCoNi, Acta Mater. **128**, 292 (2017).

[32] Z. Zhang, H. Sheng, Z. Wang, B. Gludovatz, Z. Zhang, E. P. George, Q. Yu, S. X. Mao, and R. O. Ritchie, Dislocation mechanisms and 3D twin architectures generate exceptional strength-ductility-toughness combination in CrCoNi medium-entropy alloy, Nat. Commun. **8**, 14390 (2017).

[33] J. Ding, Q. Yu, M. Asta, and R. O. Ritchie, Tunable stacking fault energies by tailoring local chemical order in CrCoNi medium-entropy alloys, Proc. Natl. Acad. Sci. USA **115**, 8919 (2018).