

Discovering chemical site occupancy- modulus correlations in Ni based intermetallics via statistical learning methods

Scott R. Broderick^a, Aakash Kumar^{b,c,1}, Adedapo A. Oni^d, James M. LeBeau^d, Susan B. Sinnott^{b,c}, Krishna Rajan^{a,*}

^a Department of Materials Design and Innovation, University at Buffalo – The State University of New York, Buffalo, NY 14260, USA

^b Department of Materials Science and Engineering, University of Florida, Gainesville, FL 32611, USA

^c Department of Materials Science and Engineering, The Pennsylvania State University, University Park, PA 16802, USA

^d Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC 37695, USA

ARTICLE INFO

Keywords:

Materials informatics

Density functional theory

Density of state

Ni-base alloys

Scanning transmission electron microscopy

ABSTRACT

We show how one may extract spectral features from the density of states (DOS) of $L1_2$ - Ni_3Al alloys that can serve as signatures or electronic “fingerprints” which capture the correlation between site occupancy of dopants and elastic properties. Based on this correlation, we have developed a computational approach for rapidly identifying the impact of the selection of dopant chemistries on bulk moduli of intermetallics. Our results show for example that Cr preferentially occupies the Al site in Ni_3Al which is confirmed by scanning transmission electron microscopy (STEM) energy dispersed X-ray spectroscopy (EDS) analysis. We further show that this preference is due to a sensitivity of Cr to the DOS at -1.7 and 0.2 eV relative to the Fermi energy. In terms of similarity in chemistry-property correlations, we find Cr has a similar effect to Ce when occupying an Al site, while Cr occupying a Ni site has similar correlation as La on a Ni site. This logic can be utilized in targeted design of new alloy chemistries based on similar property correlations and for targeted DOS modification.

1. Introduction

In this paper, we extract the signatures of the DOS spectra which dictate the elastic moduli of alloys. Further, the DOS signatures of alloying site preference are identified, with the focus here on Ni-based alloys. In our previous work, we developed the relationship between the chemistry and crystal structure and the DOS signatures based on the characteristics of the elemental constituents [1–3]. Combining these threads, we are able to accelerate the chemical design of multi-component systems for targeted properties by developing chemistry-electronic structure-property relationships, thereby reducing the number of electronic structure calculations required by rapidly screening the enormous chemical search space.

Nickel-based superalloys, which are of interest for their high-temperature behavior, have two phases present: γ -Ni matrix and γ' - Ni_3Al . The Ni_3Al phase has an $L1_2$ structure with Al at the corners and Ni at the face centers, and is the principal reason for the high temperature strength of superalloys due to the yield stress anomaly, where the yield stress increases with increasing temperature up to a limit [4]. Different alloying elements can modify the elastic properties of Ni_3Al [5–7], with

the superalloys having a complex composition with generally around 10 alloying elements. It is important to investigate the alloying behavior of the different elements in terms of their site-preference and elastic properties. A method for predicting site preference of a dopant atom in this Ni_3Al phase was suggested by Ruban and Skriver [8], based on the energy change of moving an atom X from a Ni site to an Al site ($E_X^{Ni \rightarrow Al}$). Later, Gleeson et al. [9] showed that this single parameter $E_X^{Ni \rightarrow Al}$ can be used to predict the site preference of a dopant atom. This parameter has been studied for many transition elements [7–10]. This method assumes the Wagner-Schottky model to be true [11] according to which the formation enthalpy of the ternary alloy Ni-Al-X is a linear function of the formation enthalpy of Ni_3Al and that of the point defect.

$$\Delta H = \Delta H_{Ni_3Al} + \sum_d H_d x_d \quad [1]$$

Here d is the point defect under consideration, and x_d its atomic concentration. H_d is the formation enthalpy of the defect.

The energy change is defined as:

$$E_X^{Ni \rightarrow Al} = E(Ni_{24}Al_7X) + E(Ni_{23}Al_8) - E(Ni_{23}XAl_8) - E(Ni_{24}Al_8) \quad [2]$$

* Corresponding author.

E-mail address: krajan3@buffalo.edu (K. Rajan).

¹ Present address: Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA.

When this quantity is negative, the atom preferentially occupies the Al site, while when it is positive the atom occupies either the Ni site or both sites. More precisely,

- a) $E_X^{Ni \rightarrow Al} < 0$, X occupies Al site
- b) $E_X^{Ni \rightarrow Al} > H_{AlNi} + H_{NiAl} = 1.12$ eV [7], X occupies Ni site
- c) $0 < E_X^{Ni \rightarrow Al} < H_{AlNi} + H_{NiAl}$, X atoms can occupy both sites

Some dopants of interest for Ni₃Al include B for enhanced mechanical strength [12,13], Cr for improved oxidation resistance [4,13], and Ce for both better mechanical strength and oxidation resistance [12,14]. Zr provides strengthening and resistance to creep [15,16] and La has shown to improve resistance to oxidation and hot corrosion [13]. While it is established that Ce forms a second phase CeNi₅ at concentrations exceeding 0.63 wt% Ce [17], Ce addition contributes to the desired properties. However, it is a rare earth element and a challenge exists to replace Ce with a non-rare earth element while maintaining the desired properties. While Ce and Cr are substitutional, B occupies interstitial voids owing to its small size. As a result, the replacement of Ce with Cr and B for oxidation resistance and mechanical strength is not possible, due to their different site occupancies. This defines the challenge that is addressed by this paper, namely to assess the change in properties that result due to changes to both composition and site occupancy.

The following doping chemistries and sites discussed here are: undoped Ni₃Al, Ce, La, Cr and Zr doped on the Ni site, Ce, La, Cr and Zr doped on Al site, and B doped on tetrahedral interstitial and octahedral interstitial voids. Beyond the properties of interest, these chemistries were selected so that their additional differences such as size differences and periodic table groups are integrated into the informatics analysis. For each of these compounds, the electronic Density of States (DOS) and bulk modulus (B_m) were calculated via Density Functional Theory (DFT). An informatics analysis is then applied to the DOS spectra for assessing the possible compound chemistries that can exist, comparing behavior of different dopant chemistries, and linking dopant site occupancies with mechanical behavior. We have previously applied an informatics analysis to the DOS spectra [1–3], although those previous works did not consider site occupancies. Other approaches combining informatics algorithms with electronic structure calculations sought to accelerate the calculations so that data can be generated in a high-throughput manner, and to search for appropriate crystal structures or structure determination from the significantly increased knowledge-base [18–20]. An alternate informatics approach for predicting the bulk and shear moduli has recently been demonstrated by de Jong et al. [21,22], where a massive number of property values were predicted based on series of descriptors including formation energy, volume-per-atom, other elemental properties, and relative elemental positioning within the periodic table.

By employing the entire DOS spectra as we do here, the span of reversible property descriptors are theoretically included without any prior assumptions. As the DOS represents all electronic interactions of a system, it should contain information on all electronic properties [23–25]. However, the understanding of how these properties are captured by the DOS is not well understood. One example of a property which is known to be at least qualitatively represented within the DOS spectra is bulk modulus [26], where the difference between Fermi energy and antibonding-bonding transition correlates to B_m . Some physical explanation of the connection between B_m and DOS spectra can thus be provided, although it remains qualitative and limited. Building on this concept is one of the primary objectives in linking DOS spectra with an informatics approach. By contrast, we use informatics approaches to relate patterns in the DOS spectra to the bonding of alloy systems and the corresponding properties.

Cr has previously been shown to have an Al site preference via calculation of $E_X^{Ni \rightarrow Al}$ [5–10]. Calculation of this quantity for Ce substitution has not been reported. To predict this site occupancy, we

must understand why Cr occupies the Al site and not the Ni site. In addition to assessing the calculations at 0 K, we perform experimental analysis of Ni₃Al-Cr to ensure that the modeling of site occupancy holds at higher temperatures. While the focus of this paper is limited to site preference, we also recognize that the concentration of defects is critical to consider since the properties discussed here are largely dependent on the final equilibrium state of the system. Future work will address this added layer of consideration by including chemical potentials to describe the thermodynamic interactions between the γ and γ' phases. This paper introduces an informatics driven approach for extracting previously undefined chemistry-electronic structure-property correlations, even though the systems considered have been extensively studied previously. We discuss how this added layer of information which is not available solely from the unprocessed DFT output allows us to accelerate the design of new Ni-base alloys for specific applications.

2. Methods

2.1. Density functional theory calculations

The supercell consisted of $2 \times 2 \times 2$ unit cells of Ni₃Al, shown in Fig. 1 (a). Cr, Zr, La and Ce dopants were placed in either the Al or Ni site (Fig. 1 (b) and (c)), while B was placed in the tetrahedral and octahedral void (Fig. 1c and (d)). Cr, Zr, La and Ce were varied with concentrations of 3.125 at%, 6.25 at% and 9.375 at%, which corresponds to 1, 2 and 3 dopant atoms substituting either Ni or Al site in Ni₃Al. For B, the concentration used was 3.03 at%. The concentrations were chosen keeping in mind the compositions used in superalloys [27].

The Vienna Ab initio Simulation Package (VASP) [28–31] was used to perform the density functional theory (DFT) calculations. The exchange-correlation functional by Perdew, Burke and Ernzerhof was employed within generalized gradient approximation (PBE-GGA) [32,33]. Projector Augmented Wave (PAW) scheme [34] was used to account for the core electrons, while the energy cut-off for the plane wave basis was chosen to be 350 eV, which was determined to be a suitable value. A Monkhorst-Pack [35] K-mesh of $6 \times 6 \times 6$ was used. All calculations were spin-polarized and the atoms were relaxed with a convergence criteria of 10^{-7} eV for energy and 10^{-3} eV/Å for the forces on atoms. The elastic constants were then extracted from the stress-strain relationship by straining the relaxed structures to obtain $\bar{C}_{11}$, $\bar{C}_{12}$, and $\bar{C}_{44}$, where $\bar{C}_{11} = (C_{11} + C_{22} + C_{33})/3$, $\bar{C}_{12} = (C_{12} + C_{13} + C_{23})/3$ and $\bar{C}_{44} = (C_{44} + C_{55} + C_{66})/3$. The convergence achieved was within an error bar of 8 GPa except in the case of Ce-doped Ni₃Al on the Ni-site where the error bar range is 10–15 GPa. Complete details of the elastic constant calculations and the convergence achieved are provided in our prior report [7]. The density of states (DOS) was calculated with the energy convergence criterion increased to 10^{-8} eV and a Monkhorst-Pack K-point mesh of $9 \times 9 \times 9$ for both pure and doped Ni₃Al systems. The Fermi-energy was shifted to 0 eV for all the calculations for consistency.

2.2. Informatics

To identify the underlying governing features within the DOS spectra, we employed the entire DOS spectra to quantitatively analyze the DOS without any prior assumptions by treating the DOS as spectral data, building on our prior work in analyzing spectral data [36,37]. We apply eigenvalue decomposition techniques such as principal component analysis (PCA) [36–41]. As we have demonstrated in our earlier studies [1–3,36–38], this permits one to identify the right projection of data from which meaningful features associated with the DOS can be identified. Partial Least Squares (PLS) were further applied to this high dimensional data space [42–50]. These were used in a predictive informatics mode to link the DOS spectra with elastic constants. By applying PCA to the DOS spectra, the strongest patterns in the data can be

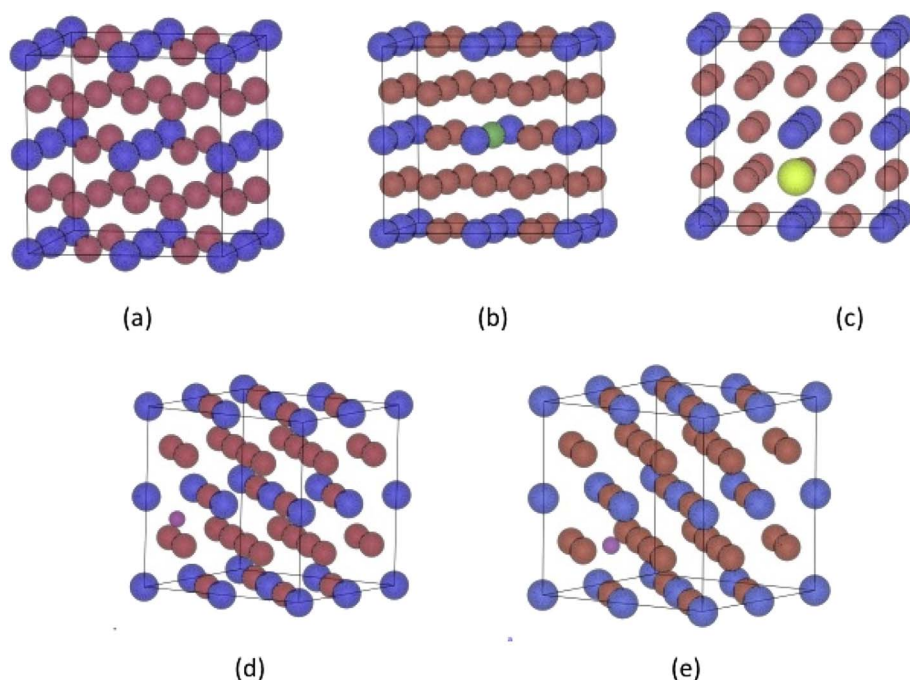


Fig. 1. Crystal structure and site substitution for (a) L1₂-Ni₃Al, which serves as input into the DFT modeling; (b) Cr atom substituting for an Al atom; (c) Ce atom substituting for Ni atom, B atom occupying (d) a tetrahedral void, and (e) an octahedral void.

identified in a limited number of dimensions.

PCA operates by performing an eigenvector decomposition of the data. As such, the principal components (PCs) capturing the most information are associated with the largest eigenvalues of the covariance matrix and their corresponding eigenvectors. The original data is decomposed into two matrices of interest for this paper: the scores and loadings. The loadings and scores matrices contain the principle patterns within the DOS curves and the scaling of those patterns to create the final DOS curve, respectively. In PLS the training data is converted to a data matrix with orthogonalized axes, which are based on capturing the maximum amount of information in fewer dimensions. The relationships discovered in the training data can be applied to a test dataset based on a projection of the data onto a high-dimensional hyperplane within the orthogonalized axis-system. Typical linear regression models do not properly account for the co-linearity between the descriptors, and as a result the isolated impact of each descriptor on the property cannot be accurately known.

The ability to predict properties of alloy systems from an input of DOS requires the integration of PCA and PLS. This approach represents a hybrid informatics approach because properties are determined as a summation of the PLS coefficients (the prediction coefficients of the PC scores values on the elastic constants) and the PCA weightings (*ie.* Property = $f[\Sigma(\text{Component of PLS result} * \text{Component of PCA result})]$) as opposed to isolating the components (*ie.* Property = $f[\Sigma(\text{Component of PLS result}) * \Sigma(\text{Component of PCA result})]$). That is, in the final development of an equation, the PLS and PCA components of the analysis cannot be separately extracted. This hybrid capacity of the approach is utilized in this paper to connect elastic constants with the corresponding DOS spectra.

2.3. Experiment

A Ni-Al-Cr superalloy crystal with nominal composition of Ni-75.7 Al-16.7 Cr-7.5 at % was grown by the directional Bridgman technique. The as-cast sample was then annealed at 1250 °C ± 10 °C for 14 days in He environment. TEM specimens were prepared by conventional mechanical polishing followed by low energy ion milling recipe. Atomic resolution EDS analysis was performed with a probe-corrected FEI Titan G2 60-300 kV S/TEM equipped with a high-brightness Schottky field emission gun, operated at 200 kV. The convergence angle and HAADF

detector inner semi-angle are 21 mrad and 77 mrad, respectively. The EDS elemental maps were acquired with a beam current of ~0.1 nA. A 5-pixel smoothing filter was applied to the EDS elemental maps during post-processing to reduce noise using Bruker esprit software package.

3. Results

DFT calculations were performed for undoped Ni₃Al, 0.03% B doped in interstitial voids, Ni₂₄Al_{8-x}Cr_x, Ni_{24-x}Al₈Cr_x, Ni₂₄Al_{8-x}Zr_x, Ni_{24-x}Al₈Zr_x, Ni₂₄Al_{8-x}Ce_x, Ni_{24-x}Al₈Ce_x, Ni₂₄Al_{8-x}La_x and Ni_{24-x}Al₈La_x, for $x = 1, 2, 3$. The output of the calculations were DOS spectra (Fig. 2) and bulk moduli (Table 1). From visual inspection of the spectra, the change due to site occupancy is not significantly evident and emphasizes the need to apply manifold learning analysis. The doping of B causes an increase in DOS at low energies, while the addition of Ce causes significant increase above the Fermi energy ($E_F = 0$ eV) due to the introduction of *f*-states. However, these visual interpretations do not provide any information on the correlation between the change in DOS spectra with changing chemistry and the change in elastic properties. Identifying those correlations provides a more critical design guideline for modifying dopant chemistry.

The informatics analysis was performed on 27 compounds, with the result shown in Fig. 3. Each point within the DOS spectra was considered as a separate descriptor, so that the dimensionality of the problem was on the order of 10^3 . Sampling a larger number of points from the DOS spectra results in a higher resolution result, but also a less robust result as specific features may be over-fit as opposed to capturing the general and physically significant trends. Further, the sensitivity to exact DFT parameters is reduced, and our objective of developing a generalized approach is enhanced. Therefore, after performing multiple informatics simulations to compare accuracy versus robustness, this number of points was identified as capturing the same results as larger descriptor sets while enhancing the potential robustness. The principal components (PC1 and PC2) provide two new uncorrelated descriptors of the data, and the scores values provide a parameterization of the DOS spectra which is based on maximizing information extraction. Relationships in the data are then defined in terms of Euclidean distance and trajectory. PC1 captures changes in the DOS due to Ce doping, as evidenced by Ce increase along PC1, while other doping chemistries have little change along this axis. PC2 captures changes in DOS due to

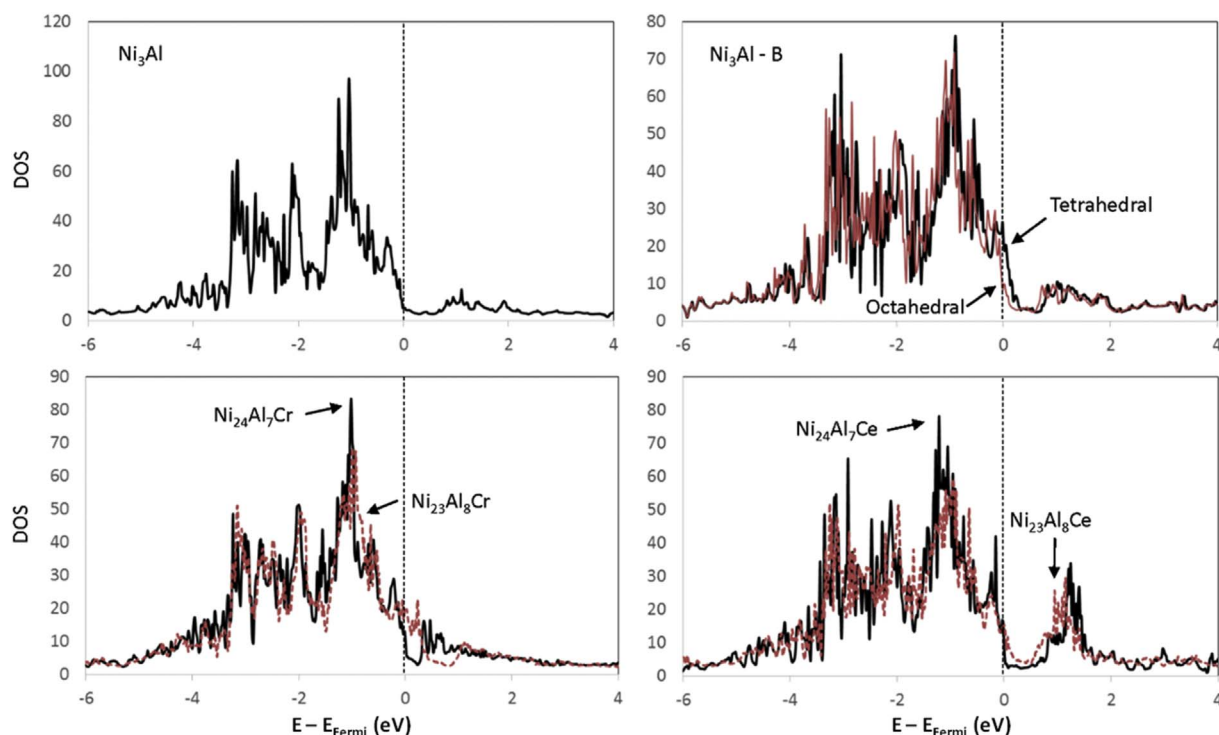


Fig. 2. Input DOS for the informatics analysis. While changes in the spectra are evident with changing dopant chemistry, the changes due to site occupancy are much more subtle. Informatics is applied to these spectra so that hidden features which control site occupancy energy are evident and possible chemistries are rapidly determined.

Table 1

DFT calculated elastic properties [7]. Using these properties for dopant chemistry and site, correlations between mappings of DOS spectra and mechanical properties can be made, developing a quantitative structure-property relationship for new compounds.

	$\bar{C}_{11}$	$\bar{C}_{12}$	$\bar{C}_{44}$	B
Ni3Al	246	155	130	185.33
Ni3Al(B) ^a	255	148	118	183.67
Ni3Al(B) ^b	255	149	118	184.22
Ni ₂₄ Al ₇ Cr	252	154	131	186.67
Ni ₂₄ Al ₆ Cr ₂	262	156	137	191.11
Ni ₂₄ Al ₅ Cr ₃	266	156	137	192.56
Ni ₂₃ Al ₈ Cr	233	152	125	179.33
Ni ₂₂ Al ₈ Cr ₂	239	161	125	186.78
Ni ₂₁ Al ₈ Cr ₃	232	152	121	178.78
Ni ₂₄ Al ₇ Zr	246	152	123	183.33
Ni ₂₄ Al ₆ Zr ₂	243	145	119	177.89
Ni ₂₄ Al ₅ Zr ₃	248	147	111	180.89
Ni ₂₃ Al ₈ Zr	225	149	115	174.67
Ni ₂₂ Al ₈ Zr ₂	220	142	104	167.89
Ni ₂₁ Al ₈ Zr ₃	207	138	95	160.67
Ni ₂₄ Al ₇ Ce	235	141	113	172.44
Ni ₂₄ Al ₆ Ce ₂	236	141	97	202.20
Ni ₂₄ Al ₅ Ce ₃	229	135	85	207.19
Ni ₂₃ Al ₈ Ce	221	132	104	190.44
Ni ₂₂ Al ₈ Ce ₂	213	126	80	190.44
Ni ₂₁ Al ₈ Ce ₃	178	110	70	190.44
Ni ₂₄ Al ₇ La	229	134	109	165.67
Ni ₂₄ Al ₆ La ₂	213	124	92	153.44
Ni ₂₄ Al ₅ La ₃	198	114	79	142.11
Ni ₂₃ Al ₈ La	213	137	102	162.22
Ni ₂₂ Al ₈ La ₂	198	129	88	151.89
Ni ₂₁ Al ₈ La ₃	212	97	59	135.22

^a 0.03 atomic percent B in octahedral void.

^b 0.03 atomic percent B in tetrahedral void.

other dopings with La having the largest impact. This demonstrates the impact in rare earth additions. Doping B in the octahedral site has a larger effect as compared to doping in the tetrahedral site, as the compound with B in the tetrahedral site sits closer to pure Ni₃Al.

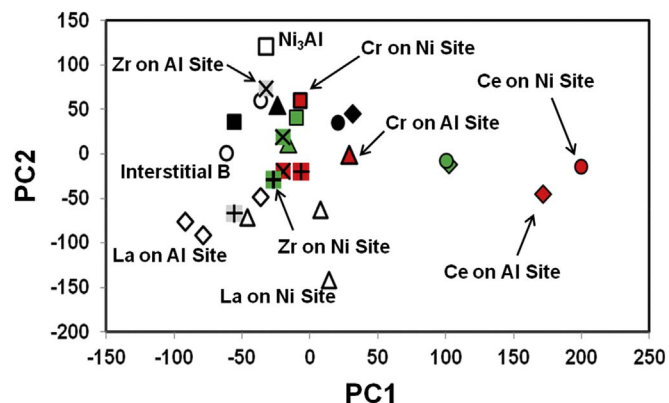


Fig. 3. Principal component projection showing the correlations associated with site occupancy and site chemistry. Black/gray filled shapes correspond to 3.125 at% dopant, green is 6.25 at% dopant and red filled shapes are 9.375 at% dopant concentration. Filled squares are Cr on Ni site, triangles are Cr on Al site, circles are Ce on Ni site, diamonds are Ce on Al site, 'X' are Zr on Al site and '+' are Zr on Ni site. Unfilled diamonds are La on Al site, unfilled triangles are La on Ni site, and unfilled circles are B on interstitial sites. From this we identify Ce as occupying either site, Cr occupying only Al site, Zr occupying only Al site, and La not having a strongly stable additional configuration. As an example, the reason Cr does not occupy Ni site is described by the DOS, as evidenced by little relationship between Cr content and PC values, while Cr on Al site forms a clear trend from undoped Ni₃Al to increasing Cr content. This interpretation can be applied to each chemistry, with stability defined by clear trends associated with additive composition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The trajectory in this map with increasing dopant concentration provides an explanation of the energetics of site substitution. The interpretation of site occupancy is that the increase of dopant continues to make the material more dissimilar to Ni₃Al. However, if it is not energetically favorable for the atom to be located on that site, the DOS will not change in such a consistent manner. When looking at Cr addition, we find that the alloy becomes less like Ni₃Al in a consistent manner with concentration of Cr on the Al site. However, when adding

Cr on the Ni site, there is no clear pattern, as demonstrated by the arced trajectory. The lack of clear systematics represents the unfavorable energetics of Cr addition to the Ni site. Building on this interpretation that the trajectory between PC values and dopant chemistry captures the defect formation energy, we predict that Ce is able to occupy either site due to the linear trajectories of both site substitutions. Note that the relationships with chemistry and site occupancy cannot be identified from the raw DOS spectra, and Cr on the Ni site would not be identified as an outlier chemistry without informatics. As the informatics analysis is extracting the strongest physics based trends within the DOS spectra, it is fairly insensitive to minor changes in the DOS which would result from variation of DFT parameters, resulting in a robust generalized approach. Further, the predictions from the data mining are in good agreement with the training cases of no occupancy of Cr and Zr on Ni site, strengthening the argument that the DFT calculation mesh and supercell sizes were sufficient as an input into the informatics model.

To test the predictions for Cr, we experimentally assessed the site substitution via EDS. The typical microstructure of the ternary Ni-Al-Cr compound consisted of a disordered (fcc structure) γ -matrix with spherical ordered ($L1_2$ structure) γ' -precipitates. EDS was acquired in the ordered region of the Ni-Al-Cr sample (Fig. 4), demonstrating that Cr sits preferentially at the Al sites in Ni_3Al . This result is consistent with experimental observations in Ref. [51] and agrees with our prediction. While the DFT calculations are performed for 0 K, the experimental agreement indicates that the DFT and informatics predictions are reasonable for materials at operational temperatures. Therefore, we expect that the prediction of Ce site occupancy is valid at operational temperatures as well.

As a further testing of the relationship of our mapping with bulk modulus, we compare the change in PC space based only on DOS input (Fig. 3) with the change in moduli listed in Table 1. From this the following results are found: (i) Cr prefers Al site and modulus increases with addition of Cr on Al site and decreases with addition on Ni site; (ii) Zr prefers Al site and modulus has larger decrease with Zr added to Ni site; (iii) Ce can occupy either site and modulus does not decrease with Ce addition to either site; and (iv) La decreases stability when added to either site and modulus decreases significantly with increased Al to either site. Therefore, we have demonstrated here that by mining only the DOS, we predict the site occupancy (prediction matching experiment) and also the relationship of site occupancy-chemistry-bulk modulus relationship (prediction matching DFT).

4. Discussion

As was shown in Figs. 3 and 4 (and as further highlighted in Fig. 5a), the Cr atoms prefer Al site occupancy. While this conclusion could be made solely from DFT calculations, we apply informatics in order to extract signature of the DOS corresponding to simultaneous change in chemistry and property. This added information which is not readily available from solely the DFT output accelerates alloy design by reducing the complex DFT output to a limited number of key features or parameters. By interrogating the corresponding eigenspectra from the spectral decomposition (Fig. 5b), we can extract these signatures. The eigenspectra define the primary patterns within the DOS spectra based on maximizing the information captured. As the DOS spectra were

mean centered prior to analysis, if each respective eigenspectra is multiplied by the corresponding scores parameter and then summed, the result is the initial mean centered spectra. Therefore, those features in the eigenspectra with the most significant magnitudes are the most important for defining the axes of Fig. 5a. For PC1, the peak corresponding with the f-states is the most significant. This matches with the interpretation that PC1 is capturing primarily effects due to Ce doping, as Ce is the only element considered here with f-electrons. For purposes of defining site occupancy in Ni_3Al -Cr, PC2 is the critical axis to consider. When increasing the Cr concentration at the Al sites, the PC2 value continues to increase. However, when increasing the Cr concentration at the Ni sites, the PC2 value does not change significantly. To understand why Cr occupies the Al site, we must identify why PC2 does not change with increasing Cr content in terms of the DOS spectra.

To link this mapping with mechanical properties of site occupancy, a model was developed for predicting bulk modulus as a function of PC values. The resulting equation was: $B_m = f(-0.003*PC1 - 0.17*PC2)$. This modulus value was obtained by considering only doped systems, and not including Ni_3Al . Since the coefficient of both PC1 and PC2 is negative (i.e. -0.003 and -0.17), B is inversely correlated with both parameters. Also, given the relative values of the coefficients, we find that PC2 is more significant for determining B_m , i.e. a change in PC2 will affect B_m more than an equivalent change in PC1. From this result, we identify that increasing Ce composition on the Ni site has a larger effect on B_m than increasing Ce composition on the Al site. This is not evident when looking at Table 1. This map therefore integrates dopant chemistry, site substitution, and mechanical property. The findings associated with Ce doping are that Ce can occupy either Ni or Al site, while the modulus is more sensitive to Ce on Ni site.

As there are multiple features present in the PC2 spectrum, it is instructive to compare with the relevant DOS spectra (Fig. 6). To define the change in DOS spectra with Cr concentration, the DOS values at each energy for $Ni_{23}Al_8Cr$ were subtracted from those for $Ni_{21}Al_8Cr_3$, and also for $Ni_{24}Al_7Cr$ from $Ni_{24}Al_5Cr_3$. In this representation of the DOS, we identify two notable peaks for Cr in the Al site. These two peaks, centered at approximately -1.7 eV and 0.2 eV (where this energy is defined as energy minus E_F) correspond to high intensity regions in the loadings spectra, which means that changes in the DOS at these energies will impact the PC2 values. Therefore, the difference in DOS at these energies results in a linear relationship between PC2 value and Cr composition for Al substitution and the lack of a relationship for Ni substitution. The reason that Cr goes to the Al site in Ni_3Al is because the DOS at -1.7 eV and 0.2 eV is insensitive to Cr addition at the Ni site but is sensitive to Cr addition at the Al site. In Figs. 3 and 5, we found that site occupancy is based on linear relationship between PC values and dopant concentration. As the Cr doped samples change along PC2, in order for the relationship to remain linear, the DOS must vary at energies corresponding to those with highest loadings values. For Cr substituted on Al site, we see large difference in DOS with increasing Cr content at the circled energies. However, for Cr substituted on Ni site, there is little change in DOS at those energies. The lack of changing DOS at those two energy ranges with increasing Cr content is the reason that Cr does not occupy the Ni site.

In order to simplify the visualization of the chemistry-property correlations, we developed a dendrogram which defines this similarity

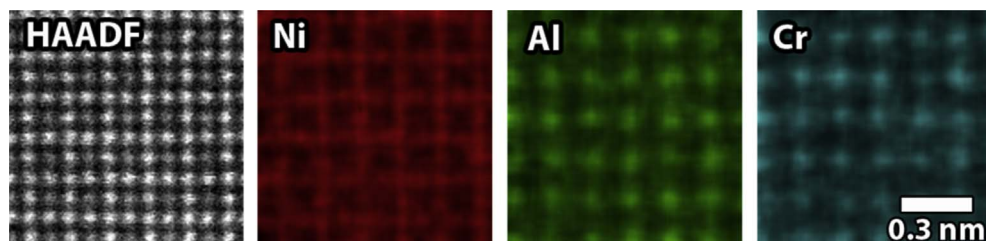


Fig. 4. EDS elemental mapping of Ni_3Al -Cr in the $\langle 100 \rangle$ projection. This result shows Cr having preferred Al substitution in Ni_3Al , agreeing with our informatics-based prediction.

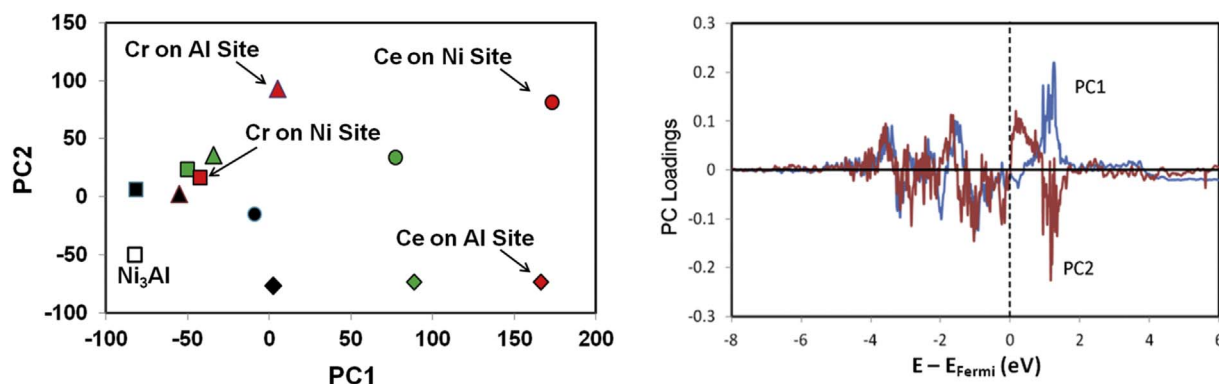


Fig. 5. This analysis corresponds with the analysis of Fig. 3 but showing fewer compounds for clarity (a) and for showing how to interpret the loadings spectra (b). Increased DOS values at energies corresponding to positive loadings will increase the respective PCs, while the inverse case is true for negative loadings values. PC1 primarily captures changes due to increasing Ce composition, with the primary feature of the PC1 eigenspectra corresponding to the f-state. PC2 defines why Cr occupies only Al sites and also describes the change in bulk modulus.

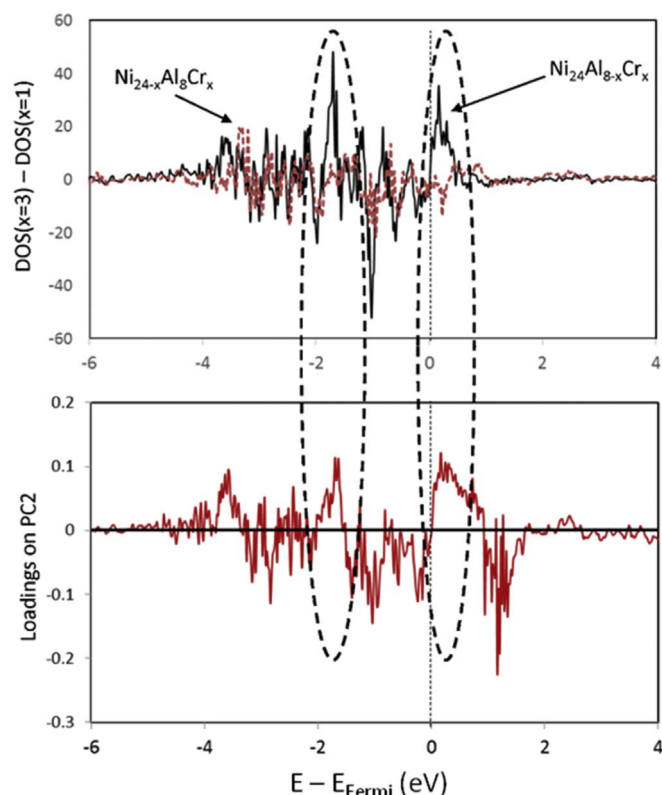


Fig. 6. Comparing eigenspectrum with DOS as function of changing site occupancy for understanding why Cr occupies Al site. The circled regions correspond with the energies responsible for the preferred Al site substitution for Cr.

(Fig. 7). The input was the change in PC space with increased dopant concentration and the resulting elastic constants. This dendrogram therefore defines similarity between compounds based solely on the chemistry-property correlation as derived from the DOS. As the connection of branches moves up in the dendrogram, the correlation decreases. For example, we find that the similarity in chemistry-property relationship is highest for Cr on an Al site and Ce on an Al site. These are next most similar to Zr on a Ni site. Following this logic, Zr shows the least site occupancy effect of any dopant, with La also having a small occupancy effect. In terms of identifying chemical substitutes for rare earth elements, we find that Cr on a Ni site is on the same branch as La on either site and Ce on a Ni site. This therefore provides a design guideline for replacing rare earth elements, although as discussed Cr preferentially occupies the Al site. However, the strategy for replacing the rare earth element with minimal impact on property is to identify

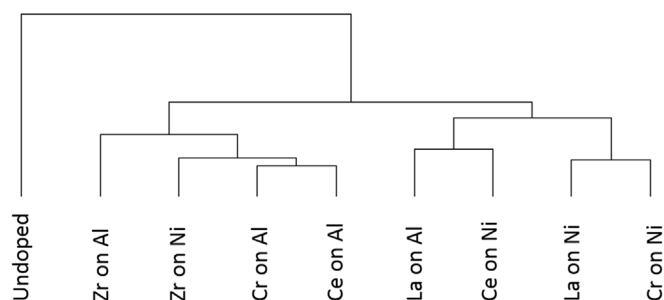


Fig. 7. Dendrogram comparing similarity between dopants at specific sites in terms of the effect on the DOS spectra and the elastic moduli. The inputs into developing the dendrogram are the change in PC positions in Fig. 3 with added dopant concentration, and the resulting elastic constants. The lines connecting branches nearer the bottom of the figure indicate higher degree of correlation (eg. Cr on Al site and Ce on Al site are most correlated, while there is minimal correlation between undoped Ni₃Al and doped Ni₃Al. From this, we find that Cr has the most similarity to rare earth dopants, while Zr and La are the least impacted by site occupancy.

an element with similar characteristics to Cr and which may occupy a Ni site, or else to identify a combination of elements to modify the couple critical signatures of the DOS associated with Cr site occupancy which we identified. Future work will incorporate elements which preferentially occupy Ni site, based on this identified chemical modification strategy.

5. Conclusions

In this paper, we developed an approach for identifying the signatures of the DOS which capture the correlation between site occupancy energetics and elastic moduli. From these signatures, we assessed the similarity between different chemistries of $\text{Li}_2\text{-Ni}_3\text{Al}$. As an example of the application of this work for chemical design, we assessed the similarity in the chemistry-property correlation between elements in order to identify potential rare earth replacements. We find Cr on the Ni site as the most similar to the rare earth containing compounds, although Cr on the Al site is also similar to Ce on the Al site. From our analysis, we identified that Cr preferentially occupies the Al site. In order to understand what features of the DOS spectra to modify through doping in order to leverage the similarity with rare earth elements, we extracted the key signatures at -1.7 and 0.2 eV for Cr doping. This provides guidelines for further chemical modification for targeted design. This work has demonstrated how the application of data dimensionality reduction methods can serve as an accelerated feature selection approach to interrogate DOS spectra coupled to a mechanistic interpretation of linking chemistry and site occupancy to design alloy chemistries with targeted mechanical properties.

Acknowledgements

We gratefully acknowledge support of this work under Air Force Office of Scientific Research (AFOSR) Grant No. FA9550-12-1-0456 and NSF Grant No. DMR-13-07811. SB and KR acknowledge support from NSF DIBBs Award Number ACI-1640867. KR also acknowledges support from the Erich Bloch Endowed Chair at the University at Buffalo-State University of New York. All the DFT calculations were performed using computational resources provided by the University of Florida Research Computing (<http://researchcomputing.ufl.edu/>). AAO and JML acknowledge the use of the Analytical Instrumentation Facility (AIF) at North Carolina State University, which was supported by the State of North Carolina and the National Science Foundation.

References

- [1] S.R. Broderick, K. Rajan, Eigenvalue decomposition of spectral features in density of states curves, *Europhys. Lett.* 95 (2011) 57005.
- [2] S.R. Broderick, H. Aourag, K. Rajan, Classification of oxide compounds through data-mining density of states spectra, *J. Am. Ceram. Soc.* 94 (2011) 2974.
- [3] S.R. Broderick, H. Aourag, K. Rajan, Data mining density of states spectra for crystal structure classification: an inverse problem approach, *Stat. Anal. Data Min.* 6 (2009) 353–360.
- [4] R.C. Reed, *The Superalloys Fundamentals and Applications*, Cambridge University Press, Cambridge, 2006.
- [5] D.E. Kim, S.L. Shang, Z.K. Liu, Effects of alloying elements on elastic properties of Ni3Al by first principles calculations, *Intermetallics* 18 (2010) 1163–1171.
- [6] Q. Wu, S. Li, Alloying element additions to Ni3Al: site preferences and effects on elastic properties from first-principles calculations, *Comput. Mater. Sci.* 53 (2012) 436–443.
- [7] A. Kumar, A. Chernatynskiy, M. Hong, S.R. Phillpot, S.B. Sinnott, An ab initio investigation of the effect of alloying elements on the elastic properties and magnetic behavior of Ni3Al, *Comput. Mater. Sci.* 101 (2015) 39–46.
- [8] A.V. Ruban, H.L. Skriver, Calculated site substitution in γ -Ni3Al, *Solid State Commun.* 99 (1996) 813–817.
- [9] C. Jiang, B. Gleeson, Site preference of transition metal elements in Ni3Al, *Scr. Mater.* 55 (2006) 433–436.
- [10] C. Jiang, D.J. Sordet, B. Gleeson, Site preference of ternary alloying elements in Ni3Al: a first-principles study, *Acta Mater.* 54 (2006) 1147–1154.
- [11] C. Wagner, W. Schottky, Theorie der geordneten Mischphasen, *Z. Phys. Chem. B* 11 (1930) 163.
- [12] Z. Yuan, S. Song, R.G. Faulkner, Z. Yu, Combined effects of cerium and boron on the mechanical properties and oxidation behaviour of Ni3Al alloys, *J. Mater. Sci.* 33 (1998) 463–469.
- [13] M.J. Donachie, *Superalloys: a Technical Guide*, second ed., ASM International, 2002.
- [14] Y. Zhaxi, S. Shenhua, X. Tianhui, Z. Xiju, Y. Zongsun, Influence of cerium on ductility of Ni3Al, *J. Mater. Sci. Lett.* 13 (1994) 1717–1719.
- [15] H.-E. Huang, C.-H. Koo, Effect of zirconium on microstructure and mechanical properties of cast fine-grain CM 247 LC superalloy, *Mater. Trans.* 45 (2004) 554–561.
- [16] P.J. Zhou, J.J. Yu, X.F. Sun, H.R. Guan, Z.Q. Hu, Roles of Zr and Y in cast microstructure of M951 nickel-based superalloy, *Trans. Nonferr. Metal. Soc. China* 22 (2012) 1594–1598.
- [17] D.M. Anliker, J.B. Newkirk, The effects of Cerium on the microstructure of INCO 901 superalloy, *Metall. Trans. A* 7A (1976) 1711–1718.
- [18] C.C. Fischer, et al., Predicting crystal structure by merging data mining with quantum mechanics, *Nat. Mater.* 5 (2006) 641.
- [19] G.H. Jóhannesson, et al., Combined electronic structure and evolutionary search approach to materials design, *Phys. Rev. Lett.* 88 (2002) 255506.
- [20] M. D'Avenzac, A. Zunger, Identifying the minimum-energy atomic configuration on a lattice: lamarckian twist on Darwinian evolution, *Phys. Rev. B* 78 (2008) 064102.
- [21] M. de Jong, et al., A statistical learning framework for materials science: application to elastic moduli of k-nary inorganic polycrystalline compounds, *Sci. Rep.* 6 (2016) 34256.
- [22] M. de Jong, et al., Charting the complete elastic properties of inorganic crystalline compounds, *Sci. Data* 2 (2015) 15009.
- [23] J.R. Alvarez, P. Rez, Calculation of electronic properties of boundaries in Ni3Al, *Acta Mater.* 49 (2001) 795.
- [24] W. Zhou, H. Wu, T. Yildirim, Electronic, dynamical and thermal properties of ultra-incompressible superhard rhenium diboride: a combined first-principles and neutron scattering study, *Phys. Rev. B* 76 (2007) 184113.
- [25] S.F. Matar, M.A. Subramanian, Calculated electronic properties of the mixed perovskite oxides: CaCu3Ti4O12 (T = Ti, Cr, Mn, Ru) within the DFT, *Mater. Lett.* 58 (2004) 746.
- [26] L. Cheng-Bin, et al., First principles study on the charge density and the bulk modulus of the transition metals and their carbides and nitrides, *Chin. Phys. Lett.* 14 (2005) 2287.
- [27] P. Caron, O. Lavigne, Recent studies at Onera on superalloys for single crystal turbine blades, *Aersosp. Lab.* 3 (2011) 1–14.
- [28] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* 54 (1996) 11169.
- [29] G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.* 6 (1996) 15–50.
- [30] G. Kresse, J. Hafner, Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium, *Phys. Rev. B* 49 (1994) 14251.
- [31] G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals, *Phys. Rev. B* 47 (1993) 558.
- [32] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* 77 (1996) 3865.
- [33] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* 78 (1997) 1396.
- [34] P.E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* 50 (1994) 17953.
- [35] H.J. Monkhorst, J.D. Pack, Special points for Brillouin-zone integration, *Phys. Rev. B* 13 (1976) 5188–5192.
- [36] S.R. Broderick, C. Suh, J. Provine, C.S. Roper, R. Maboudian, R.T. Howe, K. Rajan, Application of principal component analysis to a full profile correlative analysis of FTIR spectra, *Surf. Interface Anal.* 44 (2012) 365–371.
- [37] S.R. Broderick, J.R. Nowers, B. Narasimhan, K. Rajan, Tracking chemical processing pathways in combinatorial polymer libraries via data mining, *J. Comb. Chem.* 12 (2010) 270–277.
- [38] M. Ashton, R.G. Hennig, S.R. Broderick, K. Rajan, S. Sinnott, Computational discovery of stable M(2)AX phases, *Phys. Rev. B* 94 (2016) 054116.
- [39] A. Daffertshofer, C.J. Lamoth, O.G. Meijer, P.J. Beek, PCA in studying coordination and variability: a tutorial, *Clin. Biomech.* 19 (2004) 415–428.
- [40] L. Ericksson, T. Byrne, E. Johansson, J. Trygg, C. Vikstrom, *Multi- and Megavariate Data Analysis: Principles, Applications, Umetrics Ab, Umea*, 2001.
- [41] H. Berthiaux, V. Mosorov, L. Tomczak, C. Gatamel, J.F. Demeyre, Principal component analysis for characterising homogeneity in powder mixing using image processing techniques, *Chem. Eng. Process.* 45 (2006) 397–403.
- [42] S. Wold, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of chemometrics, *Chemom. Intell. Lab. Syst.* 58 (2001) 109.
- [43] D.V. Nguyen, D.M. Rocke, Tumor classification by partial least squares using microarray gene expression data, *Bioinformatics* 18 (2002) 39–50.
- [44] R. Rosipal, N. Kramer, C. Saunders (Ed.), *Overview and Recent Advances in Partial Least Squares, in Subspace, Latent Structure and Feature Selection Techniques*, Springer, Berlin/Heidelberg, 2006, pp. 34–51.
- [45] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, *Anal. Chim. Acta* 185 (1986) 1–17.
- [46] S. de Jong, SIMPS, an alternative approach to partial least squares regression, *Chemom. Intell. Lab. Syst.* 18 (1993) 251–263.
- [47] A. Phatak, S.D. Jong, The geometry of partial-least squares, *J. Chemom.* 11 (1997) 311–338.
- [48] P.V. Balachandran, S.R. Broderick, K. Rajan, Identifying the ‘inorganic gene’ for high-temperature piezoelectric perovskites through statistical learning, *Proc. R. Soc. A* 467 (2011) 2271–2290.
- [49] O. Wodo, S. Broderick, K. Rajan, Microstructural informatics for accelerating the discovery of processing-microstructure-property relationships, *MRS Bull.* 41 (2016) 603–609.
- [50] S. Broderick, K. Rajan, Informatics derived materials databases for multifunctional properties, *Sci. Technol. Adv. Mater.* 16 (2015) 013501.
- [51] A.A. Oni, X. Sang, S.V. Raju, S. Dumpala, S. Broderick, A. Kumar, S. Sinnott, S. Saxena, K. Rajan, J.M. LeBeau, Large area strain analysis using scanning transmission electron microscopy across multiple images, *App. Phys. Lett.* 106 (2015) 011601.