POLYPUBLIE
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

Titre: Title: The activation-relaxation technique: ART nouveau and kinetic ART

Auteurs: Authors: Normand Mousseau, Laurent Karim Béland, Peter Brommer, Jean-François Joly, Fedwa El-Mellouhi, Eduardo Machado-Charry, Mihai-Cosmin Marinica, & Pascal Pochet

Date: 2012

Type: Article de revue / Article

Référence: Citation: Mousseau, N., Béland, L. K., Brommer, P., Joly, J.-F., El-Mellouhi, F., Machado-Charry, E., Marinica, M.-C., & Pochet, P. (2012). The activation-relaxation technique: ART nouveau and kinetic ART. Journal of Atomic, Molecular, and Optical Physics, 2012, 1-14. https://doi.org/10.1155/2012/925278

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: PolyPublie URL: https://publications.polymtl.ca/5039/

Version: Version officielle de l'éditeur / Published version
Révisé par les pairs / Refereed

Conditions d'utilisation: Terms of Use: CC BY

Document publié chez l'éditeur officiel
Document issued by the official publisher

Titre de la revue: Journal Title: Journal of Atomic, Molecular, and Optical Physics (vol. 2012)

Maison d'édition: Publisher: Hindawi

URL officiel: Official URL: https://doi.org/10.1155/2012/925278

Mention légale: Legal notice:

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal

https://publications.polymtl.ca

Hindawi Publishing Corporation
Journal of Atomic, Molecular, and Optical Physics
Volume 2012, Article ID 925278, 14 pages
doi:10.1155/2012/925278

Review Article

The Activation-Relaxation Technique: ART Nouveau and Kinetic ART

Normand Mousseau,[1] Laurent Karim Béland,[1] Peter Brommer,[1] Jean-François Joly,[1] Fedwa El-Mellouhi,[2] Eduardo Machado-Charry,[3,4] Mihai-Cosmin Marinica,[5] and Pascal Pochet[4]

[1] *Département de physique and RQMP, Université de Montréal, P.O. Box 6128, Succursale Centre-Ville, Montréal, QC, Canada H3C 3J7*
[2] *Science Program, Texas A&M at Qatar, Texas A&M Engineering Building, Education City, Doha, Qatar*
[3] *Nanosciences Foundation, 23 Rue des Martyrs, 38000 Grenoble, France*
[4] *Laboratoire de Simulation Atomistique (L_Sim), SP2M, UMR-E CEA/UJF-Grenoble 1, INAC, 38054 Grenoble, France*
[5] *CEA, DEN, Service de Recherches de Métallurgie Physique, 91191 Gif-sur-Yvette, France*

Correspondence should be addressed to Normand Mousseau, normand.mousseau@umontreal.ca

Received 6 December 2011; Accepted 17 January 2012

Academic Editor: Jan Petter Hansen

Copyright © 2012 Normand Mousseau et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The evolution of many systems is dominated by rare activated events that occur on timescale ranging from nanoseconds to the hour or more. For such systems, simulations must leave aside the full thermal description to focus specifically on mechanisms that generate a configurational change. We present here the activation relaxation technique (ART), an open-ended saddle point search algorithm, and a series of recent improvements to ART nouveau and kinetic ART, an ART-based on-the-fly off-lattice self-learning kinetic Monte Carlo method.

1. Introduction

There has been considerable interest, in the last two decades, in the development of accelerated numerical methods for sampling the energy landscape of complex materials. The goal is to understand the long-time kinetics of chemical reactions, self-assembly, defect diffusion, and so forth associated with high-dimensional systems counting many tens to many thousands of atoms.

Some of these methods are extension of molecular dynamics (MD), such as Voter's hyperdynamics [1, 2], which provides an accelerated scheme that incorporates directly thermal effects, Laio and Parrinello's metadynamics [3], an ill-named but successful algorithm that focuses on computing free-energy barriers for specific mechanisms, and basin-hopping by Wales and Doye [4] and Goedecker [5]. Most approaches, however, select to follow transition state theory and treat the thermal contributions in a quasi-harmonic fashion, focusing therefore on the search for transition states and the computation of energy barriers. A number of these methods require the knowledge of both the initial and final states. This is the case, for example, for the nudged-elastic band method [6] and the growing-string method [7]. In complex systems, such approaches are of very limited application to explore the energy landscapes, as by definition few states are known. In these situations, open-ended methods, which more or less follow the prescription first proposed by Cerjan and Miller [8] and Simons et al. [9, 10] for low-dimensional systems, are preferred. Examples are the activation-relaxation technique (ART nouveau) [11–13], which is presented here, but also the eigenvector-following method [14], a hybrid version [15], and the similar dimer method [16].

Once a method is available for finding saddle points and adjacent minima, we must decide what to do with this information. A simple approach is to sample these minima and saddle points and classify the type of events that can take place, providing basic information on the possible evolution of these systems. This approach has been applied, using ART nouveau or other open-ended methods, on a number of

materials ranging from Lennard-Jones clusters to amorphous silicon and proteins [11, 13, 17–30]. A more structured analysis of the same information, the disconnectivity tree [31], popularized by Wales and collaborators, allows one to also extract some large-scale structure for the energy landscape and classify the overall kinetics of particular systems based on the generated events [32–36]. These open-ended methods can also be coupled with various accept/reject schemes, such as the Metropolis algorithm, for finding energy minima of clusters and molecules such as proteins and bulk systems [11, 20, 37, 38].

In the long run however, most researchers are interested in understanding the underlying kinetics controlling these complex systems. To this end, it is no longer sufficient to collect these saddle points: they must be ordered and connected in some fashion to reconstruct at least a reduced representation of the energy landscape. Wales and collaborators used a different approach: starting from a large catalog of events, they construct a connectivity matrix that links together minima through saddle points, and they apply a master equation to solve the kinetics [39]. While the discrete path sampling method has the advantage of providing a complete solution to the system's kinetics, the matrix increases rapidly with the system's complexity, making it difficult to address the kinetics of large and complex problems. It has nevertheless been applied with success to describe protein folding of a number of sequences [40–42].

A more straightforward approach to generate kinetics with these methods is to apply an on-the-fly kinetic Monte Carlo procedure: in a given local minimum, a finite number of open-ended event searches are launched and the resulting barriers are used to estimate a probable timescale over which an event will take place. This approach was applied to a number of systems using various open-ended methods such as the dimer [43–45], the hybrid eigenvector-following [46], and the autonomous basin climbing methods [47]. These approaches are formally correct, if an extensive sampling of activated events is performed before each kinetic Monte Carlo (KMC) step. They are wasteful, however, as unaffected events are not recycled, increasing rapidly in cost as the number of possible moves increases. This is why there have been considerable efforts in the last few years to implement cataloging for these events, which most of the time involve off-lattice positions. The kinetic activation-relaxation technique (kinetic ART), introduced a few years ago, proposes to handle these off-lattice positions through the use of a topological catalog, which allows a discretization of events even for disordered systems while ensuring that all relevant barriers are relaxed in order to incorporate fully long-range elastic deformations at each step [48, 49]. The need for such a method is very strong, as is confirmed by the multiplications of other algorithms that also address various limitations of standard KMC published recently, including the self-learning KMC [50], the self-evolving atomistic KMC [51], and the local-environment KMC [52].

While we are fully aware of other similar algorithms, we focus in this paper on the methods developed by our group: ART nouveau and kinetic ART. First, we present the most recent implementation of ART nouveau, which has been extensively optimized and is now faster than most comparable methods [30], and two applications: to iron and to protein folding. In the second section, we discuss kinetic ART, which, at least formally, is the most flexible atomistic KMC algorithm currently available, and show how it can be applied to complex materials including amorphous systems.

2. ART Nouveau

The ART nouveau method is an open-ended algorithm for first-order saddle point search. This algorithm has been developed over the last 15 years. Based on the activation-relaxation technique proposed in 1996 [11], it incorporates an exact search for a first-order saddle point in line with the eigenvector-following and the dimer method, through the use of the Lanczos algorithm [13]. This method has been tested and characterized in a recent paper by Marinica and collaborators [29] and updated by Machado-Charry and collaborators [30], decreasing significantly its numerical cost.

ART nouveau proceeds in three steps.

(1) *Leaving the Harmonic Well.* Starting from a local-energy minimum, the configuration is deformed in order to identify a direction of negative curvature on the energy landscape indicative of the presence of a nearby first-order saddle point.

(2) *Convergence to a First-Order Saddle Point.* Once a direction of negative curvature is identified, the system is pushed along this direction and away from the initial minimum, while minimizing the energy in the $3N - 1$ perpendicular hyperplane. When successful, this step ensures that the point reached is indeed a first-order saddle point.

(3) *Relaxation into a New Minimum.* By definition, a first-order saddle point connects two local minima by a minimum-energy path. To find the new minimum, we push the configuration over the transition state and relax the total energy.

These three steps form an *event*, a set of three related configurations: the initial minimum, the transition state, and the final minimum. As described here in after, the event can be used to characterize the local energy landscape, within a Metropolis scheme or to generate a kinetic trajectory using a kinetic Monte Carlo approach, as all points generated are fully connected.

After providing this short overview, it is useful to describe with some details the three basic steps defining ART nouveau.

2.1. Leaving the Harmonic Well. The selection of a direction for leaving the harmonic well is crucial for finding rapidly a characteristic distribution of saddle points surrounding the initial minimum. This harmonic well is described as the region of the energy landscape surrounding a local minimum with only positive curvature.

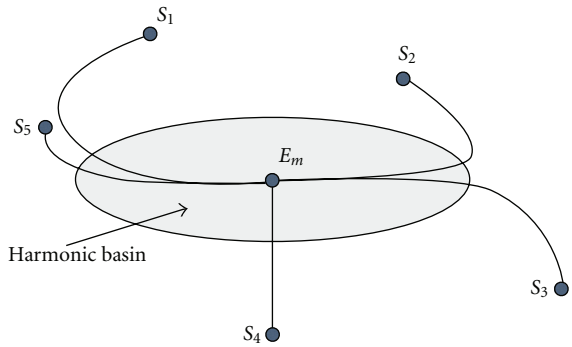


Figure 1: Sketch of minimum-energy (steepest descent) pathways from various first-order saddle points to the local energy minimum (E_m). The harmonic basin is the region within the oval surrounding the energy minimum. Except for geometric constraints, all pathways enter the harmonic basin following the lowest mode of vibration.

It is, therefore, tempting to use the eigenvectors of the Hessian matrix, corresponding to the second derivative of the energy:

$$H_{ij} = \frac{\partial^2 E}{\partial x_i \partial x_j}, \tag{1}$$

where $i, j = \{1, \ldots, 3N\}$, computed at the local minimum as a reasonable orthogonal set of sampling directions. However, this matrix does not contain any information regarding the position of nearby saddle points as the steepest-descent pathway towards this basin from each of these transition states (except for symmetric constraints) becomes rapidly parallel to one of the two directions of lowest, but positive, curvature as it moves into the harmonic basin (Figure 1).

In the absence of a discriminating basis in the selection of a given direction for launching the saddle-point search, it is best to simply use randomly chosen directions of deformation. While global random deformations are appropriate, numerous tests have shown that these lead to an oversampling of the most favorable events, generally associated with low-energy barriers, making the sampling of the energy landscape relatively costly [12, 29]. The approach chosen in ART nouveau is therefore to deform randomly select regions of the configuration, allowing nevertheless the full system to react to this deformation, in order to limit the building of strain. With this scheme, the activation can also be focused on particular subsets of a system, when needed. For example, if we are interested in the diffusion of a few defects in a crystal, it is appropriate to select, with a heavy preference, atoms near these defects for the initial deformation. Events found with both approaches are the same, but sampling is significantly more efficient with the localized procedure [29].

More precisely, the procedure implemented in ART nouveau for leaving the harmonic well is the following:

(1) an atom is first selected at random from a predefined set of regions, that can include the whole system but can be limited to, for example, the surface of a slab or a region containing defects;

(2) this atom and its neighbors, identified using a cut-off distance that can be limited to the first-shell neighbors or run through many layers, are then moved iteratively along a random direction;

(3) at every step, a slight minimisation in the perpendicular hyperplane is applied to avoid collisions and allow the full system to react to this deformation; this minimisation is not complete, as this would bring back the system onto the direction of lowest curvature of the harmonic well, and it is selected to be just sufficient to avoid unphysical conformations;

(4) at every step, we use a Lanczos procedure [54] to evaluate the lowest eigenvalue of the Hessian matrix; when this eigenvalue falls below a certain negative threshold, the system is considered to have left the harmonic well and we move to the activation regime. The negative threshold is selected to ensure that the negative eigenvalue does not vanish after a few minimisation steps in the activation regime and depends on the details of the forcefield.

2.2. Converging to a First-Order Saddle Point. Once a direction of negative curvature has been identified, it is relatively straightforward to bring the system at a first-order saddle point: the system has to be pushed along this direction while the energy is minimised in the perpendicular hyperplane. This ensures that the system will reach a first-order transition state as the force vanishes.

While straightforward, the activation can be very costly as it requires a partial knowledge of the Hessian matrix. If, for a system counting a few atoms, it is appropriate to simply compute and diagonalize the Hessian, this approach is not feasible with larger problems. To avoid this expensive calculation, a number of algorithms have been proposed for this search, including approximate projections [11], eigenvector-following [14], and hybrid eigenvector-following methods [15] such as the dimer scheme [16]. Since only the lowest direction of curvature is required, the Lanczos algorithm [54] has the advantage of building on the previous results to identify the next lowest eigendirection, decreasing considerably the computational cost of following a direction of negative curvature to a transition state.

Irrespective of system size, we find that a 15×15 Lanczos matrix is sufficient to ensure a stable activation. Since building this matrix requires forces differences, this means that 16 force evaluations, in a second-order approximation, are necessary when counting the reference position. It is possible to decrease the matrix size to as little as 4×4 [30]. In this case, the stability of Lanczos' solution is ascertained at each step by comparing the newly found direction of lowest curvature with the previous step. If the cosine between the two vectors is less than a set criterion, the Lanzcos procedure is repeated until full convergence of the eigenvector. Depending on the system, this approach can half the average number of force evaluations needed for this procedure [30].

Near the saddle point, the dual approach of activation with Lanczos and minimization in the perpendicular hyperplane with a different algorithm is not optimal. At this point,

we find it often preferable to use an integrated algorithm that can converge efficiently on inflection points. As described in [30], we have implemented recently the direct-inversion in iterative-subspace (DIIS) method [55, 56] into ART nouveau to accelerate and improve convergence. Interestingly, we have found in recent work that DIIS cannot be applied when the landscape is too rough, such as at a Si(100) surface described with an *ab initio* method, and other approaches are preferable [57]. For the appropriate systems, however, DIIS can significantly decrease the computational costs of the activation phase.

In summary, the activation and convergence to a first-order saddle point phase can be described with the following steps:

(1) using the Lanczos algorithm, the direction of negative curvature is obtained;

(2) the system is pushed slightly along this direction, with a displacement decreasing as a square root of the number of iterations, to facilitate convergence onto the saddle point;

(3) the energy is relaxed in the perpendicular hyperplane; in the first iterations after leaving the harmonic well, only a few minimization steps are taken, to avoid going back into the harmonic well and losing the direction of negative curvature;

(4) if DIIS is not used, the first three steps are repeated until the total force falls below a predefined threshold, indicating that a first-order saddle point has been reached or until the lowest eigenvalue becomes positive, indicating that the system has found its way back into the harmonic well;

(5) if DIIS is applied near the transition state, we apply the Lanczos algorithm until the negative eigenvalue has reached a minimum and has gone up for 4 sequential iterations (this criterion and the DIIS implementation are discussed in [30]); DIIS is then launched and applied until the convergence criterion is reached; a final Lanczos calculation is then performed to ensure that the point is indeed a first-order saddle point.

2.3. Relaxation into a New Minimum. Once the configuration has reached a saddle point, it is necessary to nudge slightly over to allow it to converge into a new minimum. There is clearly some flexibility here. In ART nouveau, the configuration is generally pushed over a distance of 0.15 times the distance between the saddle and the initial minimum along the eigenvector and away from the initial minimum. The system is then brought into an adjacent local energy minimum using FIRE [58], although any controlled minimisation algorithm that ensures convergence to the nearest minimum can be applied.

2.4. Characterization and Comparison. ART nouveau has been applied with success to a range of problems from Lennard-Jones clusters to proteins and amorphous silicon [11, 13, 17, 20, 21, 23, 26–30]. Its efficiency has been compared previously with a number of other open-ended and two-ended saddle-point searching methods. Olsen et al. [59] compared the efficiency of various methods as a function of the dimensionality of the studied system and found that if, for low-dimensional problems, exact methods using the knowledge of the full Hessian were preferable, those that focus solely on the direction of lowest-curvature, such as ART nouveau and the dimer method, provide a more complete set of saddle points for large-dimensional system at a much lower computational cost.

Olsen et al. also noted that for the same method, the number of required force evaluations to reach a saddle point grows with the effective dimensionality of the event taking place. Looking at Pt diffusion on a Pt(111) surface, they found that, as more atoms were allowed to move, the number of force evaluations for their ART nouveau implementation increased from 145, when a single atom was free, to 2163 with 175 free atoms [59].

In a recent paper, Machado-Charry et al. [30] compared the accelerated ART nouveau algorithm with various saddle point searching algorithms proposed in the last few years, finding a similar relation between the effective dimensionality of diffusion events and the average computational effort associated with finding a first-order saddle point. We reproduce in Table 1 the summary of the analysis in [30]. We see that the latest implementation of ART nouveau, which is many times faster than the original version [13], is extremely competitive with other saddle-searching methods, even double-ended algorithms such as the nudged-elastic band (NEB) [6] and the growing string methods [7].

Table 1 also shows a few interesting details. First, saddle points are found slightly faster using empirical rather than *ab initio* forces. This is likely due to the finite precision of the forces in quantum calculations that make the calculation of the lowest-eigenvalue and eigenvectors less stable. As already observed by Olsen et al. [59], the effective dimensionality of a system also impacts on the number of force evaluations needed to find a transition state. This effective dimensionality is related to the number of atoms that can move during an event. This is why the effective dimensionality of defect diffusion in the bulk can be lower than that at the surface (Table 1) and the computational effort for finding a saddle point does not diverge with system size, at least away from a phase transition.

2.5. Application to Iron and Protein A. ART nouveau can be applied to sample the energy landscape of complex systems and search for minimum-energy states as can be seen in the two applications presented here, to interstitial self-diffusion in iron and protein folding, both described with semiempirical potentials.

2.5.1. Interstitial Self-Diffusion in Iron. The ART nouveau method using an empirical potential is applied to the systematic search in the energy landscape of small self-interstitial clusters in iron from monointerstitial, I_1, to quadri-interstitial I_4 [29]. The explored phase space was the basin around the standard parallel configuration. The goal is to explore binding states corresponding to clustered interstitials, and not states made by separated defects. These states are not very far in energy from the

Table 1: Efficiency of various open-ended and double-ended saddle-point searching algorithms. The first four columns show results on four different systems using the latest version of ART nouveau as described by Machado et al. [30]. The next four columns are results taken from Olsen et al. [59] for their own implementation of ART nouveau and the dimer method. Two columns present results using improved version of the dimer method coupled to *ab initio* force calculations for small organic molecules [60, 61]. The last column presents results for the growing string method (GSM), a double-ended saddle search algorithm [62]. These systems are characterized as a function of their boundary condition—bulk, surface, isolated, or solution—and the effective number of degrees of freedom of each system. A comparison is made on the average number of force evaluations necessary to go from local minimum to a nearby saddle point for a successful search, $\langle f \rangle$. More important, when evaluating the efficiency of the method, is the number of force evaluations required to find a first-order saddle point, taking into account the lost events, $\langle f \rangle_s$. [1]Effective number of degrees of freedom. [2]Starting near the saddle point and not from a minimum. [3]Approximate number extracted from Figure 4 of [61] starting near the saddle point, not from a minimum.

Algo.	ART nouveau				ART*n* (Olsen)		Dimer method		Improved dimer		GSM
Ref.	Machado et al. [30]				Olsen et al. [59]				[60]	[61]	[62]
System	*a*-Si	V_{Si}	C_{20}	SiC	Pt(111)	Pt(111)	Pt(111)	Pt(111)	C_6H_{10}	PHBH/H_2O	VO_x/SiO_2
BC	Bulk	Bulk	Isol.	Surf.	Surf.	Surf.	Surf.	Surf.	Isol.	Sol.	Isol.
Pot.	SW	DFT	DFT	DFT	Morse	Morse	Morse	Morse	DFT	QM/MM	DFT
Method		PBE	LDA	PBE					B3LYP	AM1	B3LYP
DOF	3000	12[1]	60	222[1]	3	525	3	525	48	144	12[1]
$\langle f \rangle$	235	210	322	262	145	372	204	335	384[2]	425[3]	330
$\langle f \rangle_s$	670	302	718	728	145	2163	204	2148	—	—	—

lowest parallel energy configuration, since the binding energy is of the order of 1 eV. For this reason, the phase space exploration was performed using a Metropolis algorithm and a fictitious temperature. The value of the fictitious temperature is a key parameter of the method. If it is too high, most of the time is spent exploring dissociated configurations. If it is too low, it could be the case that the simulation misses clustered configurations belonging to an energy basin separated from the initial one by a high-energy saddle point. As a compromise, the fictitious temperature was set to 1200 K. In addition, when the current configuration settles at an energy 2 eV higher than the lowest-energy configuration, it is replaced with the initial configuration; in this way the system does not escape from the basin of parallel configurations.

ART nouveau revealed a large number of distinct configurations which increase rapidly with cluster size, exceeding 400, 1100, and 1500 for I_2, I_3, and I_4, respectively. The formation energy spectra of three types of clusters have few isolated low-energy configurations lying below a quasicontinuum of states. This leads to the appearance of a quasicontinuous band of states at relatively low energy above the ground state at 0.42 eV, 0.23 eV, and 0.12 for I_2, I_3, and I_4, respectively (see Figure 2(a) for I_4). This study, effectively at zero K, provides an essential complementary approach to the molecular dynamics simulations that must be run at high temperature to allow for crossing sufficiently high barriers.

ART nouveau was also used for finding transition pathways from high-energy configurations to lower-energy configurations. We have ensured a broader exploration of different ways to escape from the starting minima by performing ART nouveau simulations with different values of the fictitious Metropolis temperature, namely, 100 K, 400 K, and 800 K. Hence, we are able to provide the complete path for the unfaulting mechanisms of the self-trapped ring configurations I_{3-4}. These configurations were recently proposed for the interstitial clusters [53]: in the I_3 case the ring configuration is made of three nonparallel $\langle 110 \rangle$ dumbbells in a $\langle 111 \rangle$ plane and, in the I_4 case, obtained from I_3^{ring} by adding a $\langle 111 \rangle$ dumbbell at the center of the ring. These configurations, being composed of nonparallel dumbbells, can be considered as *faulted* loops which cannot migrate by the simple step mechanism. However, ART nouveau was able to find a transition path for the unfaulting mechanism of the I_3^{ring} (not shown) and I_4^{ring} into the standard parallel configurations $I_3^{\langle 110 \rangle}$ and $I_4^{\langle \zeta\varsigma 0 \rangle}$ (Figure 2(b)), respectively. ART nouveau generates successful un-faulting pathways with great efficiency and determines the fastest one at low temperature, in qualitative agreement with the brute force molecular dynamics calculations performed at higher temperature [53, 63].

2.5.2. Protein A. ART nouveau can also be applied to molecules such as proteins, to characterize folding [20, 38] and aggregation [37, 64, 65]. Of course, because the algorithm does not activate the full solvent, most solvent molecules see an effective $T = 0$ temperature, and proteins must be in vacuum or in implicit solvent, to avoid freezing of the surroundings. This restriction is not as severe as a large fraction of molecular dynamical simulations are also performed with implicit solvent and reduced potentials to accelerate sampling. Moreover, by leaving aside thermal fluctuations, ordered conformations are easier to identify, providing a clearer picture of various folding and aggregation pathways.

Among others, ART nouveau was applied to study folding of protein A, a fast-folding 60-residue sequence that adopts a three-helix bundle conformation in its native state. Because of its relatively low complexity, protein A has been extensively studied to understand protein folding (e.g. [66–68]). To identify folding mechanism for the full 60-residue Y15W mutant (PDB : 1SS1), 52 trajectories were launched

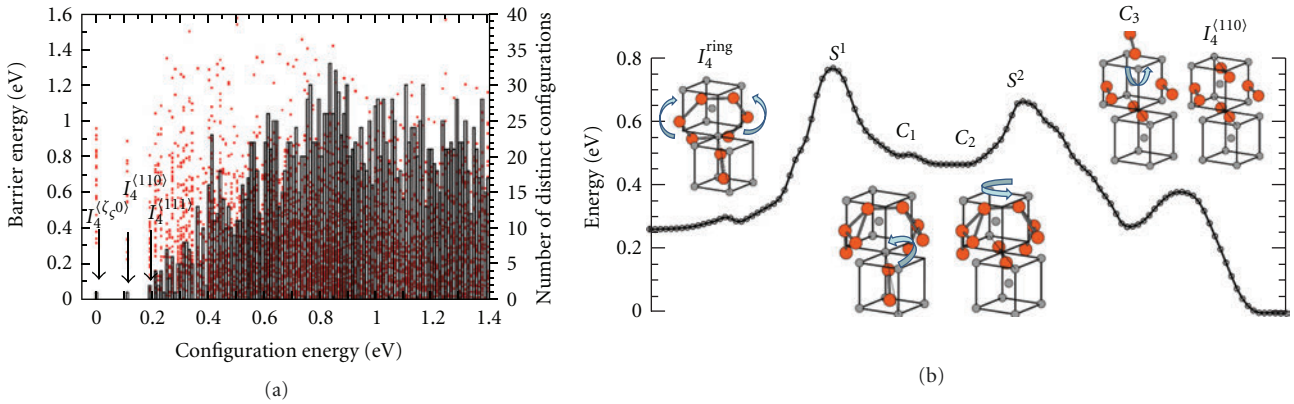


Figure 2: (a) Histogram of the relative formation energies and associated barrier energies (red squares and left vertical axis) of the ART nouveau-generated configurations for tetrainterstitials in iron modeled by the Ackland-Mendelev potential. The zero energy is the energy of the most stable configuration. The vertical axis on the right-hand side (associated with the bars) corresponds to the number of distinct minima found by ART nouveau. (b) Transition pathway for the unfaulting mechanism of the ring configuration of the tetrainterstitial, I_4^{ring}, into the most stable configuration, $I_4^{\langle\zeta\varsigma 0\rangle}$ (reproduced from [53]).

from three different initial configurations: fully extended and two random coiled conformation with right and left handedness. Simulations were run for 30 000 to 50 000 events with a Metropolis acceptance rate around 45% at a temperature of 900 K. To obtain well-relaxed conformations, each trajectory was run for a further 8000 events at 600 K. We note that, because ART nouveau does not include entropic contributions, the Metropolis temperature does not correspond directly to a physical temperature. The evolution of four folding trajectories, as characterized by the fraction of folded H1, H2, and H2 helices and the total energy, is shown in Figure 3. They show diversity in folding pathways that does not support a predetermined sequential folding. The helix formation, however, clearly occurs as the tertiary structures lock in, suggesting instead that folding of protein A is highly cooperative in agreement with the funnel picture [69].

Of the 52 trajectories, 36 adopt a structured conformation with an energy below −116 kcal/mol and belonging to one of the groups presented in Figure 4. The first two correspond to the native left-handed structure and its slightly less stable mirror image, which is also observed in other simulations [70, 71]. Configurations in Figures 4(c) and 4(d), called ϕ-shaped structures, represent intermediate states on the folding pathway as can be seen in Figure 3. The four structures on the bottom line show high β-content low-energy off-pathway conformations. These are many kcal/mol higher in energy than the native state and, clearly, have only a very low probability of forming. They can provide information useful for understanding structural differences between sequences with close homology. For example, structure in Figure 4(h) is close to 1GB1 (with an important difference in the packing of the β-hairpins). The existence of this unstable structure can explain why it is possible for the three-helix bundle to convert to protein G topology with a 59% homology between mutated sequences [72].

As a general rule, although ART nouveau does not provide kinetic information regarding trajectories, it can provide information regarding the crucial steps in relaxation, folding, and aggregation patterns for relatively large systems that are not dominated by entropic considerations. In protein A, ART nouveau managed to identify metastable ordered structures of higher energy that can shed light on close homological sequences. Finally, ART nouveau is also a very interesting tool for exploring conformation of large biological systems, if it is coupled with internal coordinates and multiscale representations [73, 74].

3. Kinetic ART

While ART nouveau is very efficient for finding saddle points, it cannot be applied directly to evaluate the dynamics of a system as the inherent bias for selecting a specific barrier over the others is not known. In the absence of such characterization, ART nouveau is ideal to sample possible events that can then serve either in a master equation or in a kinetic Monte Carlo scheme.

As mentioned in the introduction, the first approach has been attempted by Wales [39], who applied their scheme to protein folding [40–42]. Here, the main challenge is the handling of the rate matrix, which grows rapidly with the dimensionality of the system and which is very sensitive to a complete sampling of low-energy barriers. Other groups, starting with Henkelman and Jónsson, implemented a simple on-the-fly KMC approach with sampling of events after every step [43–45, 47]. This solution is time consuming, as new searches must be launched at every step, and makes little use of previously found events.

Recently, a number of kinetic Monte Carlo algorithms have been proposed to address some of these concerns by constructing a catalog: the kinetic ART [48], the self-learning KMC [50], the self-evolving atomistic KMC [51], and the local-environment KMC [75]. The challenge, for all these methods, is threefold: (1) ensuring that the most

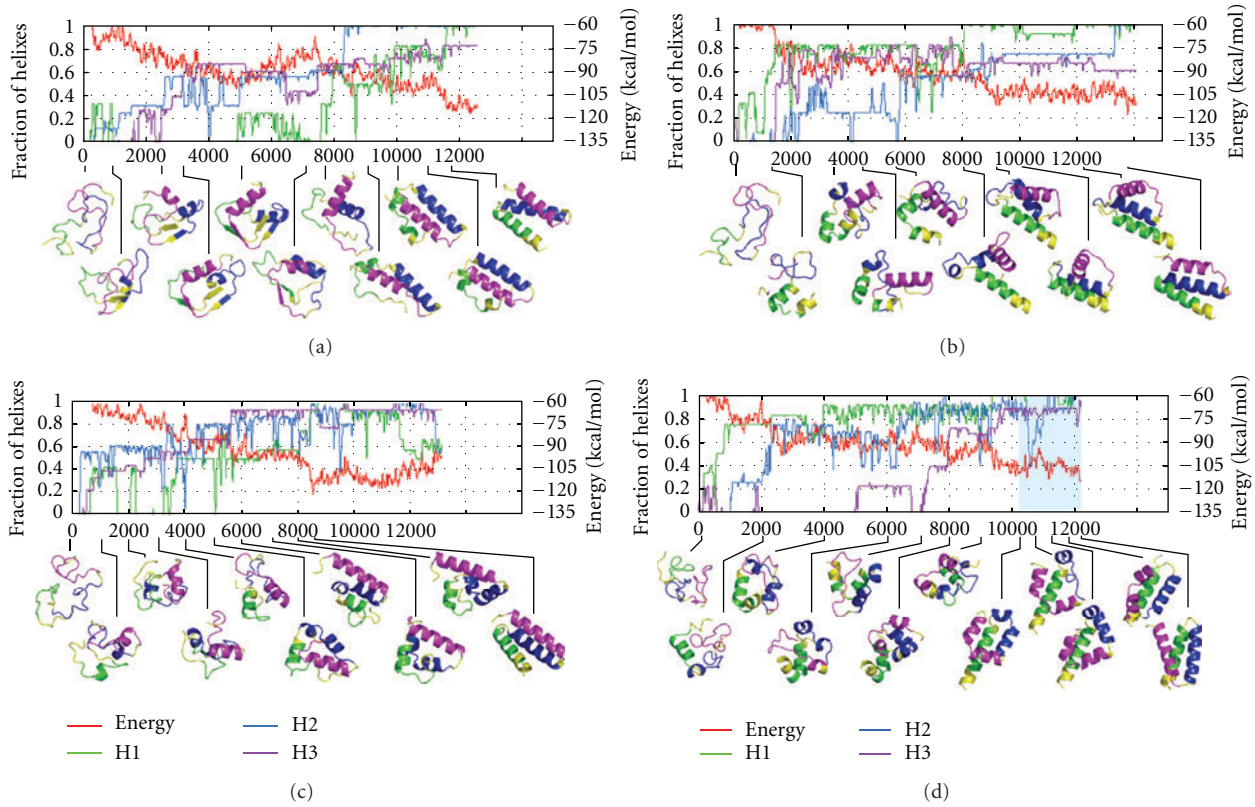


Figure 3: Evolution of trajectory #18 (a), #32 (b), #35 (c), and #25 (d). Top graphs: formation of the helical regions H1 (green), H2 (blue), and H3 (magenta) into helices and evolution of the energy level (red). The blue box of (d) indicates that a lower temperature criterion was used for that part of the simulation. All lines are smoothed by a Gaussian of width $\sigma = 0.1$ (reproduced from [38]).

relevant events are correctly included in the catalog, (2) devising a classification scheme that is efficient and can correctly discriminate events even in complex environments such as alloys and disordered materials, and (3) updating, on the fly, the energy barrier in order to fully take into account short- and long-range elastic deformations. While the various recently proposed KMC methods address some or all of these challenges, they will be rapidly evolving over the coming years. Because of this, we focus here on kinetic ART, the method we have developed and which provides one possible answer to the difficult problem of simulation long-time activated dynamics in complex environments.

Our approach attempts to minimise the computational efforts while preserving the most correct long-time kinetics, including long-range elastic effects. To do so, it is necessary to generate an event catalog that can be expanded and reused as simulations progress or new simulations are launched and, contrary to the vast majority of KMC schemes available, can handle off-lattice atomic positions. This catalog must be as compact as possible, that is, offer an efficient classification scheme, yet be sufficiently flexible to handle alloys, surfaces, and disordered environments. In an off-lattice situation, no catalog can provide precise energy barriers, as short- and long-range elastic deformations will affect the local environment. Rates, if derived from the catalog, must therefore be reevaluated at each step to include elastic effects.

This is what kinetic ART achieves by using ART nouveau for generating new events and update energy barriers, coupled with a topological classification for local environments, which allows the construction of a discrete event catalog based on a continuous geometry even in the most complex environment.

Before going into details, it is useful to present the main steps of the algorithm.

(1) Starting from a local energy-minimum, the local topology associated with each atom is analyzed. For each topology, ART nouveau searches are launched and new events are added to the catalog of events and attached to this topology. In order to ensure that the catalog is complete, event searches are not limited to new topologies: the number of searches is proportional to the logarithm of the frequency with which a topology appears in a given simulation. All events associated with the configuration at the local energy-minimum are added to the tree.

(2) All low-energy events are reconstructed and the transition state is refined to take the impact of long-range elastic deformations on the rate into account.

(3) With all barriers known, the average Poisson rate is determined according to the standard KMC

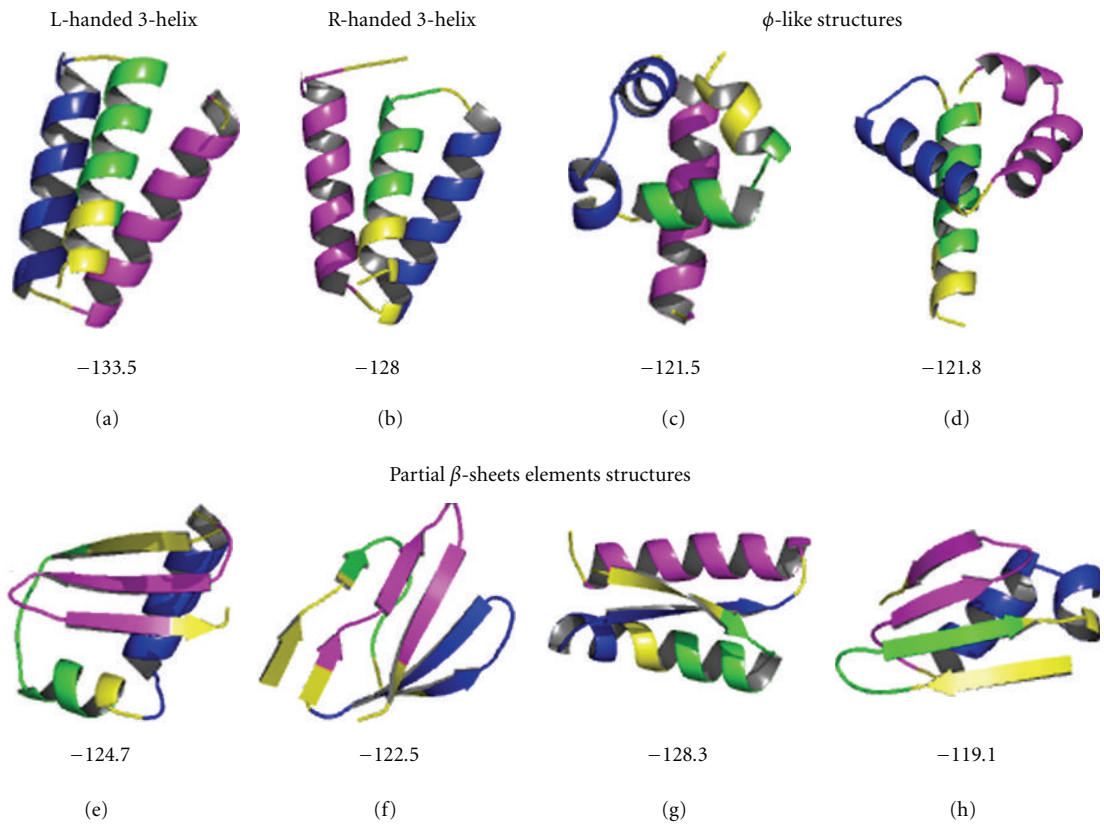


Figure 4: The structure of lowest energy found through independent simulations at 900 K followed by a energy refining simulations at 600 K. (a) The left-handed native-like bundle found in 4 simulations; (b) the right-handed mirror image found 7 times; (c) ϕ-like structures with complete H3 helix-3 simulations; (d) ϕ-like structures with complete H1 helix-4 simulations. (e) to (h) show representative structures of families of conformations with a significant β-sheet components: (e) found in 8 simulations, (f) in 3 simulations, (g) in 5 simulations, and (h) in 2 simulations. Not represented here are the 16 simulations which did not find a configuration of energy lower than −116.0 kcal/mol (reproduced from [38]).

algorithm; the clock is pushed forward and an event is selected at random, with the proper weight [76].

(4) We go back to (1).

Each of these steps involves, of course, a number of elements, which we describe in the following sections.

3.1. Topological Analysis: Constructing an Off-Lattice Event Catalog. In order to construct a catalog of events, it is necessary that these be discretized to ensure that events are recognized uniquely. For lattice-constrained motion, this requirement is straightforward to implement by focusing simply on the occupancy of the various crystalline sites. Geometry, however, is no longer a satisfactory criterion for reconstructed sites, defective and disordered systems, where atoms cannot easily be mapped back to regular lattice sites. Our solution is to use a topological description of local environments for mapping events uniquely, allowing us to describe environments of any chemical and geometrical complexity. This is done by first selecting a central atom and all its neighbors within a given cut-off radius. Atoms are then connected, following a predefined procedure that can be based loosely on first-neighbor distance or a Voronoi construction, forming a graph that is then analysed using the software NAUTY, a powerful package developed by McKay [75]. For our purposes, this package returns a unique tag associated with the topology and an ordered list of cluster atoms. The positions of the topology cluster atoms in the initial, saddle, and final states are stored.

For this discretization to work, it is necessary that a one-to-one correspondence exist between the topological description and the real-space geometry of this local region. This correspondence is enforced by the fact that the graph has to be embedded into a well-defined surrounding geometry and must correspond to either a local minimum or a first-order saddle with the given forcefield. These two elements ensure that the one-to-one correspondence is valid in the vast majority of conformations.

This correspondence may fail in two cases: for very flat minima or saddle points, there might be more than one topology associated with the geometry. In this case, the corresponding events might be overcounted. For this reason, we compare the barrier height as well as the absolute total displacement and the direction of motion of the main atom (the one that moves most during an event) both between initial and final state and between initial state and the saddle

point. We only add events to the catalog that are sufficiently different from previously found ones.

It is also possible that a single topology corresponds to more than one geometry. Most frequently, this problem appears in highly symmetric systems: two topologically identical events might differ only in the direction of motion. To resolve this issue, we reconstruct geometry and ensure that motion takes place in different directions. In this case, events are flagged as different and additional ART nouveau searches are performed to identify all topologically equivalent but geometrically different moves.

In some cases, a topology is associated with fundamentally different geometries. Here, this failure is identified automatically, since it leads to incorrect saddle-point reconstruction that cannot be converged. When this occurs, the algorithm automatically modifies the first-neighbor cutoff used for constructing the topology until the multiple geometries are separated into unique topologies. In most systems, this is extremely rare, and it is generally a symptom of an inappropriately short topology cutoff radius.

Generic and Specific Events. Once a new topology is identified, ART nouveau saddle point searches are launched a number of times from the central atom associated with this topology (to increase the efficiency of finding topologies with multiple geometries, events are launched from a randomly selected set of atoms characterized by the same topology). Depending on the system, we use between 25 and 50 searches for every $\log_{10}$ times we see a topology (e.g., 25–50 searches for a topology appearing once, 75–150 searches for a topology seen 100 times, etc.). To ensure detailed balance, these events are stored from both the initial and final minima.

These events, which are stored in the catalog, are called *generic events* and serve as the basis for reconstructing *specific events* associated with a given configuration. They can be accumulated and serve for further studies, decreasing over time the computational efforts associated with studying particular systems. For example, a catalog constructed for ion-implanted Si can be used as a starting point for amorphous silicon or an SiGe alloy (as mentioned previously, the topological analysis allows us to handle correctly alloys, by discriminating between the atomic species involved).

3.2. Reconstructing the Geometry and Refining Low-Energy Events. While representative, the transition states and energy barriers in the catalog cannot be applied exactly as they are to new configurations, as short- and long-range elastic deformations affect every barrier differently, creating favored directions even for formally isotropic defects, for example.

To include these effects, every generic event should be reconstructed for each atom and fully relaxed. This is what kinetic ART does, with one approximation: to limit the effort of refining generic into specific events, only the kinetically relevant events are reconstructed and relaxed.

After each KMC step, the event list is ordered according to the barrier energy and only the lowest-energy barrier events, representing up to given threshold (we use typically 99.9 or 99.99%) of the total rate are fully reconstructed and refined. Depending on the system and the temperature, this means that only one to ten percent of all barriers in the catalog are refined specific events. The remaining events, which contribute very little to the rate, are cloned from the generic events without adjusting the barrier.

This local reconstruction takes place in two steps. First, using the reference atomic positions of the generic event, the geometric transformation necessary to map the initial state onto the current configuration is determined. This operation is then applied to the atomic displacements between initial state and saddle point. From this first reconstruction of the saddle point geometry, the saddle point (and with it the energy barrier) is refined. In the second step, the system is relaxed into the final state by pushing it over the saddle point. If it is impossible to map the initial state onto the current configuration even though the central atoms have the same topology, we know that the one-to-one correspondence between topology and geometry has failed and apply the corrections mentioned in the previous section.

3.3. Applying the Standard KMC Procedure. The reconstruction of specific events ensures that all elastic and geometrical deformations are taken into account when constructing the list of events that is used in the kinetic Monte Carlo algorithm. This list now contains both refined low-energy specific events and clones of generic events with higher barrier energies.

Rates are determined according to transition state theory:

$$r_i = \tau_0 \exp\left(\frac{-\Delta E_i}{k_B T}\right), \tag{2}$$

where τ_0, the attempt frequency, is a system-dependent fixed constant of the order of 10^{13} s^{-1}; $\Delta E_i = E_{\text{saddle}} - E_{\text{minimum}}$ is the barrier height; k_B and T are the Boltzmann constant and the temperature, respectively. While it would be possible to extract a more precise attempt frequency from quasi-harmonic theory, we know that τ_0 varies only weakly with the chosen pathway for a large number of systems [77, 78].

These rates are then combined, following Bortz et al. [76], to extract the Poisson distributed time step over which an event will take place:

$$\Delta t = -\frac{\ln \mu}{\sum_i r_i}, \tag{3}$$

where μ is a random number in the [0,1] interval introduced to ensure a Poisson distribution and r_i is the rate associated with event i. The simulation clock is pushed forward by this Δt an event is selected with a weight proportional to its rate, and the atoms are moved accordingly. Finally, the whole system is relaxed into a minimum energy configuration, and the next step can proceed.

3.4. Other Implementation Details. The local properties of kinetic ART allow a very efficient implementation of the algorithm. First, it is straightforward to dispatch the event searching and refining to separate processors. For a typical

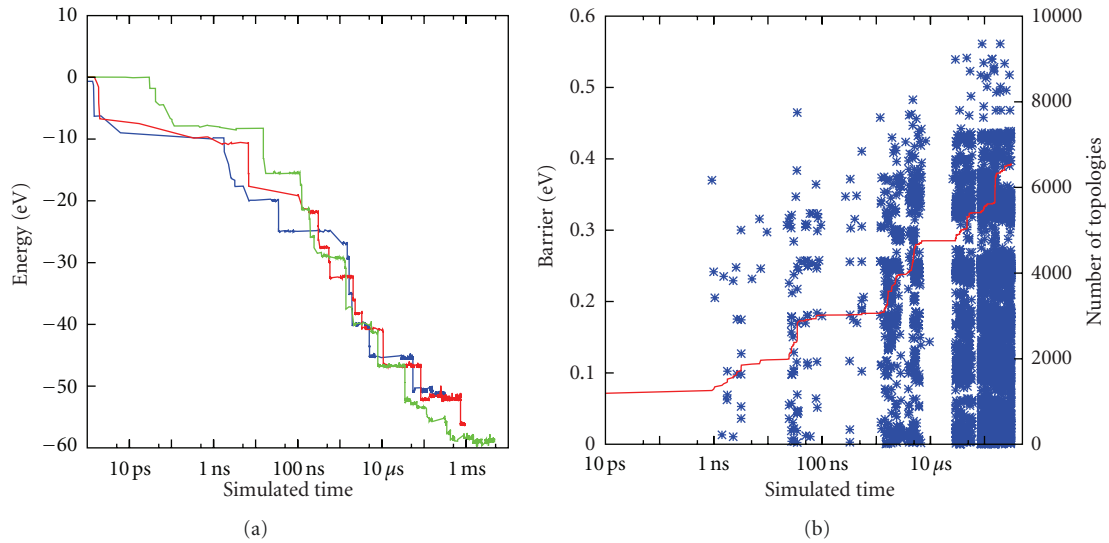


Figure 5: (a) The total potential energy of the bombarded c-Si crystal during annealing at 300 K. We show three simulations that have the same initial configuration. (b) The energy barriers crossed during a c-Si anneal at 300 K are represented by the stars. The number of topologies where we performed ART searches is shown as a continuous line.

system, between 20 and 40 processors can be used for these tasks (up to 128 when building the initial catalog), with an efficiency of 90% or more. Moreover, as events are inherently local, there is no need to compute global forces during an ART nouveau event: only nonzero forces need to be computed most of the time, with a final global minimization needed only at the end of each KMC step. This renders the algorithm for computing an event almost independent of system size for a fixed number of defects and order N overall in simulated time for a system with a constant defect density, at least with short-range potential, making it scale as efficiently as molecular dynamics in this worst case scenario.

In KMC, the residence time Δt and the selection of the next state are dominated by the lowest-energy barrier present in the system. Groups of states separated by such low-energy barriers are called basins. These pose a major obstacle to KMC simulations: basin transitions are often nondiffusive and trap the simulation in a small part of the configuration space. Additionally, the high rates also severely limit the time step, thus increasing the CPU cost without yielding meaningful physics. We developed the basin-autoconstructing mean rate method (bac-MRM) [49], based on the MRM described by Puchala et al. [79], to overcome this limitation. If an event is selected that matches the description of a basin event (small energy difference between saddle point and both initial and final state), the event is executed and then added to the current basin of events. Kinetic ART keeps all other available events in memory and adds to this list those originating from the state in which the system is now. This means that before the next KMC step, events originating from all basin states could be selected. The rates of the events are then adjusted to average over all possible numbers of intrabasin transitions. This assures that the next event and the time step are picked with the correct distribution, providing the correct kinetics even though, of course, the intrabasin trajectory information is lost. If the next event is another basin event, the basin is expanded again. This procedure is repeated, until a nonbasin event is selected (usually after all basin events have been added to the basin). The bac-MRM assures that the system will not visit any basin state twice and thus makes sure that the system can escape being trapped in configuration space.

3.5. Example: Relaxation of an Ion-Bombarded Semiconductor. We first demonstrate the capability of kinetic ART by applying it to the relaxation of a large Si box with ion implantation. In a previous work by Pothier et al. [79], a Si atom was implanted at 3 keV in a box containing 100 000 crystalline silicon (c-Si) atoms at 300 K. This system was then simulated for 900 ps with molecular dynamics and the Stillinger-Weber potential. From the final configuration of this simulation, we extracted a smaller box of 26 998 atoms which contains the vast majority of induced defects and imposed periodic boundary conditions.

Kinetic ART is applied at 300 K on this complex system. Figure 5(a) reports the evolution of three simulations with the same initial configuration over 3000 events, each reaching simulated time of over 1 ms, at a cost of 8000 CPU-hour. During this time, the system relaxes by nearly 60 eV as numerous defects crystallize. Furthermore, we see that the potential energy drops are not all identical. This indicates that a wide variety of relaxation mechanisms take place on these timescales.

Meanwhile, Figure 5(b) shows the energy barriers of the events which were executed in one of the simulations. The gaps in time are a consequence of the bac-MRM mentioned in Section 3.4. Indeed, if a basin contains some moderately high barriers, we will spend a good deal of time in it without sampling new states. This figure also shows the richness of the energetic landscape. We execute events with a great

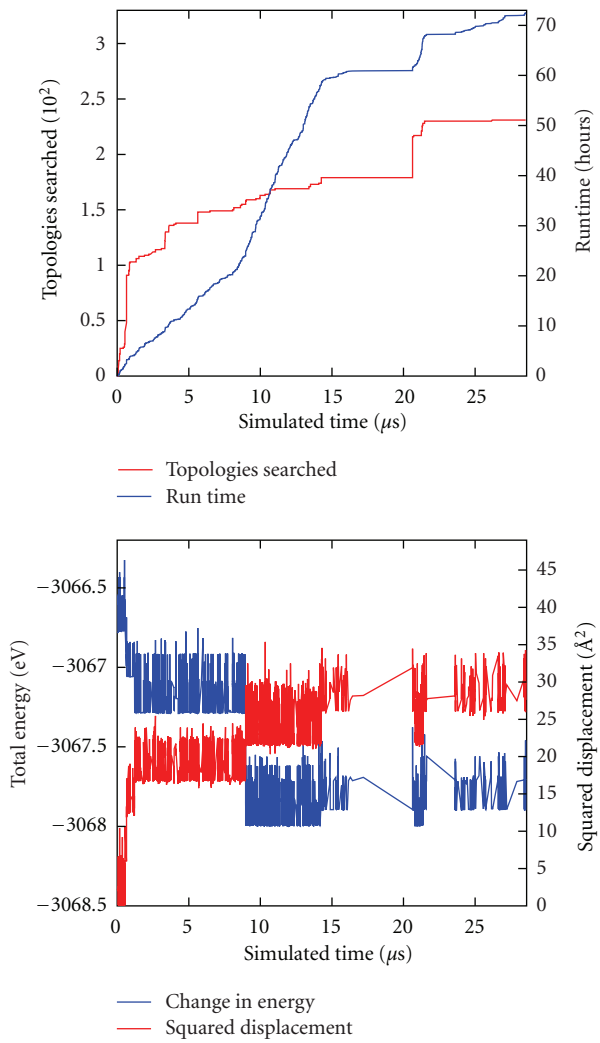


Figure 6: Simulation of an a-Si model containing 1000 atoms at 300 K. The simulation was started using a previously built event catalogue of over 87 000 events. Top: The number of new topologies searched and the total runtime as a function of simulated time. Bottom: The change in energy and the squared total displacement as a function of simulated time.

variety of energies, which show the great importance of treating elastic effects with exactitude. The great number of new topologies we encounter (even though the system lowers its energy) is another hint at the richness of this system.

While standard KMC simulations easily run for millions of steps, regularly reaching timescale of minutes or more, fundamental restrictions have limited them to simple lattice-based problems such as metal on metal growth. The possibility, with kinetic ART, to now apply these accelerated approaches to complex materials such as ion-bombarded Si, oversimulated timescale many orders of magnitude longer than what can be reached by MD, opens the door to the study in numerous systems of the long-time atomistic dynamics that was long considered out of reach. This is the case for fully disordered systems.

3.6. Example: Relaxation in Amorphous Silicon. Here, we show that even such complex materials such as amorphous silicon (a-Si), a system that has been used as a test case for ART over the years [11, 12], can be handled with kinetic ART.

Disordered systems are characterized by an extremely large number of possible configurations that exclude studies with traditional lattice-based Monte Carlo methods. Event catalogues will therefore be large, as almost each atom in a box of a few thousand atoms has its own topology. However, since the topological classification is based on the local environment, it still provides for a real gain given enough sampling.

Figure 6 shows a 28.5 μs simulation, started with a well-relaxed 1000-atom a-Si model and a preformed event catalogue of over 87 000 events at 300 K. Since the original configuration is already well explored, the simulation needs to generate only a few events every time a new topology is found, underlining the powerful capabilities of kinetic ART to generate and classify activated events in complex environments. Over the 2542 KMC steps of the simulation, only 1000 new topologies were found.

In a system such as a-Si, the continuous distribution of activated barriers means that there is not a clear separation between frequent and rare events. The energy cutoff for bac-MRM is then chosen so that the inverse of the associated rate is small compared to the desired simulated time. Since in an amorphous system flicker-like events can occur at any energy scale, the value must be adjusted as a function of the degree of relaxation and temperature. In the present case, the cutoff value was 0.35 eV, meaning that while the thermodynamics is accurate at all scales, internal dynamics on timescales lower than 120 ns is ignored. This simulation required 617 CPU-hours over six processors for a total run-time of 103 hours. As shown in Figure 6, with a well-filled catalog, only a few hundred unknown topologies are visited during the simulation.

Starting from a well-relaxed configuration, the simulation quickly reaches the μs timescales, where higher-energy flicker-like events dominate the dynamics. Still, the system finds a way to relax at three occasions. For instance, the large drop observed at 8.7 μs is the result of the system choosing an event with a large 0.51 eV barrier. Interestingly, the event involves only perfectly coordinated atoms and can be characterized as a bond-switching move where two neighboring atoms exchange neighbors, allowing an atom with a high angular strain on three of its bonds to relax near its equilibrium state.

Even though disordered systems require more computational efforts than crystalline-based defective systems, the application of kinetic ART to such systems allows the study of kinetics on timescales well beyond what could have been reached until now using more traditional molecular dynamics, as can be seen here, in a short test simulation.

4. Conclusion

We have presented the activation-relaxation technique, ART nouveau, a very efficient open-ended search method for

transition state, and have shown how it can be applied to extensively search for diffusion mechanisms and pathways as well as low-energy configurations with systems as diverse as interstitials in iron and a 60-residue protein. With recent improvements, finding a transition can take as little as 210 force evaluations and, when lost events are taken into account, between 300 and 700 force evaluations, allowing for extensive searches, with either empirical or *ab initio* forces, needed to fully describe the energy landscape of complex systems.

As the event-selection bias is unknown in ART nouveau, it cannot be used directly to study the kinetics of atomistic systems. This limitation is lifted by coupling it to a kinetic Monte Carlo scheme. Kinetic ART goes beyond this simple coupling, however, by introducing a topological classification for catalog building while fully preserving the impact of elastic deformations on the kinetics while being fully off-lattice. The efficiency of this approach was demonstrated by characterizing the relaxation of a 30 000-atom ion-bombarded box of silicon and a 1000-atom box of amorphous silicon.

Although not presented here, the algorithm also readily handles other systems such as metals and alloys. At the moment, this is done using a fixed prefactor for the attempted frequency in transition state theory. As discussed previously, such an approximation is fairly good for a number of systems. However, we are currently implementing a version of the algorithm that also evaluates prefactors. While this is more costly, it could be essential to describe correctly the kinetics of complex materials.

Both ART nouveau and kinetic ART open the door to the study of questions that were out of reach only a few years ago, either to identify diffusion mechanisms and catalytic reactions or to recreate full diffusion and relaxation pathways in complex materials such as alloys and glasses on timescale that had been out of reach until now. As discussed previously, a number of other accelerated techniques have been proposed recently and it is too early to determine whether one method will really stand out. Irrespective of this, ART nouveau and kinetic ART have already shown what they can be used for.

Acknowledgments

E. Machado-Charry, P. Pochet, and N. Mousseau acknowledge support by Nanosciences Fondation in the frame of MUSCADE Project. Partial funding was also provided by the Natural Science and Engineering Research Council of Canada, the Fonds Québécois de Recherche Nature et Technologie, and the Canada Research Chair Foundation. This work was performed using HPC resources from GENCI-CINES (Grant 2010-c2010096323) and from Calcul Québec (CQ). The ART nouveau code is freely distributed in a format appropriate for both empirical and quantum forces. Kinetic ART is still under development but a version should be available for distribution in the near future. To receive the code or for more information, please contact the authors.

References

[1] A. F. Voter, "Hyperdynamics: accelerated molecular dynamics of infrequent events," *Physical Review Letters*, vol. 78, no. 20, pp. 3908–3911, 1997.

[2] P. Tiwary and A. van de Walle, "Hybrid deterministic and stochastic approach for efficient atomistic simulations at long time scales," *Physical Review B*, vol. 84, no. 10, Article ID 100301, 2011.

[3] A. Laio and M. Parrinello, "Escaping free-energy minima," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 20, pp. 12562–12566, 2002.

[4] D. J. Wales and J. P. K. Doye, "Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms," *Journal of Physical Chemistry A*, vol. 101, no. 28, pp. 5111–5116, 1997.

[5] S. Goedecker, "Minima hopping: an efficient search method for the global minimum of the potential energy surface of complex molecular systems," *Journal of Chemical Physics*, vol. 120, no. 21, pp. 9911–9917, 2004.

[6] G. Henkelman, B. P. Uberuaga, and H. Jónsson, "Climbing image nudged elastic band method for finding saddle points and minimum energy paths," *Journal of Chemical Physics*, vol. 113, no. 22, pp. 9901–9904, 2000.

[7] A. Goodrow, A. T. Bell, and M. Head-Gordon, "Transition state-finding strategies for use with the growing string method," *Journal of Chemical Physics*, vol. 130, no. 24, Article ID 244108, 2009.

[8] C. J. Cerjan and W. H. Miller, "On finding transition states," *The Journal of Chemical Physics*, vol. 75, no. 6, pp. 2800–2801, 1981.

[9] J. Simons, P. Jørgensen, H. Taylor, and J. Ozment, "Walking on potential energy surfaces," *Journal of Physical Chemistry*, vol. 87, no. 15, pp. 2745–2753, 1983.

[10] H. Taylor and J. Simons, "Imposition of geometrical constraints on potential energy surface walking procedures," *Journal of Physical Chemistry*, vol. 89, no. 4, pp. 684–688, 1985.

[11] G. T. Barkema and N. Mousseau, "Event-based relaxation of continuous disordered systems," *Physical Review Letters*, vol. 77, no. 21, pp. 4358–4361, 1996.

[12] N. Mousseau and G. T. Barkema, "Traveling through potential energy landscapes of disordered materials: the activation-relaxation technique," *Physical Review E*, vol. 57, pp. 2419–2424, 1998.

[13] R. Malek and N. Mousseau, "Dynamics of Lennard-Jones clusters: a characterization of the activation-relaxation technique," *Physical Review E*, vol. 62, no. 6, pp. 7723–7728, 2000.

[14] J. P. K. Doye and D. J. Wales, "Surveying a potential energy surface by eigenvector-following: applications to global optimisation and the structural transformations of clusters," *Zeitschrift fur Physik D-Atoms Molecules and Clusters*, vol. 40, no. 1–4, pp. 194–197, 1997.

[15] L. J. Munro and D. J. Wales, "Defect migration in crystalline silicon," *Physical Review B*, vol. 59, no. 6, pp. 3969–3980, 1999.

[16] G. Henkelman and H. Jónsson, "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives," *Journal of Chemical Physics*, vol. 111, no. 15, pp. 7010–7022, 1999.

[17] N. Mousseau and L. J. Lewis, "Topology of amorphous tetrahedral semiconductors on intermediate length scales," *Physical Review Letters*, vol. 78, no. 8, pp. 1484–1487, 1997.

[18] Y. Kumeda, D. J. Wales, and L. J. Munro, "Transition states and rearrangement mechanisms from hybrid eigenvector-following and density functional theory.: application to C10H10 and defect migration in crystalline silicon,"

Chemical Physics Letters, vol. 341, no. 1-2, pp. 185–194, 2001.

[19] G. Henkelman and H. Jónsson, "Long time scale kinetic Monte Carlo simulations without lattice approximation and predefined event table," *Journal of Chemical Physics*, vol. 115, no. 21, pp. 9657–9666, 2001.

[20] G. Wei, P. Derreumaux, and N. Mousseau, "Sampling the complex energy landscape of a simple β-hairpin," *Journal of Chemical Physics*, vol. 119, no. 13, pp. 6403–6406, 2003.

[21] F. El-Mellouhi, N. Mousseau, and P. Ordejón, "Sampling the diffusion paths of a neutral vacancy in silicon with quantum mechanical calculations," *Physical Review B*, vol. 70, no. 20, Article ID 205202, pp. 205202–9, 2004.

[22] L. Xu, G. Henkelman, C. T. Campbell, and H. Jónsson, "Small Pd clusters, up to the tetramer at least, are highly mobile on the MgO(100) surface," *Physical Review Letters*, vol. 95, no. 14, Article ID 146103, pp. 1–4, 2005.

[23] D. Ceresoli and D. Vanderbilt, "Structural and dielectric properties of amorphous ZrO_2 and HfO_2," *Physical Review B*, vol. 74, no. 12, Article ID 125108, 2006.

[24] F. Calvo, T. V. Bogdan, V. K. De Souza, and D. J. Wales, "Equilibrium density of states and thermodynamic properties of a model glass former," *Journal of Chemical Physics*, vol. 127, no. 4, Article ID 044508, 2007.

[25] D. Mei, L. Xu, and G. Henkelman, "Dimer saddle point searches to determine the reactivity of formate on Cu(111)," *Journal of Catalysis*, vol. 258, no. 1, pp. 44–51, 2008.

[26] D. Rodney and C. Schuh, "Distribution of thermally activated plastic events in a flowing glass," *Physical Review Letters*, vol. 102, no. 23, Article ID 235503, 2009.

[27] M. Baiesi, L. Bongini, L. Casetti, and L. Tattini, "Graph theoretical analysis of the energy landscape of model polymers," *Physical Review E*, vol. 80, no. 1, Article ID 011905, 2009.

[28] H. Kallel, N. Mousseau, and F. Schiettekatte, "Evolution of the potential-energy surface of amorphous silicon," *Physical Review Letters*, vol. 105, no. 4, Article ID 045503, 2010.

[29] M.-C. Marinica, F. Willaime, and N. Mousseau, "Energy landscape of small clusters of self-interstitial dumbbells in iron," *Physical Review B*, vol. 83, no. 9, Article ID 094119, 2011.

[30] E. Machado-Charry, L. K. Béland, D. Caliste et al., "Optimized energy landscape exploration using the ab initio based activation-relaxation technique," *Journal of Chemical Physics*, vol. 135, no. 3, Article ID 034102, 2011.

[31] O. M. Becker and M. Karplus, "The topology of multidimensional potential energy surfaces: theory and application to peptide structure and kinetics," *Journal of Chemical Physics*, vol. 106, no. 4, pp. 1495–1517, 1997.

[32] M. A. Miller and D. J. Wales, "Energy landscape of a model protein," *Journal of Chemical Physics*, vol. 111, no. 14, pp. 6610–6616, 1999.

[33] D. J. Wales and T. V. Bogdan, "Potential energy and free energy landscapes," *Journal of Physical Chemistry B*, vol. 110, no. 42, pp. 20765–20776, 2006.

[34] A. T. Anghel, D. J. Wales, S. J. Jenkins, and D. A. King, "Theory of C2Hx species on Pt110 (1x2): reaction pathways for dehydrogenation," *Journal of Chemical Physics*, vol. 126, no. 4, Article ID 044710, 2007.

[35] T. James, D. J. Wales, and J. Hernández Rojas, "Energy landscapes for water clusters in a uniform electric field," *Journal of Chemical Physics*, vol. 126, no. 5, Article ID 054506, 2007.

[36] V. K. de Souza and D. J. Wales, "Connectivity in the potential energy landscape for binary Lennard-Jones systems," *Journal of Chemical Physics*, vol. 130, no. 19, Article ID 194508, 2009.

[37] S. Santini, G. Wei, N. Mousseau, and P. Derreumaux, "Pathway complexity of Alzheimer's β-amyloid Aβ16-22 peptide assembly," *Structure*, vol. 12, no. 7, pp. 1245–1255, 2004.

[38] J.-F. St-Pierre, N. Mousseau, and P. Derreumaux, "The complex folding pathways of protein A suggest a multiple-funnelled energy landscape," *Journal of Chemical Physics*, vol. 128, no. 4, Article ID 045101, 2008.

[39] D. J. Wales, "Discrete path sampling," *Molecular Physics*, vol. 100, no. 20, pp. 3285–3305, 2002.

[40] D. A. Evans and D. J. Wales, "The free energy landscape and dynamics of met-enkephalin," *Journal of Chemical Physics*, vol. 119, no. 18, pp. 9947–9955, 2003.

[41] S. A. Trygubenko and D. J. Wales, "Graph transformation method for calculating waiting times in Markov chains," *Journal of Chemical Physics*, vol. 124, no. 23, Article ID 234110, 2006.

[42] J. M. Carr and D. J. Wales, "Refined kinetic transition networks for the GB1 hairpin peptide," *Physical Chemistry Chemical Physics*, vol. 11, no. 18, pp. 3341–3354, 2009.

[43] G. Henkelman and H. Jónsson, "Theoretical calculations of dissociative adsorption of CH4 on an Ir(111) surface," *Physical Review Letters*, vol. 86, no. 4, pp. 664–667, 2001.

[44] F. Hontinfinde, A. Rapallo, and R. Ferrando, "Numerical study of growth and relaxation of small C60 nanoclusters," *Surface Science*, vol. 600, no. 5, pp. 995–1003, 2006.

[45] L. Xu, D. Mei, and G. Henkelman, "Adaptive kinetic Monte Carlo simulation of methanol decomposition on Cu(100)," *Journal of Chemical Physics*, vol. 131, no. 24, Article ID 244520, 2009.

[46] T. F. Middleton and D. J. Wales, "Comparison of kinetic Monte Carlo and molecular dynamics simulations of diffusion in a model glass former," *Journal of Chemical Physics*, vol. 120, no. 17, pp. 8134–8143, 2004.

[47] Y. Fan, A. Kushima, S. Yip, and B. Yildiz, "Mechanism of void nucleation and growth in bcc Fe: atomistic simulations at experimental time scales," *Physical Review Letters*, vol. 106, no. 12, Article ID 125501, 2011.

[48] F. El-Mellouhi, N. Mousseau, and L. J. Lewis, "Kinetic activation-relaxation technique: an off-lattice self-learning kinetic Monte Carlo algorithm," *Physical Review B*, vol. 78, no. 15, Article ID 153202, 2008.

[49] L. K. Béland, P. Brommer, F. El-Mellouhi, J. -F. Joly, and N. Mousseau, "Kinetic activation-relaxation technique," *Physical Review E*, vol. 84, no. 4, Article ID 046704, 2011.

[50] A. Kara, O. Trushin, H. Yildirim, and T. S. Rahman, "Off-lattice self-learning kinetic Monte Carlo: application to 2D cluster diffusion on the fcc(111) surface," *Journal of Physics Condensed Matter*, vol. 21, no. 8, Article ID 084213, 2009.

[51] H. Xu, Y. N. Osetsky, and R. E. Stoller, "Simulating complex atomistic processes: on-the-fly kinetic Monte Carlo scheme with selective active volumes," *Physical Review B*, vol. 84, no. 13, Article ID 132103, 2011.

[52] D. Konwar, V. J. Bhute, and A. Chatterjee, "An off-lattice, self-learning kinetic Monte Carlo method using local environments," *Journal of Chemical Physics*, vol. 135, no. 17, Article ID 174103, 2011.

[53] D. A. Terentyev, T. P. C. Klaver, P. Olsson et al., "Self-trapped interstitial-type defects in iron," *Physical Review Letters*, vol. 100, no. 14, Article ID 145503, 2008.

[54] C. Lanczos, *Applied Analysis*, Dover, New York, NY, USA, 1988.

[55] P. Pulay, "Convergence acceleration of iterative sequences. the case of SCF iteration," *Chemical Physics Letters*, vol. 73, no. 2, pp. 393–398, 1980.

[56] R. Shepard and M. Minkoff, "Some comments on the DIIS method," *Molecular Physics*, vol. 105, no. 19–22, pp. 2839–2848, 2007.

[57] E. Cancès, F. Legoll, M. C. Marinica, K. Minoukadeh, and F. Willaime, "Some improvements of the activation-relaxation technique method for finding transition pathways on potential energy surfaces," *Journal of Chemical Physics*, vol. 130, no. 11, Article ID 114711, 2009.

[58] E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch, "Structural relaxation made simple," *Physical Review Letters*, vol. 97, no. 17, Article ID 170201, 2006.

[59] R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson, and H. Jónsson, "Comparison of methods for finding saddle points without knowledge of the final states," *Journal of Chemical Physics*, vol. 121, no. 20, pp. 9776–9792, 2004.

[60] A. Heyden, A. T. Bell, and F. J. Keil, "Efficient methods for finding transition states in chemical reactions: comparison of improved dimer method and partitioned rational function optimization method," *Journal of Chemical Physics*, vol. 123, Article ID 224101, 14 pages, 2005.

[61] J. Kästner and P. Sherwood, "Superlinearly converging dimer method for transition state search," *Journal of Chemical Physics*, vol. 128, no. 1, Article ID 014106, 2008.

[62] A. Goodrow and A. T. Bell, "A theoretical investigation of the selective oxidation of methanol to formaldehyde on isolated vanadate species supported on titania," *Journal of Physical Chemistry C*, vol. 112, no. 34, pp. 13204–13214, 2008.

[63] Y. Fan, A. Kushima, and B. Yildiz, "Unfaulting mechanism of trapped self-interstitial atom clusters in bcc Fe: a kinetic study based on the potential energy landscape," *Physical Review B*, vol. 81, no. 10, Article ID 104102, 2010.

[64] A. Melquiond, G. Boucher, N. Mousseau, and P. Derreumaux, "Following the aggregation of amyloid-forming peptides by computer simulations," *The Journal of chemical physics*, vol. 122, no. 17, p. 174904, 2005.

[65] N. Mousseau and P. Derreumaux, "Exploring energy landscapes of protein folding and aggregation," *Frontiers in Bioscience*, vol. 13, no. 12, pp. 4495–4516, 2008.

[66] H. Gouda, H. Torigoe, A. Saito, M. Sato, Y. Arata, and I. Shimada, "Three-dimensional solution structure of the B domain of staphylococcal protein A: comparisons of the solution and crystal structures," *Biochemistry*, vol. 31, no. 40, pp. 9665–9672, 1992.

[67] J. K. Myers and T. G. Oas, "Preorganized secondary structure as an important determinant of fast protein folding," *Nature Structural Biology*, vol. 8, no. 6, pp. 552–558, 2001.

[68] S. Sato, T. L. Religa, V. Dagget, and A. R. Fersht, "Testing protein-folding simulations by experiment: B domain of protein A," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 18, pp. 6952–6956, 2004.

[69] J. N. Onuchic and P. G. Wolynes, "Theory of protein folding," *Current Opinion in Structural Biology*, vol. 14, no. 1, pp. 70–75, 2004.

[70] J. Lee, A. Liwo, and H. A. Scheraga, "Energy-based de novo protein folding by conformational space annealing and an off-lattice united-residue force field: application to the 10–55 fragment of staphylococcal protein A and to apo calbindin D9K," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 5, pp. 2025–2030, 1999.

[71] G. Favrin, A. Irbäck, and S. Wallin, "Folding of a small helical protein using hydrogen bonds and hydrophobicity forces," *Proteins: Structure, Function and Genetics*, vol. 47, no. 2, pp. 99–105, 2002.

[72] P. A. Alexander, D. A. Rozak, J. Orban, and P. N. Bryan, "Directed evolution of highly homologous proteins with different folds by phage display: implications for the protein folding code," *Biochemistry*, vol. 44, no. 43, pp. 14045–14054, 2005.

[73] M.-R. Yun, R. Lavery, N. Mousseau, K. Zakrzewska, and P. Derreumaux, "ARTIST: an activated method in internal coordinate space for sampling protein energy landscapes," *Proteins: Structure, Function and Genetics*, vol. 63, no. 4, pp. 967–975, 2006.

[74] L. Dupuis and N. Mousseau, "Holographic multiscale method used with non-biased atomistic force fields for simulation of large transformations in protein," *Journal of Physics*, vol. 341, no. 1, Article ID 012015, 2012.

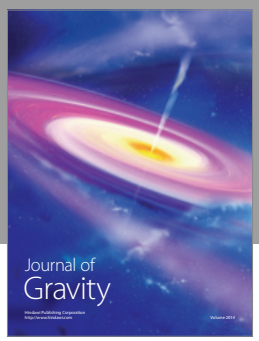
[75] B. D. McKay, "Practical graph isomorphism," *Congressus Numerantium*, vol. 30, pp. 45–87, 1981.

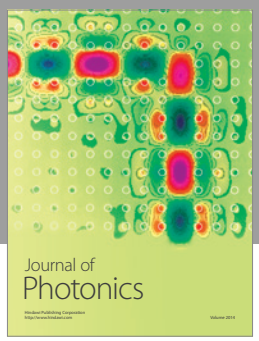
[76] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, "A new algorithm for Monte Carlo simulation of Ising spin systems," *Journal of Computational Physics*, vol. 17, no. 1, pp. 10–18, 1975.

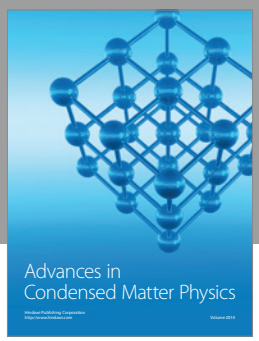
[77] Y. Song, R. Malek, and N. Mousseau, "Optimal activation and diffusion paths of perfect events in amorphous silicon," *Physical Review B*, vol. 62, no. 23, pp. 15680–15685, 2000.

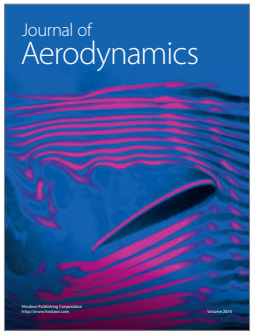
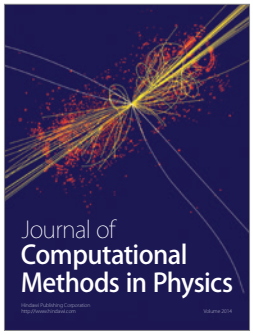
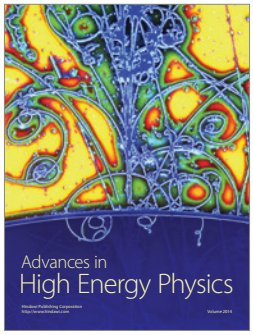
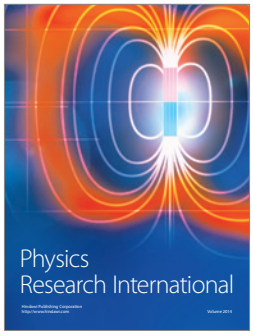
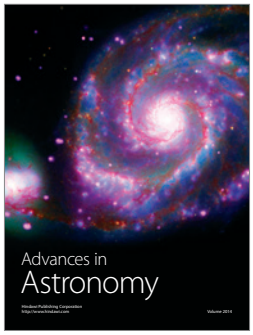
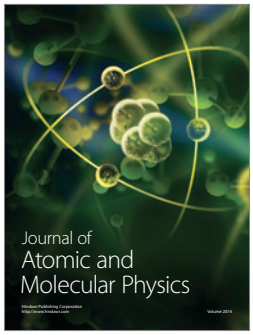
[78] H. Yildirim, A. Kara, and T. S. Rahman, "Origin of quasi-constant pre-exponential factors for adatom diffusion on Cu and Ag surfaces," *Physical Review B*, vol. 76, no. 16, Article ID 165421, 2007.

[79] B. Puchala, M. L. Falk, and K. Garikipati, "An energy basin finding algorithm for kinetic Monte Carlo acceleration," *Journal of Chemical Physics*, vol. 132, no. 13, Article ID 134104, 2010.







Hindawi

Submit your manuscripts at
http://www.hindawi.com