

# Accelerated Modeling of Lithium Diffusion in Solid State Electrolytes using Artificial Neural Networks

Karun K. Rao<sup>1,3</sup>, Yan Yao<sup>2,3</sup>, Lars C. Grabow<sup>1,3\*</sup>

<sup>1</sup>Department of Chemical and Biomolecular Engineering

<sup>2</sup>Department of Electrical and Computer Engineering

<sup>3</sup>Texas Center for Superconductivity

University of Houston, Houston, TX 77004, USA

Email: [grabow@uh.edu](mailto:grabow@uh.edu)

**Keywords:** Machine Learning, Artificial Neural Networks, Solid State Electrolyte, Ionic Diffusion, Molecular Dynamics

## Abstract

There is great interest in solid state lithium electrolytes to replace the flammable organic electrolyte for an all solid state battery. Previous efforts trying to understand the structure-function relationships resulting in high ionic conductivity materials have mainly relied on *ab initio* molecular dynamics. Such simulations, however, are computationally demanding and cannot be reasonably applied to large systems containing more than a few hundred atoms. Herein, we investigate using artificial neural networks (ANN) to accelerate the calculation of high accuracy atomic forces and energies used during molecular dynamics (MD) simulations, to eliminate the need for costly *ab initio* force and energy evaluation methods, such as density functional theory (DFT). After carefully training a robust ANN for four and five element systems, we obtain nearly identical lithium ion diffusivities for Li<sub>10</sub>GeP<sub>2</sub>S<sub>12</sub> (LGPS) when benchmarking the ANN-MD

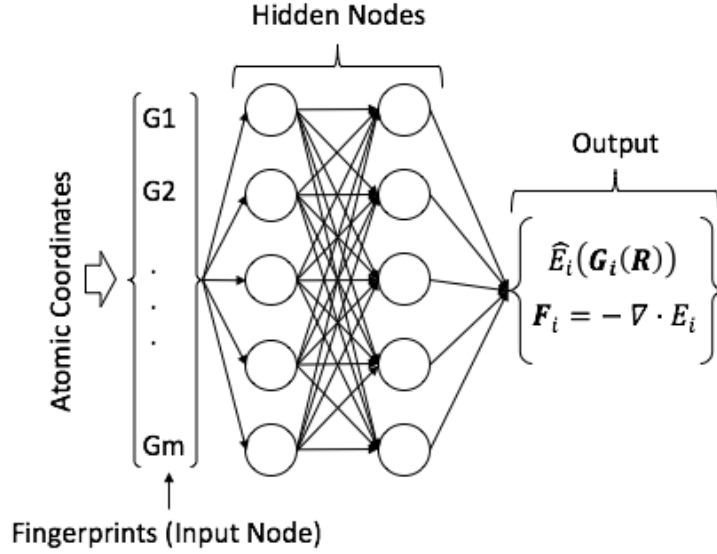
results with DFT-MD. To demonstrate the power of the outlined ANN-MD approach we apply it to a doped LGPS system to calculate the effect of concentrations of chlorine on the lithium diffusivity at a resolution that would be unrealistic to model with DFT-MD. We find that ANN-MD simulations can provide the framework to study systems that require a large number of atoms more efficiently while maintaining high accuracy.

## 1. Introduction

In the pursuit of higher energy dense batteries, the use of solid state electrolytes (SSE) has gained much attention recently to develop an all solid state battery. This electrolyte is not only safer than the conventional flammable organic electrolyte in a lithium-ion battery, but could enable the use of metallic lithium as the anode which would increase both specific energy and capacity of the battery.<sup>[1-6]</sup> One well studied superionic conductor,  $\text{Li}_{10}\text{GeP}_2\text{S}_{12}$  (LGPS), has an ionic conductivity of  $1.2 \text{ mS cm}^{-1}$ , which approaches the conductivity of traditional ether based organic liquid electrolytes.<sup>[7]</sup> Understanding the underlying mechanisms of the high conductivity could enable the design of even higher conductivity materials that could be used in commercial applications.

The first mechanistic analysis of diffusion in LGPS showed through *ab initio* density functional theory (DFT) molecular dynamics (MD) simulations that LGPS is a three-dimensional superionic conductor with lithium primarily diffusing along the *c*-axis and minimal (but non-zero) diffusion in the *ab* plane.<sup>[8]</sup> Similar first principle studies have been performed on other superionic lithium conductors, and electrolytes to study diffusion mechanisms.<sup>[9-14]</sup> The main limitation of these DFT based studies is that simulations are run at high temperatures to reduce the MD simulation times and are usually performed on one unit cell due to the expense of the DFT based force calculations. Although useful, DFT-MD simulations cannot be used to probe effects that require simulating several hundred atoms.

One possible method to accelerate MD simulations and reduce the associated computational cost is to replace the DFT based force calculation with a force field generated from an artificial neural network (ANN). The first ANN was described mathematically in 1960 and has since been used extensively in many tasks such as image recognition, data mining and other



**Figure 1.** Artificial Neural Network diagram. From left to right, the atomic coordinates uniquely determine the input (fingerprint) node. Input node values are then passed through several hidden layers which are combined to determine the atomic contribution to the energy. The topology shown is (5,5): two layers, each with 5 nodes.

complex tasks.<sup>[15]</sup> In material science, ANNs and other machine learning models are also applied to further data-mining of DFT data and to more efficiently predict material properties.<sup>[16–18]</sup> Comprehensive overviews of the implementation of ANNs used to calculate atomic energies and forces are available in the literature<sup>[19–22]</sup>, and only a brief summary will be presented here.

The goal of an ANN trained in an ‘atom centered’ mode is to model the potential energy surface of a structure by estimating the energy contributions from each atom. Given a training image of atomic coordinates the energy is predicted by Equation 1<sup>[23]</sup>:

$$E = \sum_{i=1}^N \hat{E}_i(\mathbf{G}_i(\mathbf{R}_i)) \quad (1)$$

where  $N$  is the number of atoms in the image and  $\mathbf{G}_i(\mathbf{R}_i)$  represents a feature vector of the given input atomic coordinates. This feature vector (or fingerprint) is a map of the coordinates based on a neighbor list of each atoms within a cutoff radius. Fingerprints are passed to the ANN of several hidden layers and nodes to determine the atom contribution to the total energy,  $\hat{E}_i$  (**Figure 1**) and

forces on each atom that are consistent with the gradient of the potential energy surface. It is important to state that hidden nodes in the artificial neural network may have no physical meaning. The optimal weights for the ANN are found by training, or a regression of the weights to minimize errors from energies and forces predicted by the ANN and those calculated from other sources, such as DFT.

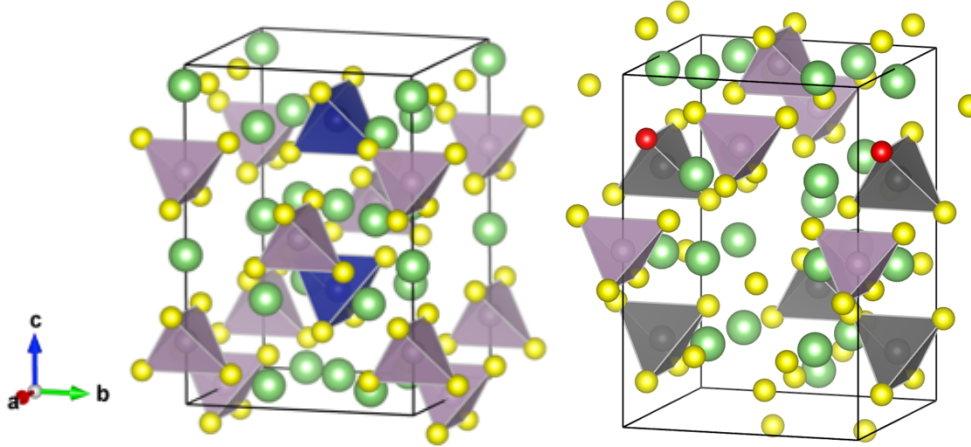
Implementation of ANNs for modeling atomistic simulations have only recently been explored as increased computational capacity has enabled the generation of a large, consistent, sets of training data. ANN based force fields for crystalline and other periodic systems have largely been limited to two or three atomic species<sup>[24]</sup> and are used mainly to predict the energy of structures rather than forces on individual atoms.<sup>[25–27]</sup> As far as the authors are aware, there have been no neural networks trained to five element systems and have only been sparsely used to directly study kinetic phenomena.<sup>[28,29]</sup>

Herein we propose using ANNs to reduce the need for computationally expensive DFT methods to perform MD simulations. We train the machine learning algorithm by using a feed-forward ANN to generate force fields for the elements in LGPS as a model system. After training a robust neural network we calculate nearly identical diffusivities when benchmarking ANN-MD compared to DFT-MD. We also compare the computational cost of a force calculation from the ANN versus *ab initio* DFT to highlight the benefit ANNs display in large systems. By taking advantage of linear ANN scaling we extended the use of these networks to study doping in a related LGPS material ( $\text{Li}_{9.54}\text{Si}_{1.74}\text{P}_{1.44}\text{S}_{11.7}\text{Cl}_{0.3}$ ) by using a trained ANN to predict optimal dopant concentrations that could not have previously been studied using traditional DFT methods.

## 2. Methods

### 2.1 DFT Simulations

Spin paired DFT calculations were performed using the Vienna ab initio Simulation Package (VASP).<sup>[30–33]</sup> The projector augmented wave (PAW) method with the Perdew-Burke-Ernzerhof (PBE) functional were used to solve the Kohn-Sham equations. All calculations were performed with a planewave cutoff energy of 600 eV to ensure high accuracy force and dielectric constant calculations. Partial occupancies near the Fermi level were determined by Gaussian smearing<sup>[34]</sup> with a Fermi temperature of  $k_b T = 0.2$  eV. The Brillion zone sampling was Gamma centered with a  $k$ -point mesh for LGPS and the doped-LGPS of  $2 \times 2 \times 3$  for one unit cell. A  $1 \times 1 \times 1$   $k$ -point mesh was used for all larger supercell calculations. Initial crystal structures were obtained from the materials project database and are shown in **Figure 2**.<sup>[35]</sup>



**Figure 2.** Crystal structures of (A) LGPS and (B)  $\text{Li}_{20}\text{Si}_3\text{P}_3\text{S}_{23}\text{Cl}$ . Green, yellow, and red are lithium, sulfur and chlorine, respectively, and purple, blue, and grey represent  $\text{PS}_4$ ,  $\text{GeS}_4$ , and  $\text{SiS}_4$  tetrahedra, respectively. The total number of sulfur tetrahedra in both materials is the same, as is the number of lithium atoms in each unit cell.

## 2.2 Molecular dynamics

Diffusion characteristics in all simulated materials were predicted using molecular dynamics simulations in the canonical  $NVT$  ensemble.<sup>[8]</sup> *Ab initio* molecular dynamics simulations were performed in VASP, while the molecular dynamics using neural networks for force calculations were performed with the Atomic Simulation Environment (ASE).<sup>[36]</sup> The temperature

of the canonical ensemble was controlled using the Nosé-Hoover thermostat implemented in the Parrinello-Rohman methods.<sup>[37–39]</sup> The simulation flow was as follows:

1. Assign all atoms random initial velocities at  $T = 100$  K according to a Maxwell-Boltzmann distribution
2. Raise the temperature to the target by linear velocity scaling over 4 ps
3. Allow the system to equilibrate at the final temperature for 1 ps
4. Run the  $NVT$  ensemble using the Nosé-Hoover thermostat for 20 ps

The time step chosen for all simulations was 2 fs. All molecular dynamics simulation trajectories were analyzed using the diffusion analyzer module in the Pymatgen library,<sup>[40]</sup> and pair distribution functions were calculated using VMD.<sup>[41]</sup>

## 2.3 Artificial Neural Networks

Training and construction of an artificial neural network and fingerprint analysis was done using the Atomistic Machine Learning Package (AMP).<sup>[42]</sup> The descriptors, and neural network topology were varied and tested to optimize the ANN (see Results section), but all networks used a cutoff radius of 6.5 Å, and employed default cosine cutoff functions, tanh activation functions, and loss function weights. Regression was done for each neural network training using the Broyden–Fletcher–Goldfarb–Shanno algorithm.<sup>[43]</sup>

## 3. Results and Discussion

### 3.1 Training

The most intensive task in using an ANN is the training, which, if done properly, will create a well-conditioned model that can predict the same forces and energies as found in DFT. Training data was generated by running a DFT-MD simulation of LGPS as described in the Methods

section, but allowing for large temperature fluctuations to provide a wide space of atomic structures for training. This simulation generated 12,000 training images of which 80% were randomly selected for training the ANN, and the other 20% were used to measure the test error. The training data was further split with 10% assigned to the validation set.

**Table 1.** Descriptor/fingerprint comparison between Zernike and Gaussian descriptors with varying number of symmetry functions. The number of functions corresponds to the value of  $m$  in the  $Gm$  vector of Figure 1. Distribution of outputs for (A) and (B) are given in **Figure S1**.

Descriptor	Number of Symmetry Functions	Topology	Force RMSE
Zernike (A)	12	(5,5)	0.4410
Zernike	20	(5,5)	0.2100
Zernike	49	(5,5)	0.1933
Gaussian (B)	56	(5,5)	0.0742
Gaussian	72	(5,5)	0.0822

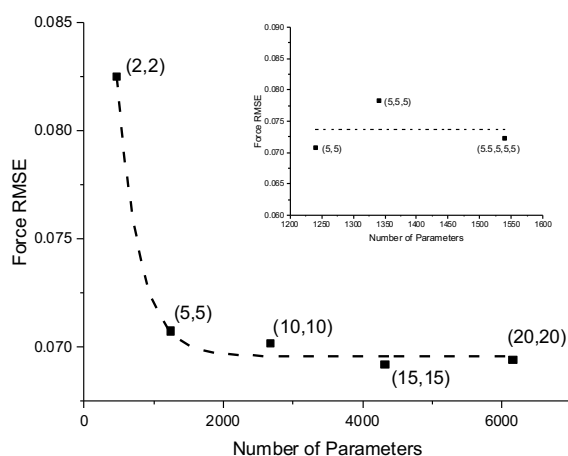
The selection of descriptors is crucial in developing a robust neural network model and to determine which interactions are relevant.<sup>[44]</sup> Both the Zernike and Gaussian descriptors are sensitive to 2 and 3 atom interactions, while only the Zernike descriptor depends on 4 atom dihedral angles.<sup>[42]</sup> The network topology used to compare descriptors was a (5,5) structure with two hidden layers each with 5 nodes. We compare the root mean squared error for forces across all training images (Training Force RMSE) while varying the number of symmetry functions. The Gaussian methodology of describing the atomistic neighborhood with 76 symmetry function resulted in the lowest error (**Table 1**) but takes double the time to compute due to the exponential scaling of calculating descriptors. Increasing the number of symmetry functions in either the Zernike or Gaussian descriptors will continue to decrease errors, but at diminished utility and with



increasing computational cost to calculate the increased number of symmetry functions. Individual node analysis for each descriptor component showed a wide variety of outputs, with minimal redundancies between nodes (**Figure S1**).

Although the universal approximation theorem states that an arbitrary large hidden layer structure can fit any data to arbitrary accuracy,<sup>[45]</sup> increasing the complexity of the topology increases both the computational expense of training and increases the probability of overfitting. Initially, a Gaussian-56 descriptor was used and two hidden layers while varying the number of nodes per layer. By systematically increasing the number of nodes per layer, we find for the LGPS system, 5 nodes per layer to be the optimal where increasing the number of nodes beyond 5 yields little to no further decrease in the overall training error (**Figure 3**). We subsequently increase the number of layers fixing the number of nodes per layer and find that 2 hidden layers sufficiently minimizes the training error.

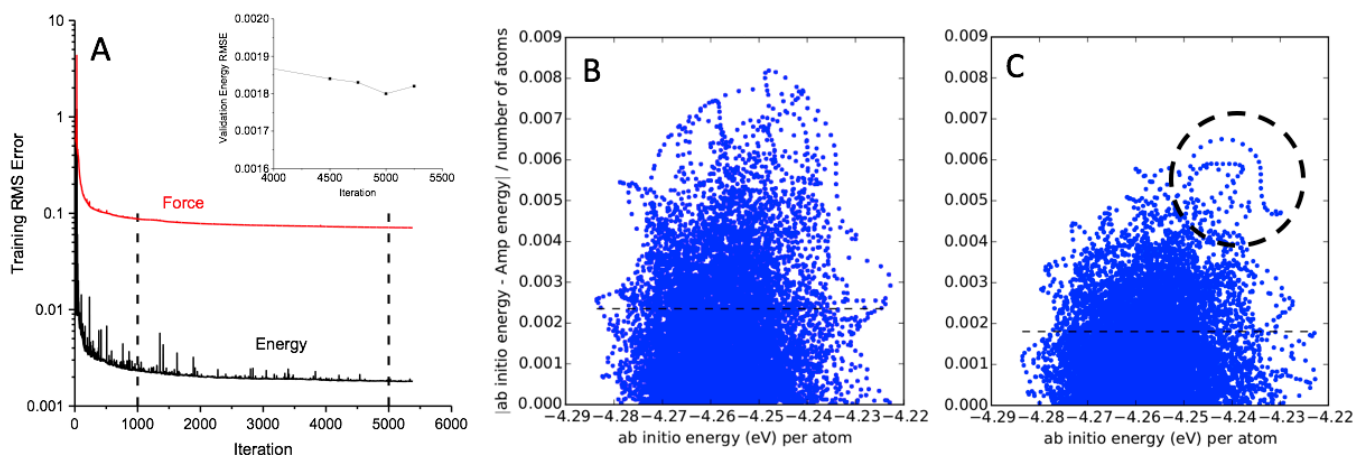
For a given topology there is an optimal training that minimizes both training and validation error.<sup>[23]</sup> When we carefully optimized the residual force and energy errors, we observed an initial exponential decrease as a function of training iteration with most of the minimization in



**Figure 3.** Impact of topology on training error. When increasing the number of nodes per layer there is a diminishing effect on error past five nodes. Additionally, there is no apparent correlation between number of layers and error. Lines added as a guide.

error occurring in the first 1,000 steps. After 1,000 steps, the rate of error minimization decreases (Figure 4a). We select different models along the regression and measure the validation error, that is, the error of the other 960 images in the validation set. The energy validation error decreased until reaching a minimum value of 0.0018 eV after the 5,000<sup>th</sup> iteration. The validation error increased slightly to 0.0019 eV after the 5,300<sup>th</sup> iteration suggesting that continued regression would have led to the ANN becoming over fit.

The energy residuals of the validation set at different iterations of the regression also provide some additional insight into the quality of the training. With a well fit model, the residuals are expected to be normally distributed around the mean, whereas observed patterns suggest the model is biased or missing physically relevant descriptors. After 1,000 iterations, the energy error residuals are generally evenly distributed, but there appear to be points that form clear curves and patterns that occur at all ranges of energies (Figure 4b). After increased training (5,000 iterations), the residual errors of those same points decrease (i.e. the whole distribution is ‘squished’ down), but the curved patterns are still present, especially at the two energy value extremes (Figure 4c).



**Figure 4.** (A) Training and validation errors during training of a (5,5) Gaussian neural network. The inset shows the validation error near the end of training with a slight increase after the 5,000<sup>th</sup> iteration (B) and (C) show the residuals at iteration 1,000 and 5,000, respectively. One pattern is highlighted to show what images could be explicitly included in future training iterations. Dashed black lines show root mean squared error.

These patterns in the residuals may be explained as under sampled geometries in training since only a small fraction of the validation geometries (< 5%) form the residual patterns.

For the LGPS system, we chose the optimum network at the 5,000<sup>th</sup> iteration because it had both the lowest validation energy and force errors of 0.0018 eV and 0.077 eV Å<sup>-1</sup>, respectively. The final testing error of the network was similar to that of the validation with energy and force errors of 0.0022 eV and 0.112 eV Å<sup>-1</sup>. The absolute magnitude of the error is lower than the uncertainty in the original DFT data, and are not expected to impact the calculation of thermodynamic properties which are averaged over the course of an MD simulation.

### 3.2 Validation of ANN-MD Predicted Diffusivity

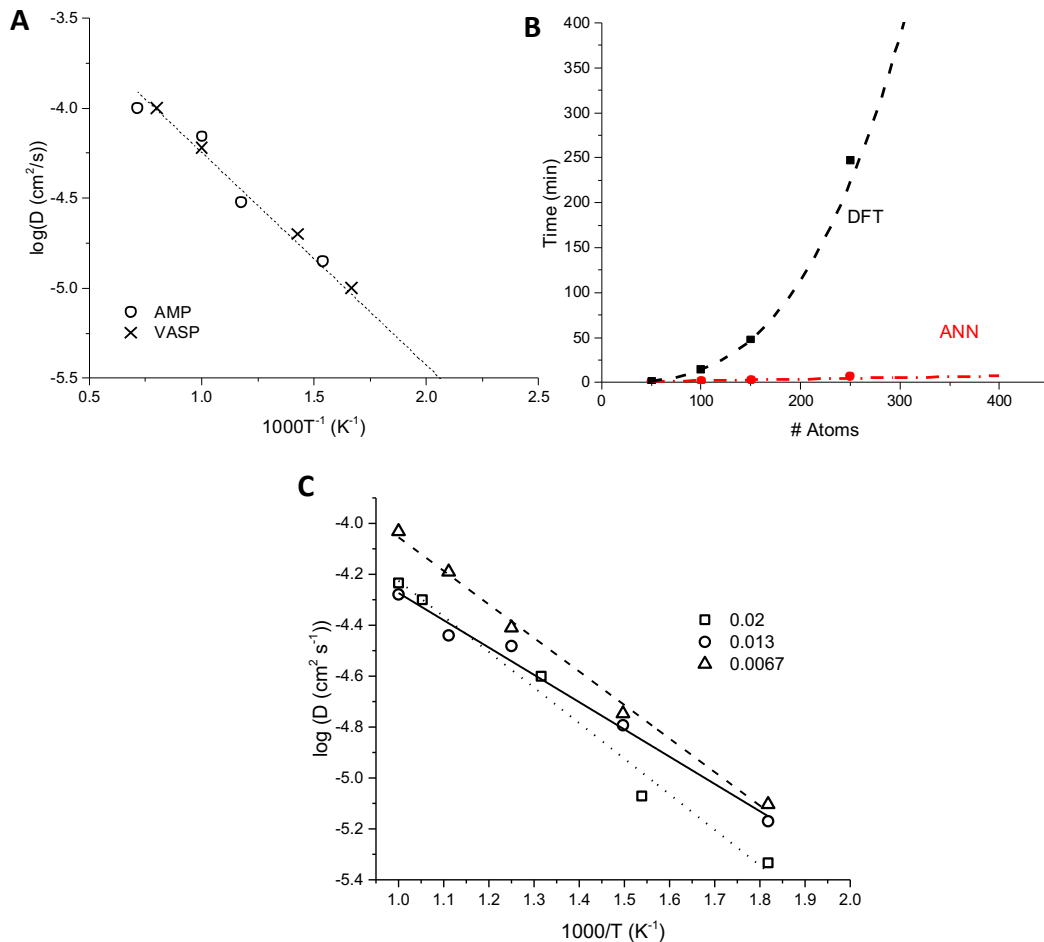
To validate the applicability and accuracy of the trained ANN for calculating temperature dependent Li diffusivities in LGPS, we performed MD simulations at elevated temperatures with both DFT and ANN generated forces. High temperatures were selected to enable faster convergence. We extract the self-diffusivity coefficient from the averaged mean squared displacement of lithium atoms along the MD trajectory according to Equation 2:

$$D = \frac{1}{2dt} \frac{1}{N} \sum_i [r_i(t + t_o) - r_i(t_o)]^2 \quad (2)$$

with errors calculated from the uncertainty of the linear fit to the averaged mean square displacement. The diffusivity in equation 2 is related to the conductivity by the Nernst-Einstein relationship in Equation 3:

$$\sigma = \frac{F^2 c D}{RT} \quad (3)$$

The comparison between Arrhenius data from DFT and the ANN show remarkable agreement (**Figure 5a**) with similar predicted activation energies (0.28 eV from DFT, and 0.26 eV from the ANN), and extrapolated room temperature conductivities (10 mS cm<sup>-1</sup> from DFT and



**Figure 5.** (A) Arrhenius plot comparing the diffusivity behavior predicted between DFT-MD (VASP) and ANN-MD (AMP). (B) Time per force calculation as a function of system size. (C) Arrhenius plot for  $\text{Li}_{20}\text{Si}_3\text{P}_3\text{S}_{23}\text{Cl}$  system at various atomic concentrations of chlorine.

8.8mS  $\text{cm}^{-1}$  from ANN) supporting the use of the ANN model to replace DFT as the force calculator in MD simulations.

### 3.3 Computational Cost Advantage of Neural Networks

Increasing the number of atoms in the simulation reveals the overall calculation of forces in the ANN scales linearly with the number of atoms in the system (**Figure 5b**). The scaling of the ANN is approximately linear because the cutoff radius used is of the same size as a unit cell of LGPS and will therefore show a significant time advantage over standard DFT based approaches which scale  $O(n^3)$ .<sup>[46]</sup> On a single Intel E5-2660v3 node with 20 cores and 128 GB memory, the time required for a single force calculations for 50 atoms is about 60 seconds for both DFT and

the ANN, so any efficiency gains of the ANN are only apparent when studying larger systems. ANN-MD simulations become much more useful tools when studying large systems where the physics studied cannot be accurately modeled with less than 100 atoms such as low dopant concentrations, grain boundaries<sup>[47]</sup>, or low vacancy concentrations. As a proof of concept, we study the diffusion characteristics in a related doped-LGPS material with ANN-MD which takes advantage of this scaling behavior compared to DFT-MD.

### 3.4 Substituted system

Recently, a new material,  $\text{Li}_{9.54}\text{Si}_{1.74}\text{P}_{1.44}\text{S}_{11.7}\text{Cl}_{0.3}$  has been reported showing similar and slightly higher conductivity to that of LGPS. The doped-LGPS crystal structure has Si substitution in the Ge positions, and a Cl (8g) site in the  $\text{P}(2b)\text{X}_4$  tetrahedra.<sup>[48]</sup> The exact reason why Si and Cl doping yield higher conductivity is unknown, and it is possible that varying degrees of chlorine substitution will result in different conductivity. Simulating this material would require a calculation of 100 atoms which begins to reach the limit of DFT-MD. To demonstrate the power of the outlined ANN-MD approach we apply it to this doped LGPS system to calculate the effect of concentrations of chlorine on the lithium diffusivity at a resolution that would be unrealistic to model with DFT-MD. Because this material contains different elements than LGPS, the previously trained forcefield is not transferrable and we must train a different ANN to model this material.

We ran a DFT-MD simulation of one unit cell of  $\text{Li}_{20}\text{Si}_3\text{P}_3\text{S}_{23}\text{Cl}$  from 300 – 1000 K in VASP to create 10,000 images, of which 6,000 were randomly selected to perform the training, and 4,000 were used for testing. The neural network was trained following a similar procedure outlined above, where the final ANN was defined with a Gaussian-56 descriptor with a (5,5) hidden layer topology, resulting in a final training force RMSE of  $0.052 \text{ eV } \text{\AA}^{-1}$ , and testing error of  $0.0422 \text{ eV } \text{\AA}^{-1}$ . With a single unit cell, there are 50 simulated atoms allowing chlorine

concentration changes only in 2% intervals. Increasing the number of atoms in the simulation box increases the dopant resolution we can simulate.

After training the artificial neural network model for this system, we ran ANN-MD simulations at multiple temperatures on  $\text{Li}_{20}\text{Si}_3\text{P}_3\text{S}_{23}\text{Cl}$  with decreasing concentrations of chlorine: 2%, 1.3%, and 0.67%. These concentrations correspond to three, two and one chlorine atom respectively each in a  $3 \times 1 \times 1$  supercell. The unit cell was repeated along the  $c$ -axis during simulations because it is the dominant axis of diffusion. Each chlorine removed was substituted with a sulfur atom to complete the tetrahedral arrangements, and one silicon was substituted with phosphorous for each chlorine to maintain charge neutrality. From the ANN-MD simulations, we similarly constructed Arrhenius plots for lithium ions (**Figure 5c**) to extract activation energy and extrapolated conductivities. We found that 1.3% chlorine substitution had the lowest activation energy (0.24 eV) and highest predicted room temperature conductivity of  $24.22 \text{ mS cm}^{-1}$  (range of  $9 - 53 \text{ mS cm}^{-1}$ ) which is in remarkable agreement with the experimental value of  $25 \text{ mS cm}^{-1}$ . Both 2% and 0.67% chlorine substitution had similar activation energies of 0.32 eV and 0.29 eV, respectively, corresponding to lower extrapolated room temperature conductivities of 14.35 and  $20.04 \text{ mS cm}^{-1}$ , respectively (**Table 2**).

To begin understanding the observed optima in conductivity as a function of doping concentration, we calculated an estimated activation energy of lithium site hopping using a modified Anderson-Stuart model.<sup>[49,50]</sup> The electrostatic binding energy of the transition state relative to the initial state can be estimated as Equation 4:

$$\Delta E_B = \left( -1 + \frac{2(r_{\text{cation}} + r_{\text{anion}})}{\lambda_{\text{cation}} + 2(r_{\text{cation}} + r_{\text{anion}})} \right) \frac{Z_{\text{anion}} Z_{\text{cation}} e^2}{4 \pi \epsilon_o \epsilon_r} \left( \frac{1}{r_{\text{cation}} + r_{\text{anion}}} \right) \quad (4)$$

where  $Z_{\text{anion}}$  and  $Z_{\text{cation}}$  are the partial charges on the lithium and chlorine atoms (+1 and -1 respectively),  $\epsilon_o$  and  $\epsilon_r$  are the permittivity of free space and dielectric constant, and  $e$  is the

**Table 2.** Summary and comparison of calculated diffusion parameters for Cl-doped LSiPS and LGPS. Conductivities in brackets represent 95% error intervals from error in fitting the Arrhenius slope and uncertainty in temperature measurements.

Material	Method	Activation Energy [eV]	Conductivity [mS cm <sup>-1</sup> @ 300K]	Diffusivity [cm <sup>2</sup> s <sup>-1</sup> @ 300K]
LSiPS - 2% Cl	ANN-MD	0.32	14.35 [8 - 23]	$3.6 \times 10^{-8}$
LSiPS - 1.3% Cl	ANN-MD	0.24	24.22 [9 - 53]	$2.55 \times 10^{-7}$
LSiPS - 0.67% Cl	ANN-MD	0.29	20.04 [14 - 32]	$8.00 \times 10^{-8}$
Li <sub>9.54</sub> Si <sub>1.74</sub> P <sub>1.44</sub> S <sub>11.7</sub> Cl <sub>0.3</sub> (1.2% Cl)	Experimental		25 <sup>[48]</sup>	
LGPS	DFT-MD	0.28	10 [2 - 38]	$7.88 \times 10^{-8}$
LGPS	ANN-MD	0.26	8.8 [2 - 40]	$6.82 \times 10^{-8}$
LGPS <sup>[8]</sup>	DFT-MD	0.21	9 [2 - 40]	$7.10 \times 10^{-8}$

elementary charge. The effective ionic radii of the lithium and chlorine atoms ,  $r_{cation}$  and  $r_{anion}$ , are estimated from the total molar volume in Equation 5:

$$r_{\alpha} = 2 \left( \frac{3V_m N_{\alpha}}{4\pi N_T} \right)^{\frac{1}{3}} \quad (5)$$

where  $V_m$  is the cell simulation volume,  $N_T$  is the total number of atoms in the cell, and  $N_{\alpha}$  is the number of  $\alpha$  atoms (either number of lithium or chlorine atoms). Finally,  $\lambda_{cation}$  is the average distance between stable lithium sites. The activation energy is equal to the electrostatic binding energy difference plus the strain energy of the transition state. In this analysis, we neglect the strain energy contribution which is generally an order of magnitude lower than the electrostatic binding energy.<sup>[51]</sup>

The ANN-MD simulations at each Cl doping concentration revealed similar average distances between stable sites at 3.25 Å, and were calculated from the first peak of the Li-Li pair

correlation function<sup>[52]</sup> averaged over the simulation (**Figure S2**). The similar distance between stable sites at all concentrations may be expected due to the equal concentrations of lithium and similar overall geometry in each case. Therefore, the different chlorine concentrations influence equation 4 only by changing the values of  $r_{anion}$  and  $\epsilon_r$ . From equation 4, increasing both  $r_{anion}$  and the dielectric constant decrease the activation energy barrier. As the chlorine concentration increases,  $r_{anion}$  increases from 0.63 to 0.79 to 0.91 Å at 0.67% 1.3% and 2% Cl, respectively. However, the DFT-calculated dielectric constants were 5.13, 5.24, and 4.92 at 0.67%, 1.3%, and 2% Cl, respectively, and do not follow a clear trend. Using these calculated values, the electrostatic activation energies were found to be 0.32, 0.28, and 0.29 eV, which are in good agreement with the observed diffusion activation energies calculated from the Arrhenius plot of the ANN-MD simulation and are summarized in **Table 3**. The competing effects of changes in the dielectric constant and increasing effective anion radius at different doping levels result in an optimal concentration in which the transition state energy is minimized and conductivity is maximized.

Although the effect of concentration on the overall conductivity behavior is small in this studied system, it nonetheless demonstrates the need to consider a wide range of concentrations of doping when modeling and designing electrolytes. Studying defects and doped systems with sufficiently fine resolution requires simulations in larger simulations cells that are not amenable to **Table 3**. Calculated theoretical electrostatic energy barriers for  $\text{Li}^+$  hopping in the doped LSiPS system obtained from Eq. 4.

Material	$\lambda_{cation}(\text{\AA})$	$r_{anion}(\text{\AA})$	$\epsilon_r$	$\Delta E_S(\text{eV})$
LSiPS - 2% Cl	3.35	0.63	5.14	0.31
LSiPS - 1.3% Cl	3.35	0.79	5.24	0.28
LSiPS - 0.67% Cl	3.35	0.91	4.92	0.29



standard DFT-MD approaches and yet these atomic scale modifications that impact the ionic conductivity require DFT level accuracy to accurately simulate. For example, it has been shown in other lithium superionic conductors that defects and vacancies<sup>[53,54]</sup> play an important role in the energetics of lithium site hopping and can significantly decrease activation energy barriers to diffusion and increase the concentration of ion carriers.<sup>[55]</sup> Probing the diffusion characteristics with the accuracy of DFT-MD calculations on such low Cl concentrations and for other atomic scale modifications would not have been feasible for so many atoms but are now becoming accessible using ANN-based force fields.

#### 4. Conclusion

We demonstrate a methodology for training robust artificial neural networks to probe diffusion coefficients in solid state electrolytes to reduce the computational expense of running *ab initio* techniques. To the best of our knowledge we report the first ANN trained to accurately predict a transport property from force training in a 5-element system which speaks to the high adaptability and future possibilities of the ANNs. Diffusion coefficients calculated from molecular dynamics simulations with forces generated from the ANN are consistent with those calculated from DFT-derived forces and are in good agreement with experimentally obtained diffusion coefficients. The overall force error between the ANN and DFT was measured to be  $\sim 0.07$  eV  $\text{\AA}^{-1}$ , and further optimization and training can allow for even lower errors, if required. Additionally, we demonstrate that ANNs can be used to find optimal dopant concentrations in concentration regimes inaccessible to conventional DFT-MD.

The main limitation of this ANN approach is that each trained model is system specific. Trained ANNs cannot accurately predict energies and forces for geometries significantly different from the training data, nor can it work with elements not included in the training data. Therefore,

ANNs are not currently good tools for predicting new structures or testing the effects of elements substitutions or doping with different elements. Training ANNs also require large data sets, but as computational resources continue to increase and considering that training the ANN is a one-time computational cost, it may soon become feasible to train with a larger cutoff radius to enable even higher accuracy calculations.

We expect this methodology to be particularly useful when probing effects that can only be measured or simulated with more atoms such as effects at grain boundaries<sup>[47]</sup>, higher resolution of doping (for elements already present in the system), or viscosity.<sup>[56]</sup> Another advantage of the ability to simulate larger systems is that the simulation temperature does not need to be artificially increased to sample rare events. Although ANNs in MD simulations remain a relatively new technique, they promise to be highly computationally efficient while maintaining good accuracy. Especially as more data becomes available, it is natural to expect these ANNs to become more accurate and versatile in their applications.

## **Acknowledgements**

K.K.R. acknowledges the NASA Space Technology Research Fellowship (Grant # 80NSSC17K0148), Y.Y acknowledges the Seed Funding for Advanced Computing (SeFAC) grant from the Center for Advanced Computing and Data Systems (now Hewlett Packard Enterprise Data Science Institute) at the University of Houston, and L.C.G acknowledges partial support from the Texas Center for Superconductivity (TcSUH). Computational resources were provided by the uHPC cluster managed by the University of Houston and acquired through NSF-MRI award number 1531814. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.<sup>[57]</sup> We also acknowledge the use of computational resources provided by the National Energy Research

Scientific Computing (NERSC) Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under contract number DE-AC02-05CH11231. The authors acknowledge the use of the Maxwell/Opuntia/Sabine Cluster and the advanced support from the Research Computing Data Core at the University of Houston to carry out the research presented here. We also acknowledge useful discussions with Dr. Jeremy Palmer.

## References

- [1] J. G. Kim, B. Son, S. Mukherjee, N. Schuppert, A. Bates, O. Kwon, M. J. Choi, H. Y. Chung, S. Park, *J. Power Sources* **2015**, 282, 299.
- [2] Y. Seino, T. Ota, K. Takada, A. Hayashi, M. Tatsumisago, *Energy Environ. Sci.* **2014**, 7, 627.
- [3] W. S. Tang, M. Matsuo, H. Wu, V. Stavila, W. Zhou, A. A. Talin, A. V. Soloninin, R. V. Skoryunov, O. A. Babanova, A. V. Skripov, A. Unemoto, S. I. Orimo, T. J. Udovic, *Adv. Energy Mater.* **2016**, 6, 1.
- [4] C. Sun, J. Liu, Y. Gong, D. P. Wilkinson, J. Zhang, *Nano Energy* **2017**, 33, 363.
- [5] J. F. M. Oudenhoven, L. Baggetto, P. H. L. Notten, *Adv. Energy Mater.* **2011**, 1, 10.
- [6] W. G. Zeier, J. Janek, *Nat. Energy* **2016**, 1, 16141.
- [7] N. Kamaya, K. Homma, Y. Yamakawa, M. Hirayama, R. Kanno, M. Yonemura, T. Kamiyama, Y. Kato, S. Hama, K. Kawamoto, A. Mitsui, *Nat. Mater.* **2011**, 10, 682.
- [8] Y. Mo, S. P. Ong, G. Ceder, *Chem. Mater.* **2012**, 24, 15.
- [9] D. Di Stefano, A. Miglio, K. Robeyns, Y. Filinchuk, M. Lechartier, A. Senyshyn, H. Ishida, S. Spannenberger, D. Prutsch, S. Lunghammer, D. Rettenwander, M. Wilkening, B. Roling, Y. Kato, G. Hautier, *Chem* **2019**, 5, 2450.
- [10] Z. Zhu, I. H. Chu, S. P. Ong, *Chem. Mater.* **2017**, 29, 2474.

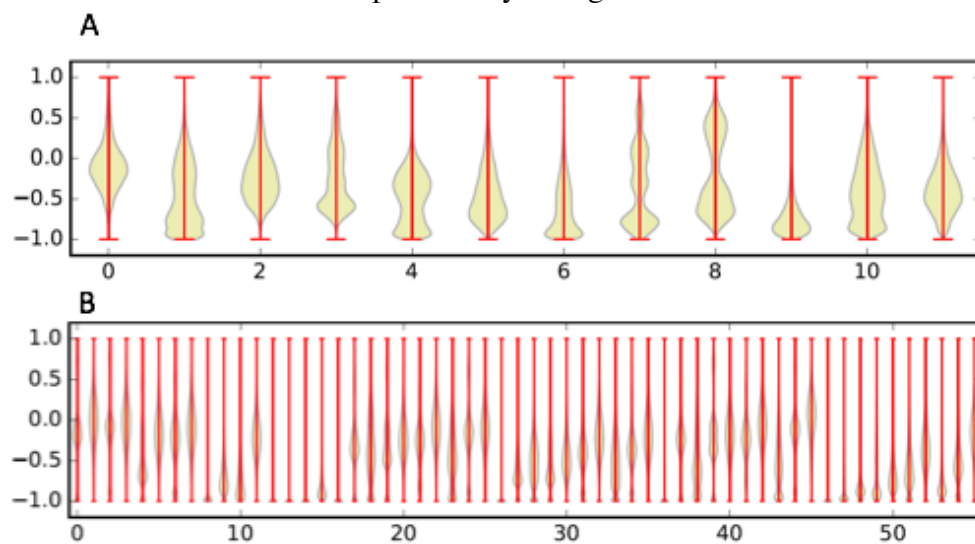
- [11] Y. Wang, W. D. Richards, S. P. Ong, L. J. Miara, J. C. Kim, Y. Mo, G. Ceder, *Nat. Mater.* **2015**, *14*, 1026.
- [12] A.-Y. Song, Y. Xiao, K. Turcheniuk, P. Upadhy, A. Ramanujapuram, J. Benson, A. Magasinski, M. Olguin, L. Meda, O. Borodin, G. Yushin, *Adv. Energy Mater.* **2017**, *1700971*, 1700971.
- [13] J. B. Varley, K. Kweon, P. Mehta, P. Shea, T. W. Heo, T. J. Udovic, V. Stavila, B. C. Wood, *ACS Energy Lett.* **2017**, *2*, 250.
- [14] Z. Jiang, K. Klyukin, V. Alexandrov, *J. Chem. Phys.* **2016**, *145*, 114303.
- [15] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. Van Den Driessche, T. Graepel, D. Hassabis, *Nature* **2017**, *550*, 354.
- [16] Z. Wang, Y. Adachi, Z. Chen, *Adv. Theory Simulations* **2020**, *3*, 1900197.
- [17] X. Xu, N. Gupta, *Adv. Theory Simulations* **2019**, *2*, 1800131.
- [18] A. Mansouri Tehrani, A. O. Oliynyk, M. Parry, Z. Rizvi, S. Couper, F. Lin, L. Miyagi, T. D. Sparks, J. Brgoch, *J. Am. Chem. Soc.* **2018**, *140*, 9844.
- [19] J. Behler, M. Parrinello, *Phys. Rev. Lett.* **2007**, *98*, 1.
- [20] N. Artrith, A. Urban, *Comput. Mater. Sci.* **2016**, *114*, 135.
- [21] J. Behler, in *Int. J. Quantum Chem.*, **2015**, pp. 1032–1050.
- [22] C. M. Handley, P. L. A. Popelier, *J. Phys. Chem. A* **2010**, *114*, 3371.
- [23] A. A. Peterson, R. Christensen, A. Khorshidi, *Phys. Chem. Chem. Phys.* **2017**, *19*, 10978.
- [24] S. Hajinazar, J. Shao, A. N. Kolmogorov, *Phys. Rev. B* **2017**, *95*, 014114.
- [25] J. R. Boes, J. R. Kitchin, *J. Phys. Chem. C* **2017**, *121*, 3479.
- [26] J. R. Boes, J. R. Kitchin, *Mol. Simul.* **2017**, *43*, 346.

- [27] C. Chen, W. Ye, Y. Zuo, C. Zheng, S. P. Ong, *Chem. Mater.* **2019**, *31*, 3564.
- [28] T. D. Huan, R. Batra, J. Chapman, S. Krishnan, L. Chen, R. Ramprasad, *npj Comput. Mater.* **2017**, *3*, 37.
- [29] J. Wang, S. Shin, S. Lee, *Adv. Theory Simulations* **2020**, *3*, 1.
- [30] S. R. Bahn, K. W. Jacobsen, *Comput. Sci. Eng.* **2002**, *4*, 56.
- [31] G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169.
- [32] G. Kresse, J. Hafner, *Phys. Rev. B* **1994**, *49*, 14251.
- [33] G. Kresse, J. Hafner, *Phys. Rev. B* **1993**, *47*, 558.
- [34] E. D. Stevens, J. Rys, P. Coppens, *J. Am. Chem. Soc.* **1978**, *100*, 2324.
- [35] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Mater.* **2013**, *1*, 011002.
- [36] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, *J. Phys. Condens. Matter* **2017**, *29*, 273002.
- [37] S. Melchionna, G. Ciccotti, B. Lee Holian, *Mol. Phys.* **1993**, *78*, 533.
- [38] S. Melchionna, *Phys. Rev. E. Stat. Phys. Plasmas. Fluids. Relat. Interdiscip. Topics* **2000**, *61*, 6165.
- [39] B. L. Holian, A. J. De Groot, W. G. Hoover, C. G. Hoover, *Phys. Rev. A* **1990**, *41*, 4552.
- [40] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, G. Ceder, *Comput. Mater. Sci.* **2013**, *68*, 314.

- [41] Humphrey W., Dalke A., Schulten K., *J. Mol. Graph.* **1996**.
- [42] A. Khorshidi, A. A. Peterson, *Comput. Phys. Commun.* **2016**, 207, 310.
- [43] J. Nocedal, S. J. Wright, *Numerical Optimization Second Edition*, **2006**.
- [44] K. K. Rao, Q. K. Do, K. Pham, D. Maiti, L. C. Grabow, *Top. Catal.* **2020**, *Accepted*, DOI 10.1007/s11244-020-01267-2.
- [45] Tianping Chen, Hong Chen, *IEEE Trans. Neural Networks* **1995**, 6, 911.
- [46] D. Sholl, J. Steckel, Sholl, *Density Functional Theory: A Practical Introduction*, Wiley, **2009**.
- [47] C. A. J. Fisher, H. Matsubara, *J. Eur. Ceram. Soc.* **1999**, 19, 703.
- [48] Y. Kato, S. Hori, T. Saito, K. Suzuki, M. Hirayama, A. Mitsui, M. Yonemura, H. Iba, R. Kanno, *Nat. Energy* **2016**, 1, 16030.
- [49] A. Dive, Y. Zhang, Y. Yao, S. W. Martin, S. Banerjee, *Solid State Ionics* **2019**, 338, 177.
- [50] O. L. Anderson, D. A. Stuart, *J. Am. Ceram. Soc.* **1954**, 37, 573.
- [51] R. Christensen, G. Olson, S. W. Martin, *J. Phys. Chem. B* **2013**, 117, 16577.
- [52] B. G. Levine, J. E. Stone, A. Kohlmeyer, *J. Comput. Phys.* **2011**, 230, 3556.
- [53] S. E. Boulfelfel, G. Seifert, S. Leoni, *J. Mater. Chem.* **2011**, 21, 16365.
- [54] B. Zhang, R. Tan, L. Yang, J. Zheng, K. Zhang, S. Mo, Z. Lin, F. Pan, *Energy Storage Mater.* **2018**, 10, 139.
- [55] S. Wang, Q. Bai, A. M. Nolan, Y. Liu, S. Gong, Q. Sun, Y. Mo, *Angew. Chemie Int. Ed.* **2019**, 1.
- [56] J. B. Haskins, W. R. Bennett, J. J. Wu, D. M. Hernández, O. Borodin, J. D. Monk, C. W. Bauschlicher, J. W. Lawson, *J. Phys. Chem. B* **2014**, 118, 11295.
- [57] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S.

## Supporting Information

**Figure S1.** (A) and (B) show the fingerprint outputs for the Zernike, and the Gaussian descriptors corresponding to Table 1, respectively. The input vectors have length 12 and 56, respectively. There are no nodes with a constant output nor any acting as switches.



**Figure S2.** Li-Li radial distribution function from ANN-MD from 2 to 6 Å. The first peak for all Cl concentrations occurs at 3.25 Å.

