Genetic algorithm based approach to investigate doped metal oxide materials: Application to lanthanide-doped ceria

James Hooper, Arif Ismail, Javier B. Giorgi,*,† and Tom K. Woo*,‡
Centre for Catalysis Research and Innovation, Department of Chemistry, University of Ottawa, 10 Marie Curie, Ottawa, Canada K1N 6N5

(Received 23 March 2010; revised manuscript received 13 May 2010; published 9 June 2010)

A genetic algorithm (GA)-inspired method to effectively map out low-energy configurations of doped metal oxide materials is presented. Specialized mating and mutation operations that do not alter the identity of the parent metal oxide have been incorporated to efficiently sample the metal dopant and oxygen vacancy sites. The search algorithms have been tested on lanthanide-doped ceria (L=Sm,Gd,Lu) with various dopant concentrations. Using both classical and first-principles density-functional-theory (DFT) potentials, we have shown the methodology reproduces the results of recent systematic searches of doped ceria at low concentrations (3.2% L_2O_3) and identifies low-energy structures of concentrated samarium-doped ceria (3.8% and 6.6% L_2O_3) which relate to the experimental and theoretical findings published thus far. We introduce a tandem classical/DFT GA algorithm in which an inexpensive classical potential is first used to generate a fit gene pool of structures to enhance the overall efficiency of the computationally demanding DFT-based GA search.

DOI: 10.1103/PhysRevB.81.224104 PACS number(s): 61.72.Bb, 61.72.U−, 61.50.Ah

I. INTRODUCTION

Metal oxides are well known for their integral roles in a number of pioneering technologies with direct implications to the automotive, industrial, medicinal, and security sectors. Many of the application-specific and fundamental electronic or ionic properties of metal oxides can be improved upon by doping their lattice. Ceria (CeO_2), for example, is well suited for solid oxide fuel cell technology due to its high ionic conductivity.[1] When doped with a trivalent ion, intrinsic defects, characterized by oxygen vacancies, are created in ceria's doped fluorite lattice and it is the increased mobility of oxygen ions over these vacancies that is thought to be responsible for its enhanced ionic conductivity.[2] In recent studies, samarium was found to be the best dopant in this regard[3,4] and has since become a focal point of fuel-cell research intent on verifying and shedding light on the underlying principles responsible.[3,5,6]

From a theoretical standpoint, defect associations have been studied in samarium-doped ceria (SDC) from independent computational works on 3.2% SDC ($[Sm_2O_3]_1[CeO_2]_{30}$) and 2.2% SDC ($[Sm_2O_3]_1[CeO_2]_{46}$) by quantifying association energies or oxygen migration barriers via classical polarizable force fields and first-principles DFT calculations.[3,5] The 3.2% SDC model is particularly well studied because only one vacancy needs to be placed in a $2\times2\times2$ ceria simulation cell, making it feasible to perform a systematic screening of every possible configuration. The theoretical difficulties for modeling higher concentrations stem from navigating the number of ways in which more than one vacancy can be placed in the simulation cell and, more generally, systematic screenings of any doped metal oxide become unfeasible when dealing with a large number of defect-defect interactions in the simulation cell. For example, a systematic search of all possible ways to model 6.6% SDC by distributing four dopant atoms and two vacancies in a $2\times2\times2$ ceria simulation cell would involve screening approximately 3.5 billion possible configurations, ignoring symmetry-equivalent geometries. Many applications of doped metal oxides use even higher concentrations, at which both electrical and ionic conductivities are optimal in SDC, for example,[2,4,7] and the number of possible configurations becomes even higher.

The standard approach toward developing theoretical models of doped metal oxide materials still often relies on using one's intuition to manually place dopants or defects in the simulation such that all "relevant" configurations are studied.[8,9] As the concentration of the dopant increases, a manual exploration of the configurations not only becomes intractable but is more likely to miss important arrangements of dopants and defects.

A number of innovative methodologies have been developed to search out the most stable crystal structures of a system at given P-T conditions, notably those based on the genetic algorithm (GA) (Refs. 10 and 11), but they have only recently been successfully applied to study clusters, surfaces, and high-pressure solids at the DFT level of theory.[12–14] A recent GA-inspired experimental screening of doped ceria materials was also reported as part of an elaborate search for selective hydrogen oxidation catalysts.[15] These methodologies have proven to be effective at navigating nontrivial potential-energy surfaces by recursively generating lower energy structures through a number of predefined operations intended to pass on favorable structural traits from generation to generation until the overall health of the population converges. In materials design, inverse GA algorithms, algorithms designed to search out crystal structures that optimize the property in question, have been used to identify well-known or novel binary and ternary alloys with optimal heats of formation,[12] band-gap features,[16] or target structure-energy relationships.[17] A cohesive methodology that unambiguously determines ground-state structures of unknown metal oxide materials remains undeveloped.[18]

In this work, we will demonstrate that GA-inspired algorithms can readily be used to investigate doped metal oxide crystal structures and, specifically, we apply our own such

GA procedures to study lanthanide-doped ceria (LDC), focusing particularly on larger simulation cells of SDC materials. The mating and mutation routines have been designed to accommodate both dopant and vacancy defects in the ceria lattice and, in addition to SDC, gadolinium-doped ceria (GDC) and lutetium-doped ceria (LuDC) were investigated to compare our search algorithms with other potential-energy surfaces extracted from systematic, exhaustive searches of 3.2% lanthanide-doped ceria reported in the literature. The routines reproduce the results of all the systematic searches and agree qualitatively when applied to systems of similar concentrations in larger $3\times3\times3$ ceria supercells using classical potential-energy evaluations. At the DFT level of theory, even the ~100 atom $2\times2\times2$ ceria simulation cells make purely DFT-based GA trials too computationally demanding beyond 3.2% SDC concentrations. Thus, we introduce and evaluate a tandem classical potential and first-principles DFT GA method in which the classical potential is used to first screen out high-energy structures such that the final DFT-level evaluation of the structures is more efficient.

II. PROCEDURE

Genetic algorithm studies traditionally represent a crystal by a list of N atomic Cartesian coordinates, requiring no prior information about the system aside from the composition and volume of the unit cell. The crystal structures for this study were similarly represented by the Cartesian coordinates of each lattice site[12] in the fcc undoped CeO_2 lattice. A vacancy is created for every two dopant atoms so, specifically, the genetic representation of a generic LDC structure is constructed by assigning N dopant atoms to distinct ceria lattice sites and $N/2$ vacancies to distinct oxygen sites. A collection of LDC structures, or "genes," is herein referred to as a "population." The initial populations are constructed by substituting dopants and oxygen vacancies at randomly selected Ce and O sites on the undoped lattice. For example, the 3.2% and 6.6% L_2O_3 concentrations are built by introducing one and two oxygen vacancies, respectively, to a 96-atom CeO_2 $2\times2\times2$ supercell.

For structure evolution, the number of atoms was kept constant for the duration of the procedure wherein, at each step, randomly chosen mutation or mating operations were used to combine structural motifs of one or two randomly chosen "parent" genes. The mating/mutation operations were designed to accommodate both dopant-type and vacancy-type defects and preserve the lanthanide coordinates from the original $F_{m\text{-}3m}$ symmetry lattice. The mating and mutation operations used in this study were the following: (i) randomly perturbing the Cartesian coordinates of a chosen subset of atoms from the parent to make the offspring; (ii) swapping individual Cartesian coordinates of a randomly chosen subset of Ln atoms (Ln=lanthanide) with those of a randomly chosen subset of Ce atoms; (iii) swapping individual Cartesian coordinates of a randomly chosen subset of oxygen atoms with those of a randomly chosen subset of oxygen vacancies; and (iv) swapping the Ce/Ln lattice sites of one parent structure with those of the other parent structure. Only one randomly chosen mating/mutation operation was used to generate a single offspring. Occasionally these mating procedures generated physically unreasonable crystal structures, thus offspring structures were only added to the next generation if all atomic pairs were at least 1.0 Å apart, otherwise the same mating operation was repeated. The resulting offspring structures are given a 25% chance to mutate further through geometry perturbations. Specifically, this perturbation involves randomly selecting a nonempty pool of physically relevant lattice sites, by cycling through each atom and giving it a $1/N_T$ chance of being selected (N_T is the total number of atoms), and randomly perturbing each of their x, y, and z Cartesian coordinates by a maximum of 0.2 Å.

The "fitness" of a candidate structure, the metric that determines how well the structure fits the desired profile, was determined from its relative lattice energy to the lowest-energy structure in its population, as derived from classical polarizable force fields or first-principles DFT calculations. In light of this, we acknowledge that the temperature and oxygen chemical potential are known to play key roles when assessing the stability of metal oxides and, thus, finite temperature free energies would be better suited than the enthalpies discussed in this work. Such calculations are significantly more time consuming and impractical to apply at this time but we note the GA algorithms presented here could readily be used to sample the free energy instead of the enthalpy with minimal modification. The population of subsequent generations was created in the following manner: first, the fittest structures were promoted, unperturbed, to the next generation, usually those in the top 15 percentile. To generate the rest of the population, fit structures were preferentially chosen as parent structures by selecting them from a Boltzmann weighted probability distribution.[10] The exponential term stems from the relative lattice energies, scaled by an appropriate temperature to allow at least a 25% chance to select half of the promoted structures. Unphysical structures are screened out during the mating process and the offspring structures are optimized, under variable cell conditions at zero pressure, before evaluating their fitness.[16] The genetic algorithm was typically halted when the energy of the four lowest-energy structures remained unchanged for 5 or 20 generations in the 96-atom and 324-atom simulation cells, respectively.

We used the GULP simulation package[19] for all classical molecular mechanics calculations, using the Buckingham potential to model short-range dispersion pairwise interactions, Ewald's method[20] to sum the long-range Coulombic interactions, and the shell model[21] to account for the polarizability of the O^{2-} ions. The oxygen shell had a charge of $-2.08e$ and was tied to the atomic core by a 27.29 eV/Å^2 spring constant. The Buckingham parameters, shown in Table I, were taken from Balducci *et al.*[22] and Senyshyn *et al.*[23] Equation (1) shows the functional form of the Buckingham potential. The functions were evaluated between 0.5 and 10 Å for Ce-O and O-O pairs, and between 0.5 and 6 Å for Sm-O pairs. The elevation of the lower cutoff from 0.0 to 0.5 Å was employed to help efficiently eliminate unphysical structures from our genetic algorithm; it should in no way affect the energies of the low-energy structures. The presented force field parameters have not previously been used to study samarium-doped ceria, so they were benchmarked against

TABLE I. Buckingham parameters for interatomic potential calculation.

Species	A (eV)	ρ (Å)	C (eV Å^6)
O^{2-}-O^{2-} [a]	22764.3	0.149	27.89
Ce^{4+}-O^{2-} [a]	1986.8	0.3511	20.40
Sm^{3+}-O^{2-} [b]	4040.9	0.3034	0.0

[a]Reference 22.
[b]Reference 23.

other force fields that were recently used for Ln-doped CeO_2 materials.[5] The relative lattice energies of a pool of randomly generated structures matched qualitatively. The force field was also benchmarked and validated with first-principles DFT calculations by verifying that the order (by fitness) of selected structures was similar with both methods,

$$S_{ij} = Ae^{(-r_{ij}/\rho)} - \frac{C}{r_{ij}^6}. \quad (1)$$

First-principles DFT calculations were performed with the Vienna *ab initio* simulation package (VASP).[24,25] We used the projector augmented wave method of Blöchl[26] to treat the core states, with the oxygen $2s^2$ and $2p^4$, the cerium $5s^2$, $5p^6$, $6s^2$, $5d^1$, and $4f^1$, and the samarium $5s^2$, $5p^6$, $6s^2$, and $5d^1$ electrons being treated as valence electrons. The Hubbard (U-term) correction to DFT was not employed because the primary purpose of this work is to evaluate the robustness of a GA-inspired structural search algorithm when applied to doped metal oxides. In this work we examine a number of different dopants, which, like the Hubbard correction, would change the most favorable configurations. A plane-wave cutoff of 520 eV was used in all geometry optimizations and single-point energy evaluations with the gradient-corrected exchange and correlation functionals of Perdew-Burke-Ernzerhof.[27] For energy evaluations within the genetic algorithm, the Brillouin zone was sampled only at the Γ point, whereas Monkhorst-Pack grids were constructed from $3\times3\times3$ meshes in refined calculations of the phases selected for further study. Our calculations showed Γ-point calculations are sufficient for recovering accurate geometries and qualitatively describing relative energies of 3.2% and 6.6% SDC materials. We observed that when doped with Sm, the oxygen atoms neighboring the vacancy relax and the volume of the simulation cell expands by 3.2% to assume a slightly increased lattice constant of 5.47 Å.[7] In light of these observations, the fitness of each structure was determined directly from lattice energies calculated after full ionic and lattice relaxation of each structure and the algorithm was modified to read in the optimized cell parameters and incorporate the lattice information into the genetic representation of the structure itself.[16] Variable cell geometry optimizations at zero pressure were used to evaluate the fitness of all structures, as derived from the GULP and VASP simulation packages.

III. RESULTS AND DISCUSSION

We used our GA-inspired evolutionary algorithms to recursively scan the potential energy surface (PES) of LDC materials, focusing particularly on SDC, in order to ascertain the robustness and scalability of the GA-inspired search procedure. We start by comparing the GA results with a systematic search of 3.2% SDC materials at classical and DFT levels of theory to establish the effectiveness of the algorithm.[3,5] There are ~64 000 ways to distribute two dopants and one vacancy in our 3.2% LDC simulation cells, but, when symmetry is taken into account, the total number of unique LDC structures totals only 35. This makes systematic screenings of the PES feasible but still allows for a large enough overall search space to test the GA-inspired search routines. The 3.2% SDC findings are then supported with GA searches of 3.2% GDC and LuDC structures in order to test alternative potential-energy surfaces which can be compared to the work of Wei *et al.*[5]

To illustrate the algorithms' effectiveness for larger systems, we then investigate 1.9% and 3.8% SDC in $3\times3\times3$ ceria simulation cells to obtain a concentration comparable with the well-studied 3.2% SDC. Ignoring symmetry, there are 5.97×10^{12} and 3.00×10^{25} possible ways to distribute Sm atoms and oxygen vacancies about Ce and O lattice sites in the 1.9% and 3.8% SDC simulation cells, respectively, making them suitable demonstrations of the method's applicability. And finally, the GA search routines are used to explore the previously unexplored 6.6% SDC in a $2\times2\times2$ ceria supercell at an effective DFT level of theory, by introducing two vacancies into the $2\times2\times2$ ceria simulation cell and using tandem classical and density-functional-theory potential-energy evaluations as part of a "tandem," or dual-level,[28] optimization scheme. The classical evaluations are used to screen high-energy structures out the GA trial's gene pool and increase its overall efficiency by requiring fewer computationally expensive DFT calculations to reach convergence.

A. Explorations of 3.2% LDC materials

We first ran our GA-inspired structure searches on 3.2% SDC, hoping to validate the evolutionary protocols in the code. Using the undoped ceria $2\times2\times2$ supercell to generate initial structures, seven separate GA structure searches were performed with classical energy evaluations. With a population size of 20 individuals, the trials ran for, on average, just under 20 generations. First-principles DFT optimizations were then run on the best ten structures from the GA trials, starting from their classically derived cell and atomic positions.

The six lowest-energy structures recovered from these optimization runs are shown in Table II, labeled as structures S1–S6. The lattice energies are reported relative to the lowest-energy structure recovered from the GA trials. For the remainder of the text, we will classify LDC materials from the relative distribution of dopant atoms about vacancies in the lattice, specifically by referencing the dopant-vacancy distances as nearest neighbor (NN), next-nearest neighbor (NNN), and so on. For example, there are two such dopant-vacancy distances to account for in 3.2% LDC materials constructed from $2\times2\times2$ ceria simulation cells, so the SDC structure with a vacancy positioned NN to both samarium atoms will be labeled NN-NN.

TABLE II. Relative lattice energies and relevant structural information of 3.2% SDC structure recovered from GA search.

Structure	E_{GULP} [a] (eV/cell)	E_{VASP} [b] (eV/cell)	Sm-Sm distance (Å)	Sm-vac distance[c] (Å)
S1	0.000	0.000	6.7	NN-NNN
S2	0.042	0.007	3.8	NN-NN
S3	0.098	0.023	6.7	NNN-NNN
S4	0.103	0.028	7.6	NNN-NNN
S5	0.158	0.107	3.8	NNN-NNN
S6	0.216	0.042	3.8	NN-NNN

[a]GULP constant pressure optimization with the potentials from Refs. 22 and 23.
[b]Relative total electronic energies from VASP.
[c]NN denotes nearest neighbor (~2.4 Å) and NNN denotes next-nearest neighbor (~4.5 Å).

Prior first-principles calculations have found that in 3.2% SDC materials the dopant atoms prefer to occupy the NN site to the oxygen vacancy[3] as in the NN-NN structure, structure S2. Our results agree with this DFT study of Andersson *et al.* but note that a limited number of configurations were considered in that work (namely, structures S2, S5, and S6) and more recent DFT and classical studies claim a vacancy preference at the NNN site in SDC materials.[5,6] The NN-NNN SDC structure, structure S1, was recovered as the lowest-energy structure from our GA trials. In a systematic screening of all possible structures of 3.2% SDC at both classical and DFT levels of theory, we noted the same optimal Sm-Sm distance and NN-NNN positions with respect to the vacancy in the lowest-energy structure. The second, third, and fourth most fit structures from the GA trials, structures S2–S4 in Table II, were also recovered in the same energetic order within the systematic searches. The systematic search results therefore parallel the GA results, implying the evolutionary protocols in our GA are sufficient for exhaustively exploring the PES for 3.2% SDC materials. Structures S1 and S2 are shown in Fig. 1.

Encouragingly, Table II shows the relative DFT energies of the lowest-energy structures from the GA trials agree well with those predicted by the polarizable force field we assembled from Refs. 22 and 23. However, Wei *et al.*[5] reported 3.2% SDC favors structures that place the vacancy NNN to the dopant atoms, as in structure S4. To address this discrepancy we ran further GA trials and a systematic search on 3.2% SDC using the force field parameters reported by Wei *et al.* The lowest-energy structures from both search methods are given in Table III.

Our GA results match qualitatively with those from the systematic search, correctly recovering and ranking each of the structures in the correct energetic sequence. To further demonstrate the effectiveness of the genetic algorithm with alternative potentials we investigated the relative lattice energies of two other doped ceria materials for which a systematic search has already been completed.[5,6] Since the dopant's preferred site with respect to the vacancy crosses over from NNN to NN near gadolinium,[5,6] we chose to study 3.2% GDC and LuDC to complement the 3.2% SDC results discussed in Table III. For each dopant, we ran two GA trials with 15 genes in each population, allowing them to run for, on average, 15 generations, until the fittest structure remained unchanged for five generations. The lowest-energy LuDC and GDC structures extracted from our GA trials are shown in Tables IV and V, labeled L1 through L8 and G1 through G6, respectively. For ease of comparison, the last column of Tables IV and V refers to the structure notation used in Ref. 5.

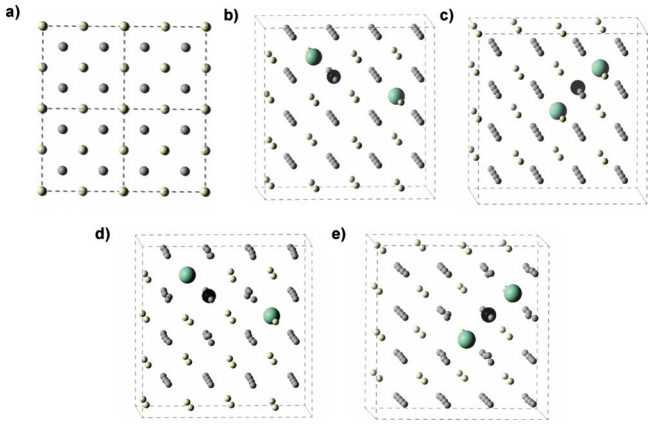

FIG. 1. (Color online) Extended structures of SDC. (a) $2\times2\times2$ supercell of CeO_2 crystal structure viewed along the [001] axis, the unrelaxed lattice of (b) structure S1 and (c) structure S2, and the relaxed geometries of (d) structure S1 and (e) structure S2. The Sm (large green balls), vacancy (medium black balls), cerium (small white balls), and oxygen (small gray balls) lattice positions are shown in all figures.

The recovered low-energy structures agree with our own systematic searches of LuDC and GDC with Wei *et al.*'s potentials, all of the structures shown in Tables IV and V

TABLE III. Relative lattice energies and relevant structural information of 3.2% SDC structure recovered from GA search using the force field parameters from Ref. 5.

Structure	E_{GULP} [a] (eV/cell)	Sm-Sm distance (Å)	Sm-vac distance[b] (Å)	Structure symbol[c]
S4	0.000	7.6	NNN-NNN	4_2_2
S3	0.001	6.7	NNN-NNN	3_2_2
S5[d]	0.016	3.8	NNN-NNN	1_2_2
S1	0.030	6.7	NN-NNN	3_1_2
S5[d]	0.090	3.8	NNN-NNN	1_2_2
S5[d]	0.091	3.8	NNN-NNN	1_2_2
S2	0.125	4.1	NN-NN	1_1_1
S6	0.138	3.8	NN-NNN	1_1_2

[a]GULP constant pressure optimization with the potentials from Ref. 5.
[b]NN denotes nearest neighbor (~2.4 Å) and NNN denotes next-nearest neighbor (~4.5 Å).
[c]Refers to the naming convention used in Ref. 5.
[d]Structures were found that had the same Sm-Sm and Sm-vacancy distances but were marginally different due to in-plane oxygen relaxations around the vacancy.

TABLE IV. Relative lattice energies and relevant structural information of 3.2% LuDC structures recovered from GA searches using the force field parameters from Ref. 5.

Structure	E_{GULP} [a] (eV/cell)	Sm-Sm distance (Å)	Sm-vacancy distance[b] (Å)	Structure symbol[c]
L1	0.000	4.1	NN-NN	1_1_1
L2	0.229	5.3	NN-NNN	2_1_2
L3	0.234	6.6	NN-NNN	3_1_2
L4	0.241	9.1	NN-NNNNN	5_1_4
L5	0.242	6.6	NN-NNNN	3_1_3
L6	0.267	7.6	NN-NNNN	4_1_3
L7	0.270	6.4	NN-NNNNN	3_1_4
L8	0.309	3.8	NN-NNN	1_1_2

[a]GULP constant pressure optimization with the potentials from Ref. 5.
[b]NN denotes nearest neighbor (~2.4 Å), NNN denotes next-nearest neighbor (~4.5 Å), and so forth.
[c]Refers to the naming convention used in Ref. 5.

were recovered, in the same energetic order, from the systematic searches. Overall, the NN-NN structure is favored strongly (by 0.229 eV/cell) in LuDC, not favored (thermoneutral with other structures) in GDC, and disfavored in SDC by 0.125 eV/cell. Therefore, the GA trials successfully reproduce the trends obtained from systematic screenings of such LDC materials reported in Ref. 5 and reinforce the notion that the methodology used in this work effectively explores low-energy 3.2% LDC materials. In the remaining work, we use these routines to replace systematic searches of larger systems, specifically those of 1.9% and 3.8% SDC in $3\times3\times3$ ceria simulation cells or 6.6% LDC materials in $2\times2\times2$ simulation cells.

B. Explorations of LDC materials using larger supercells

To test the scalability of our GA-inspired doping procedures, we ran several GA trials on Sm-doped $3\times3\times3$ ceria

TABLE V. Relative lattice energies and relevant structural information of 3.2% GDC structures recovered from GA searches using the force field parameters from Ref. 5.

Structure	E_{GULP} [a] (eV/cell)	Sm-Sm distance (Å)	Sm-vacancy distance[b] (Å)	Structure symbol[c]
G1	0.000	6.7	NN-NNN	3_1_2
G2	0.020	4.1	NN-NN	1_1_1
G3	0.029	6.7	NNN-NNN	3_2_2
G4	0.029	7.6	NNN-NNN	4_2_2
G5	0.048	3.8	NNN-NNN	1_2_2
G6	0.097	3.8	NN-NNN	1_1_2

[a]GULP constant pressure optimization with the potentials from Ref. 5.
[b]NN denotes nearest neighbor (~2.4 Å), NNN denotes next-nearest neighbor (~4.5 Å), and so forth.
[c]Refers to the naming convention used in Ref. 5.

TABLE VI. Relative energies of most thermodynamically favored structures of 1.9% SDC in a $3\times3\times3$ ceria simulation cell.

Rank in opt GA	E_{GULP} (opt) (eV/cell)	Sm-vac distances[a]	Sm-Sm distances[a]	vac-vac distance (Å)
B1	0.000	6 NNN, 1 NN	2 NNN, 2 NN	~6.1
B2	0.056	6 NNN, 0 NN	4 NNN, 0 NN	~6.1
B3	0.071	6 NNN, 1 NN	2 NNN, 2 NN	~6.7
B4	0.089	5 NNN, 1 NN	3 NNN, 1 NN	~6.7
B5	0.106	6 NNN, 1 NN	3 NNN, 2 NN	~6.1
B6	0.144	6 NNN, 0 NN	1 NNN, 1 NN	~6.7
B7	0.205	6 NNN, 0 NN	2 NNN, 1 NN	~6.7
B8[b]	1.500	0 NNN, 2 NN	4 NNN, 1 NN	~6.1

[a]Sm-vacancy distance columns show the number of Sm with that particular distance followed by a letter code for the distance value (NN is on average 2.4 Å and NNN is 4.5 Å).
[b]A select high-energy structure (B8) is shown for comparison.

supercells in order to simulate larger systems with concentrations comparable with the 3.2% SDC materials discussed above. When one, two, or four vacancy defects are introduced to the $3\times3\times3$ ceria lattice, this simulates 0.9%, 1.9%, and 3.8% SDC concentrations, respectively. Initially, all the GA trials were set up to use 100-gene population sizes with classically derived fitness metrics from the force field parameters assembled from Refs. 22 and 23. The trials were run until the fittest structure remained unchanged for 60 generations in order to ensure sufficient convergence of the trial within the larger search spaces; the effect of the simulation parameters on convergence is discussed further for a test scenario in Sec. III C. When two GA trials were run on 0.9% SDC, they converged after ten generations and recovered the NN-NNN structure as the lowest-energy structure. The results were analogous to those reported from placing one vacancy in a $2\times2\times2$ simulation cell.

Several GA trials were then run on 1.9% SDC in order to determine how large a population was needed to map out the system's potential-energy surface. Four sets of three trials were set up such that each trial stored either 100, 175, 250, or 325 structures, respectively, in their populations. All the trials with 175 or more genes per population converged after, at most, 100 generations to the same low-energy structure while all but one of the smaller trials converged to higher-energy structures. The 20 lowest-energy structures recovered from these trials were then read into the initial population of a further GA trial, which was run for 250 generations to test for lower energy structures. Although several new intermediate structures were located, no alternative to the lowest-enthalpy structure was recovered from the extra trial. The lowest-energy structures from the 1.9% SDC structure search, labeled B1–B7, are shown in Table VI. An additional GA trial was set up to store 500 genes per population and recovered no additional relevant information but did converge after only 30 generations.

The results show that several energetically competitive structures were recovered from the GA trials, namely, structures B1–B5. To quantify the structures, we consider the number of dopant atoms in the first three coordination shells

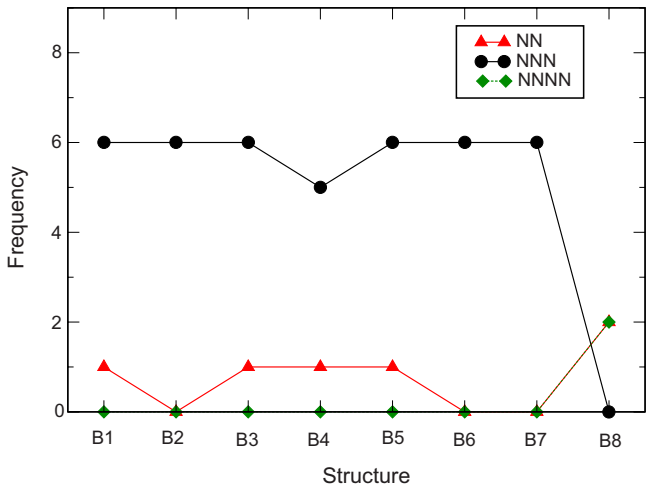

FIG. 2. (Color online) Plot showing the number of dopant Sm atoms in first three coordination shells about each vacancy in structures B1–B8 of 1.9% SDC in a $3\times3\times3$ ceria simulation cell.

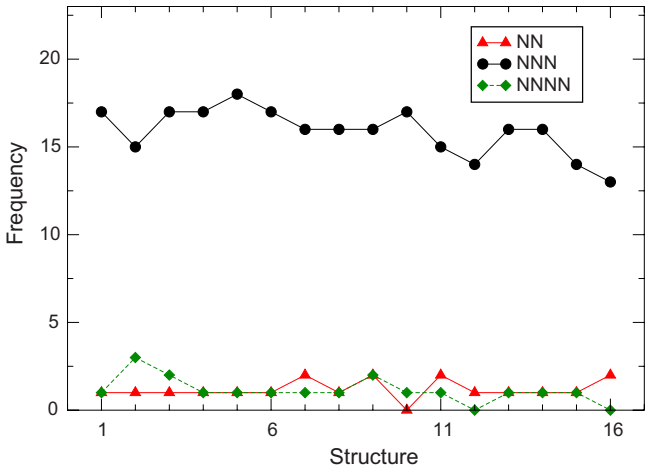

FIG. 3. (Color online) Plot showing the number of dopant Sm atoms in the first three coordination shells about each vacancy in 3.8% SDC structures. Structures increase in energy from left to right (see text for more details).

relative to each vacancy, note that only the structural information from the first two coordination shells is shown in Table VI. Figure 2 shows the sum of the number of dopant Sm atoms in the first three coordination shells of each vacancy in structures B1–B8. The results portray multiple energetically competitive configurations of SDC with dopant atoms predominantly positioned NNN to the vacancies. This type of behavior could not be captured from the 3.2% SDC materials derived from $2\times2\times2$ simulation cells since there are only two dopant-vacancy interactions in the simulation cell. Overall, there are either no NN vacancy-dopant interactions, as in structures B2 and B6, or a singleton NN vacancy-dopant interaction, as in structures B1 and B3. Structure B1 was recovered in all of the trials except two of those which stored 100 genes per population; structures B6 and B7 were recovered in those trials. A high-energy structure, B8, is shown in Table VI and Fig. 2 as a reference to a structure without NNN dopant/vacancy interactions.

Several GA trials using classical energy evaluations were then run on systems with four vacancy defects in a $3\times3\times3$ ceria simulation cell, modeling 3.8% SDC materials and allowing a concentration comparable with the 3.2% SDC simulation cells discussed above. For the first three trials, 300 structures were stored in each population and the trials were propagated until the fittest member of the population remained unchanged for 60 generations. Each trial converged to a different energy after running for, on average, 200 generations. The best four structures from each of these first three trials were then read into the initial population of a fourth GA trial which was run for 200 generations. A new low-energy structure was found after only 50 generations and remained converged for the duration of the simulation. This structure was then read into the initial population of a fifth GA trial, in addition to the structures read in for the fourth trial, and persevered as the most fit gene throughout 250 generations. Finally, two more sets of traditional GA trials were set up with either 400 or 500 genes per population and two-of-five and four-of-five trials, respectively, converged to the same low-energy structure within 200 generations.

As expected, a number of structures were recovered from these 3.8% SDC trials. To navigate the increased complexities of these large simulation cells, we present each of the three or four lowest-energy structures from a representative pool of the trials in Fig. 3, the lowest-energy structure is included among them. The metric plotted for each structure is the sum of the distribution of dopant Sm atoms over the first three coordination shells of each vacancy. The energy of the structures increases from left to right.

Energetically, all of the presented 3.8% SDC structures fall within 0.324 eV/cell. The lowest-energy structure (structure 1), featuring a single NN dopant/vacancy association, lies 0.186 eV/cell lower in energy than the lowest recovered structure with no NN associations (structure 10 from Fig. 3). The observed trend is reminiscent of that observed in 1.9% SDC simulation cells, favoring configurations with dominant, but not exclusive, NNN dopant/vacancy interactions. In comparison with the 3.2% SDC results, these favored mixed site-preference 3.8% SDC configurations relate well, in an abstract sense, to the lowest-energy NN-NNN structure recovered from 3.2% SDC simulation cells (structure S1). However, it is clear that configurations with exclusive NNN interactions are favored over those with exclusive NN interactions in 3.8% SDC materials; this was not the case for the 3.2% SDC materials (note structure S2 in Table II). Furthermore, the 3.8% SDC results show a much larger spread of low-energy, energetically competitive configurations with mixed site preferences, this trend could not be recovered from the 3.2% SDC simulation cells due the imposed symmetry of the smaller simulation cell. In summary, although the lowest-energy structures of ~3.5% SDC in each simulation cell show mixed NN/NNN site preferences, the NN site preference implied by the results in Table II in $2\times2\times2$ simulation cells is completely reversed to the expected NNN site preference in $3\times3\times3$ simulation cells.

Regarding the reliability of the search procedure itself, the overall convergence of each trial of the large 1.9% and 3.8% SDC simulation cells to their respective low-energy structure best demonstrates the power of these evolutionary proce-

dures. On average, the lowest-energy structures from the 1.9% SDC and 3.8% SDC GA trials were recovered after only 10 000 and 80 500 energy evaluations with 250 and 400 genes, respectively, per population. Since the same energy was converged upon in several independent trials and, in addition, no alternative lower energy structures were found when the most fit candidate structures were read directly into the initial gene pools of subsequent trials, we are confident that the encoded mating and mutation routines effectively sample LDC's configurational space. The larger system sizes however were noted to prolong the rates of convergence, requiring at least 175- and 500-member populations to reliably recover the respective lowest-energy 1.9% and 3.8% SDC structures within 200 generations. This suggests that consistent or "absolute" convergence can likely not be expected in systems with larger search spaces. Encouragingly, smaller-scale GA trials (with 300-gene population sizes) which build off the results of a previous trial were shown to function remarkably well in this regard on 3.8% SDC structures, providing an alternative stratagem. Given the stochastic nature of the algorithm and considering the results presented in Fig. 3, it is reasonable to assume trials on larger systems can, nonetheless, readily be used to assess optimal defect/defect interactions even if absolute convergence cannot be guaranteed since favored structural motifs will dominate the recovered structures' overall expressions. The next section further illustrates the power of the methodology, as we use tandem classical/DFT GA trials to perform an extensive search of the more-concentrated 6.6% SDC in a $2\times2\times2$ ceria simulation cell at an effective DFT level of theory.

C. Tandem classical/DFT GA investigations of 6.6% LDC materials

The added complications from studying higher dopant concentrations in LDC materials derived from $2\times2\times2$ ceria simulation cells stem from parsing the additional vacancy-vacancy interactions introduced into the simulation cell. Higher concentrations of LDC materials are more relevant to real applications of the material[7] but the extra defects result in a marked increase in the number of possible structures the simulation cell may assume. A systematic screening of these configurations at the DFT, and even classical, level of theory is unfeasible but, nonetheless, every possible set of low-energy defect interactions is important to consider as they may adversely influence the observed dopant-vacancy association energies.

We showed that our GA-inspired algorithms effectively search the potential-energy surface of ~3.2% SDC, GDC, and LuDC materials, recovering the lowest-energy structures as predicted from several distinct polarizable force fields and DFT calculations. Their success suggests that they can readily be used to replace systematic searches of higher concentrations, namely, those of 6.6% LDC materials which are created by introducing one more vacancy, and two more dopant atoms, into the simulation cell. However, since DFT calculations are far more time consuming than classical evaluations, we have noticed from our classical-based GA trials that even GA structure searches at the DFT level of theory would prove to be too computationally demanding. Considering this shortcoming, it seems prudent to narrow the scope of the DFT GA structure searches by first using a classical-based GA search to eliminate high-energy structures and recover a pool of structures to serve as a starting point for DFT-based GA trials. The idea behind such a practice is to speed up the DFT-based GA trial by starting it from a supposedly fit gene pool, thereby limiting the number of generations needed to reach convergence and saving substantial compute time. Moreover, since the SDC force field parameters assembled from Refs. 22 and 23 reproduce the results of DFT calculations on 3.2% SDC, they show promise for screening structures that would be low in energy at the DFT level of theory. Therefore, we will use a large-scale, dual-level GA structure search to recursively scan the potential-energy surface of 6.6% SDC, using the tandem classical/DFT approach outlined above.

First, we will demonstrate that such a tandem classical/DFT GA trial can reproduce the results we obtained for 3.2% SDC. Specifically, to set up the tandem trial, we modified one GA trial of 3.2% SDC, with 15 genes per population, such that the first 20 generations used classically derived fitness values[22,23] and the last five generations used DFT-derived fitness metrics. Note that the basis for using tandem classical/DFT GA trials assumes that the "screening" potential, the classical component, describes an adequate PES that agrees qualitatively to some degree with the "refined" potential, the DFT component. For this 3.2% SDC test scenario, we know this to be the case. As expected, this tandem GA trial finds no structure lower in energy than S1. Moreover, we note the mating routines from the classical trial recovered structure S1 after five generations and the routines from the DFT portion of the trial found no structures that displaced the five lowest-energy structures from the classical portion, meaning the lowest-energy structures were recovered after evaluating only one generation at the DFT level of theory. Note that the purpose of using DFT in the algorithm is to eliminate the possibility that a true low-energy structure is removed from the gene pool or a true high-energy structure is retained in the gene pool by the classical portion of the trial; clearly, this was not a problem at the 3.2% concentration.

Considering that a systematic screening of all possible configurations of 6.6% SDC materials is unfeasible, it is impossible to conclude whether the classical potential agrees with the DFT potential. In light of this uncertainty, one must ensure that the dual-level tandem GA method would suffice even if the screening potential did a poor job describing the refined potential. To ascertain whether the methodology would be sufficient even if a poor screening potential is used, we set up a tandem classical/classical GA trial on 6.6% LDC materials for which we can independently determine a near complete profile of the local minima on both potential-energy surfaces. In this framework, we can run separate, independent GA trials to explore the potential-energy surfaces of the screening and refined potentials since classical energy evaluations are fast and can be rigorously checked. The low-energy structures from the screening potential were then used as a starting point for a further GA trial with the refined potential in order to check whether it converges to the ex-

TABLE VII. Relative lattice energies (GULP) and relative electronic energies (VASP) of the top six 6.6% SDC structures from the "optimization" genetic algorithm.

Structure	E_{GULP} (opt) (eV/cell)	E_{VASP} (opt) (eV/cell)	Vac-vac distance (Å)	Sm-vac1 distance[a]	Sm-vac2 distance[a]	Closest Sm-Sm distances[b] (occurrences)
A1	0.000	0.128	6.7	4 NNN	4 NNN	~3.9 (2)
A2	0.048	0.000	6.1	3 NNN, 1 NN	3 NNN, 1 NN	~3.9 (2)
A3	0.102	0.020	6.1	4 NNN	4 NNN	~3.9 (3)
A4	0.154	0.020	6.1	4 NNN	3 NNN, 1 NN	~3.9 (2)
A5	0.161	0.030	6.1	4 NNN	3 NNN, 1 NN	~3.9 (3)
A6	0.211	0.041	6.1	4 NNN	2 NNN, 2 NN	~3.9 (2)

[a]Sm-vacancy distance columns show the number of Sm with that particular distance followed by a letter code for the distance value (NN is on average 2.4 Å and NNN is 4.5 Å).
[b]The closest Sm-Sm distances and the number of times they appear in the cell.

pected low-energy structures. Specifically, the LuDC potential of Wei *et al.*[5] was used to screen structures for the final SDC potential of Wei *et al.*,[5] thereby setting up a situation where the two potentials favor different dopant sites with respect to the vacancy (NN vs NNN). This would simulate a situation where there is a mismatch between the classical and DFT potentials.

To reiterate, separate traditional GA trials were first run on 6.6% LuDC and SDC materials in order to understand the precise potential-energy landscape of the LuDC and SDC classical potentials. With 120-gene population sizes, all of the 6.6% LuDC and SDC GA trials converged to the same low-energy structure after 20 generations. Larger GA trials, which were set up to store 150–200 individuals per population and run for 100 generations, recovered the same structures. A series of further trials were then run with smaller population sizes on 6.6% SDC, spanning 80, 50, 40, and 25 genes per population. All five 80-gene and four of the five 50-gene GA trials recovered the lowest-energy structure within 40 generations but only three of the five 40-gene and one of the five 25-gene trials recovered the same 6.6% SDC structure within 50 generations, requiring ~100 generations to reach convergence instead.

The lowest-energy LuDC structure had two distinct NN-NN Lu/vacancy associates, neighboring Lu atoms with a vacancy positioned NN to both of them, to create four overall NN Lu-vacancy interactions. The 20 lowest-energy LuDC structures were noted to all have at least three and at most five NN dopant-vacancy interactions. On the other hand, all of the 6.6% SDC trials converged to a low-energy structure with exclusive NNN dopant-vacancy interactions such that the vacancies are separated by approximately 6 Å, namely, structure A3 from Table VII. The next lowest-energy structure also featured exclusive NNN dopant-vacancy interactions but a longer vacancy-vacancy interaction, namely, structure A1 from Table VII. Both structures were recovered in all four 120-gene GA trials performed on 6.6% SDC. Each of the 20 lowest-energy structures heavily favored NNN interactions, with only one such structure possessing more than one NN interaction.

With the goal of recovering the lowest-energy 6.6% SDC structure, four tandem classical/classical GA trials with 50 genes per population were then set up for 6.6% SDC by reading in the 16 lowest-energy LDC structures from the 6.6% LuDC search into their initial gene pools and padding the rest of the population with randomly generated genes. The population size was chosen under the premise that this trial represents the second, more computationally expensive, portion of the tandem GA trial. In such a case, the smallest population size that reliably reproduces the lowest-energy structure is desired. Encouragingly, all four tandem GA trials converged to the expected low-energy SDC structures (A1 and A3) after 40 generations. The longer search time for the latter potential (40 vs 20 generations from the stand alone trials) is expected since the population size is smaller and the initial structures read into the gene pool from the LuDC trials are, on average, more unfavorable on the SDC PES than randomly generated structures. An additional set of tandem trials were then set up by reading in either the 25 or 40 lowest-energy structures from the LuDC trials and neglecting to introduce any randomly generated genes into the initial population. Only one of the four GA trials with 25 genes per populations recovered structure A3, within 50 generations, and two of the four GA trials with 40 genes per population recovered the expected structures. These results indicate that care must be taken when using a screening potential that does not correlate well with the final potential since convergence to the final structure can be hampered by starting from a heavily biased high-energy gene pool. Even so, the parameters we used in our tandem classical/classical GA trials did successfully reproduce the expected results even with an insufficient screening potential and should suffice for tandem classical/DFT GA investigations of 6.6% SDC with a potential that correlates well with DFT calculations at lower concentrations.

Returning to the objective of using the force field assembled from Refs. 22 and 23 to screen 6.6% SDC structures for DFT GA trials, seven classical GA trials, using classically derived fitness metrics,[22,23] and two subsequent DFT GA trials, using DFT-derived fitness metrics, were performed on 6.6% SDC. The five classical GA trials used population sizes of 80 or 120 individuals and were run for 40–50 generations. The six lowest-energy structures recovered collectively from these trials are shown in Table VII, labeled as structures A1–

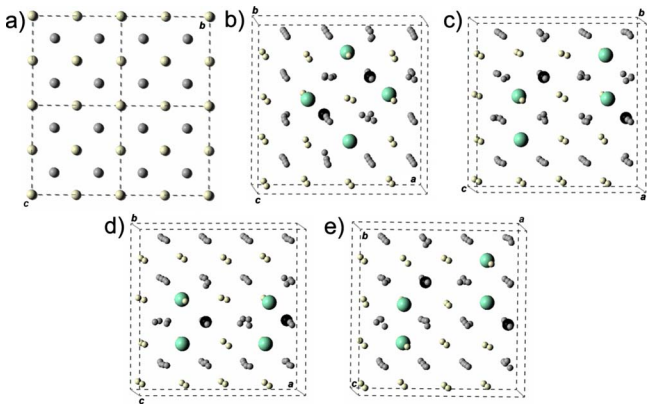

FIG. 4. (Color online) Extended structures of 6.6% SDC. (a) $2\times2\times2$ ceria supercell. Two views of structure A1 are shown projected along the (b) [001] axis and (c) [100] axis. Two views of structure A2 are shown projected along the (d) [001] axis and (e) [010] axis.

A6. Structure A2 was recovered in all five of the trials, structure A1 was recovered in three of the trials, and all the remaining presented structures were recovered in at least three of the trials. The DFT-based GA trials were then set up with a collection of low-energy structures from the classical GA trials already in their gene pools, setting up a tandem GA trial as discussed above. Specifically, the top 16 structures from the classical GAs and 34 randomly generated structures (for a total of 50 genes) were read into the initial populations. Upon their execution, it was noted that the mating routines from the DFT portion of the trials actually found no structures to displace the three lowest-energy structures carried over from the classical portion over the collective 25 generations from the trials. This was deemed sufficiently converged since the benchmark calculations on the tandem 6.6% LuDC/SDC GA trials each converged after such a standard was met.

First-principles DFT calculations do, however, disagree with our selected classical force field for 6.6% SDC materials, penalizing structure A1 in favor of all the other structures shown in Table VII. In general, the NNN vacancy positions with respect to the Sm atoms are still heavily favored but some NN interactions are seen as well. The lowest-energy structure A2, shown in Fig. 4 along with structure A1, has one such NN dopant/vacancy interaction. Note that when all these structures were input directly into the gene pool of GA trials employing DFT calculations, thereby applying further mating/mutation operations at the DFT level of theory, structure A2 persevered as the lowest-energy structure. We emphasize that although the classical components of the GA trials recovered a seemingly low-energy structure which is strongly disfavored at the DFT level (structure A1), it did successfully recover the remaining five lowest-energy structures in the same order as seen in the DFT portion of the GA trials. Importantly, this means the classical screening did effectively recover energetically favorable candidates. No examples of structures favored at the DFT level of theory but strongly disfavored at the classical level of theory were found.

The difference between the classical and DFT methodologies is best captured by comparing structures A1 and A3, which differ only by the relative positions of the vacancies. Classically, the short-range dispersion and coulombic interactions are indeed lower in energy in structure A3 by 0.169 eV/cell and 0.113 eV/cell, respectively. The long-range coulombic interactions, however, heavily favor structure A1 (by 0.383 eV/cell), thereby tilting the total energy to favor structure A1. Thus, our force field seems to overestimate long-range vacancy coulombic interactions. However, as discussed above, it correctly ranks the remaining low-energy structures and provides a quick, manageable way to screen the PES of 6.6% SDC before using DFT calculations. Given that a genetic algorithm using only first-principles methods is unfeasible due to limited computational resources, unless the PES is extremely small, classical methods are necessary but not sufficient for "exhaustive" first-principles evaluations.

Our results for 6.6% SDC parallel that of our 3.2%, 0.9%, 1.9%, and 3.8% SDC studies, finding that configurations with dopants positioned at both the NNN and NN positions to the vacancies, such that the NNN positions are favored, are the most thermodynamically stable configurations. The pool of low-energy, near-degenerate, structures clearly show the NNN site is preferred in SDC materials, agreeing qualitatively with recent classical and DFT investigations of less concentrated SDC in the literature.[5,6] Additionally, these results present the new notion that there is no heavily favored dopant/vacancy associate distribution in SDC, as supported by the recovery of several similar configurations with minimal, near-equivalent association energies.

Overall, our genetic algorithm was able to automate the search process of 6.6% LDC materials. We reiterate such concentrations have not been previously systematically investigated due to the sheer number of possible configurations. By automating a "dual-level" evolutionary search process, we were able to feasibly screen out unfavorable candidate structures with classical methods and then apply these same evolutionary procedures with higher-theory DFT calculations. Concerning this approach, we noted that when favored structures are screened out at the lower level of theory, the convergence of the GA slows down considerably but does not significantly alter its course in the long run as long as one is careful to introduce a pool of randomly generated structures into the gene pool when the higher level of theory is introduced.

IV. CONCLUSIONS

In summary, we have implemented and tested a specialized GA-inspired search procedure for doped metal oxides which allows the study of generic defect association complexes in metal oxide materials at experimentally relevant dopant concentrations. The key aspect behind these search techniques was the incorporation of mating and mutation routines that preserve the identity of the parent metal oxide itself; the specific mating operations applied herein are readily transferable to any doped metal oxide framework.

To demonstrate the methodology's effectiveness, we have shown such structure searches for low-energy structures of doped ceria reproduce the results of systematic searches of 3.2% samarium-, gadolinium-, and lutetium-doped ceria de-

rived from 96-atom simulation cells. In doing so, we used a number of distinct potentials and successfully recovered the well-documented crossover, across the lanthanide series, from NNN to NN dopant site preference to the vacancy at Gd^{3+}.[5] Furthermore, we found the NN-NNN structure was the most stable 3.2% SDC structure at the DFT level of theory, representing a balance of the competing effects of dopant-vacancy and dopant-dopant interactions. By using tandem classical and density-functional-theory energy calculations, the GA-inspired search procedures successfully mapped out 6.6% SDC at the DFT level of theory. The use of a classical optimization scheme in the early stages of the GA greatly accelerated the search process and proved to be an excellent screen for the more rigorous and computationally expensive DFT optimizations that followed. Rather than solely using the lowest-energy structures from the classical GA search, it was found that it was important to introduce a number of random structures into the gene pool to accommodate differences in the classical and DFT potential-energy surfaces.

Building on the preliminary 3.2% LDC results, we further showed the GA's potential by navigating SDC materials derived from large 324-atom simulation cells with as many as 12 defects, corresponding to concentrations upwards of 3.8% SDC. We found consistent convergence of several independent GA trials to the same low-energy structure, but, as expected, prolonged rates of convergence to the extent that consistent or absolute convergence cannot be expected in higher concentrations with population sizes sporting less than 500 structures. We note that trials on larger systems can, nonetheless, readily be used to assess optimal defect/defect interactions since the stochastic nature of the algorithm ensures favored structural motifs will dominate the recovered structures' overall expressions.

To illustrate the potential impact of assessing defect interactions from GA structure searches, we note that knowledge of the atomic positions of SDC provided herein, for example, can provide the ground work for future computational studies of these materials for use as catalysts and solid electrolytes. When the GA was applied to SDC, the NNN dopant site with respect to the vacancy site tended to be clearly preferred over the NN site in the larger simulation cells, as expected from the literature.[5,6] It was particularly interesting to observe the emergence these NNN interactions in the larger $3\times3\times3$ ceria cells at a similar concentration to those studied in prior calculations on $2\times2\times2$ simulation cells (3.8% vs 3.2% SDC). In both 3.8% and 6.6% SDC, there is only one NN dopant/vacancy association in the most stable configurations, further implying there are multiple similarly structured, energetically competitive configurations of SDC. With respect to SDC's use as solid electrolytes, these competitive structures would undoubtedly impact the material's oxygen mobility since their existence provides an abundance of SDC structures with zero relative association energies. Given the linear dependence of the migration barrier on trivalent lanthanide dopants as one progresses along the lanthanide series, this would then single out SDC as being a promising material for minimizing the asymmetric thermodynamic profile of migration barriers, as discussed by Wei *et al.*[5] and Nakayama and Martin.[6] Further DFT studies on this particular nuance for SDC materials are underway.

Regarding further possible applications of GA-inspired search methodologies toward other metal oxides, we note, when it becomes feasible to do so, the GA algorithms presented here could readily be used to sample the free energy instead of the enthalpy to assess scenarios where the temperature and oxygen chemical potential are known to play key roles. Furthermore, this methodology not only allows general studies of defect association energies but lays a further foundation to readily be incorporated into proposed "inverse optimization" schemes[16] in order to optimize a particular property of an oxide material.

*Corresponding author.
†javier.giorgi@uottawa.ca
‡twoo@uottawa.ca

[1] H. Inaba and H. Tagawa, Solid State Ionics **83**, 1 (1996).
[2] H. Yahiro, Y. Eguchi, K. Eguichi, and H. Arai, J. Appl. Electrochem. **18**, 527 (1988).
[3] D. A. Andersson, S. I. Simak, N. V. Skorodumova, I. A. Abrikosov, and B. Johansson, Proc. Natl. Acad. Sci. U.S.A. **103**, 3518 (2006).
[4] K. Muthukkumaran, R. Bokalawela, T. Mathews, and S. Selladurai, J. Mater. Sci. **42**, 7461 (2007).
[5] X. Wei, W. Pan, L. Cheng, and B. Li, Solid State Ionics **180**, 13 (2009).
[6] M. Nakayama and M. Martin, Phys. Chem. Chem. Phys. **11**, 3241 (2009).
[7] Y.-P. Fu, S.-B. Wen, and C.-H. Lu, J. Am. Ceram. Soc. **91**, 127 (2008).
[8] Z. Yang, G. Lao, Z. Lu, and T. K. Woo, J. Phys.: Condens. Matter **20**, 035210 (2008).
[9] Z. Yang, T. K. Woo, and K. Hermansson, J. Chem. Phys. **124**, 224704 (2006).
[10] D. M. Deaven and K. M. Ho, Phys. Rev. Lett. **75**, 288 (1995).
[11] Y. Zeiri, Phys. Rev. E **51**, R2769 (1995).
[12] G. H. Jóhannesson, T. Bligaard, A. V. Ruban, H. L. Skriver, K. W. Jacobsen, and J. K. Nørskov, Phys. Rev. Lett. **88**, 255506 (2002).
[13] A. R. Oganov and C. W. Glass, J. Chem. Phys. **124**, 244704 (2006).
[14] G. Trimarchi and A. Zunger, Phys. Rev. B **75**, 104113 (2007).
[15] J. Beckers, F. Clerc, J. H. Blank, and G. Rothenberg, Adv. Synth. Catal. **350**, 2237 (2008).
[16] K. Kim, P. A. Graf, and W. B. Jones, J. Comput. Phys. **208**, 735 (2005).
[17] S. V. Dudiy and A. Zunger, Phys. Rev. Lett. **97**, 046401 (2006).
[18] J. K. Nørskov, T. Bligaard, J. Rossmeisl, and C. H. Christensen, Nat. Chem. **1**, 37 (2009).
[19] J. D. Gale, J. Chem. Soc., Faraday Trans. **93**, 629 (1997).

[20] P. P. Ewald, Ann. Phys. **64**, 253 (1921).
[21] J. D. Gale, Philos. Mag. B **73**, 3 (1996).
[22] G. Balducci, J. Kaspar, P. Fornasiero, M. Graziani, and M. S. Islam, J. Phys. Chem. B **102**, 557 (1998).
[23] A. Senyshyn, A. R. Oganov, L. Vasylechko, H. Ehrenberg, U. Bismayer, M. Berkowski, and A. Matkovskii, J. Phys.: Condens. Matter **16**, 253 (2004).
[24] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).
[25] G. Kresse and J. Furthmuller, Comput. Mater. Sci. **6**, 15 (1996).
[26] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[27] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[28] D. Balamurugan, W. Yang, and D. N. Beratan, J. Chem. Phys. **129**, 174105 (2008).