

Utilization of Artificial Neural Network to explore the compositional space of hollandite-structured materials for radionuclide Cs incorporation

Dipta B. Ghosh^{1,2,*}, Bijaya B. Karki^{1,2,3} and Jianwei Wang^{2,3,*}

¹School of Electrical Engineering and Computer Science,

²Department of Geology and Geophysics, ³Center for Computation and Technology

Louisiana State University, Baton Rouge, LA 70803, USA

*Corresponding author

Abstract:

Hollandite with the general formula $A_2B_8O_{16}$ is known for its potential to immobilize radionuclide Cs in the tunnel along the z -axis of the crystal structure. The effective Cs incorporation in a hollandite phase with an optimal loading capacity and the long term stability depends significantly on the B-site cations, which, in addition to providing optimal structural compatibility, must ensure the phase's resistance to chemical weathering in an aqueous environment that includes external thermodynamic conditions such as temperature and solution chemistry. Based on the importance of the B-site cations, we explored in detail the possible B-site compositions by employing Artificial Neural Network (ANN) simulations and crystal chemistry principles. With a set of 91 experimentally determined data collected on hollandite that is available in open literature, we trained the network and subsequently tested the predictive power of the trained network. Relying on the successful outcomes of the trained network at the testing phase, we further utilized the trained network to map the dependence of the tunnel size, which was used as a criterion for Cs compatibility in the channel, in a wide compositional space encompassing eighteen 3+ cations and fifteen 4+ cations. By combining the Cs compatibility and the structural tolerance factor for hollandite structure, the predicted B-site compositions, comprising of cations spanning across the depth and breadth of the periodic table, can be employed as a guide in the search for optimal hollandite composition for Cs immobilization.

Introduction:

Safe storage and permanent disposal of nuclear wastes require appropriate isolation and immobilization of the nuclear wastes in hosts such as ceramic materials and the glass matrix. High degree of thermal stability and chemical stability under varying aqueous environments are essential in order to effectively immobilize the nuclear-waste form. This is because, exposure to aqueous solutions with different pH conditions is likely unavoidable under natural geological repositories and a significant amount of heat is generated in the radioactive decay process. Borosilicate glasses, which have been used to dispose nuclear wastes have some disadvantages because some of the radionuclides such as ^{137}Cs and ^{129}I become volatile at borosilicate glass melting temperatures and also due to their low solubility for these radionuclides [1].

Ceramic waste-forms such as synroc have instead been suggested as a potential alternative to circumvent the aforesaid drawback and withstand higher waste loading [2–6]. In particular, to immobilize Cs^+ and its transmuted radionuclide, Ba^{2+} , among several proposed ceramic waste forms, e.g., sodium zirconium phosphate (NZN-CsZr₂(PO₄)₃) [7], pollucite ((Cs,Ca)AlSi₂O₆) [8] and hollandites $[\text{Ba}_x\text{Cs}_y][(\text{Ti}(\text{or Fe}), \text{Al})_{(2x+y)}^{3+}\text{Ti}_{(8-2x-y)}^{4+}]\text{O}_{16}$ [9, 10], hollandites have emerged as a viable option for Cs immobilization.

With the general formula $\text{A}_2\text{B}_8\text{O}_{16}$, hollandites can potentially incorporate a wide-range of cations in the A and B sites of the structure (Fig. 1). For instance, the A-sites can have either mono or divalent cations such as: Na^+ , Cs^+ , Rb^+ , Ba^{2+} , etc., and the B-sites can accommodate di-, tri- and tetra-valent cations such as: Mg^{2+} , Fe^{2+} , Fe^{3+} , Al^{3+} , Ti^{3+} , Ti^{4+} , Si^{4+} , among others. Owing to the tunnel (A-site) surrounded by octahedral cation-anion units, hollandite structure is adept for immobilizing cesium (Cs). The ionic radius of Cs is however relatively large (1.74 Å for eight coordinated Cs^+). In order to stabilize the element inside the tunnel site and for relatively high Cs loading, the tunnel size should be such that there is no size mismatch between Cs and the channel, or no residual internal stress at the local site and in the overall lattice, which could occur as a result of the mismatch. That is where the role of the B site cation(s) is utterly significant, i.e., by appropriate substitutions with a mixture of suitable cation(s), the tunnel/channel size is optimized and the structure is stabilized, so that the phase is thermodynamically stable and preferably it is also less prone to the environmental degradation.

Substantial experimental efforts have been given with regard to the compositional optimization of the B-site and stability [9, 11–22]. For instance, studies on Barium hollandites have indicated that substitution

of Ti^{4+} with Mg^{2+} , Al^{3+} or Fe^{3+} increases the thermodynamic stability of the hollandite phase, with the Fe^{3+} substituted phase being most stable [14]. The results of this study [14] have also indicated that the single phase Al- and Fe- (but not Mg-) hollandites are relatively stable with respect to the phase assemblages containing $BaTiO_3$ perovskite. In case of Ti-based hollandites with Cs and Ba in the A site, it was reported that Cs incorporation in the hollandite phase was nearly perfect (~100% Cs retention) for Fe^{3+} -hollandite, in contrast, Cs retention was only 46% in case of Cr^{3+} -hollandite, with rest of the Cs being present in some other Cs-rich segregated phase [13]. In another study on Ti^{4+} -based Cs-Ba hollandites, with Cr^{3+} , Fe^{3+} and Al^{3+} cations in the B site, the Cr-bearing hollandites have been reported to be better (in terms of relative durability) candidates than the Al-, Fe- or Cr-Al-Fe- bearing hollandites for Cs incorporation and retention [15]. The difference in synthesis temperature, i.e., solid state sintering [13] vs. melt processing [15], seems to be the apparent source of contradiction between these two experimental data. It is worth noting here that despite using melt processing at 1500 °C, Cr-hollandites were not melted at all, while the Fe-hollandites were completely melted and Cr-Al-Fe hollandites were only partially melted in the study of Amoroso et al., [15]. Instead, all the samples were processed by solid state sintering at 1200 °C in the study of Aubin-Chevaldonnet et al., [13]. In these cases, it is intriguing to see how the composition and processing temperature and methods can lead to the differences in Cs incorporation and retention [13, 15, 23].

The suitability of hollandite to immobilize Cs has also been recently assessed in terms of thermodynamic stability [21] and its tolerance with regard to radiation [22]. The study on $Ba_xCs_yZn_{(x+y/2)}Ti_{(8-x-y/2)}O_{16}$ ($0 < x < 1.33$; $0 < y < 1.33$) indicated that Zn-doped hollandites are not only thermodynamically stable with respect to their parent oxides, the thermodynamic stability also increases with increasing Cs content [21]. On the issue of radiation damage, elevating the temperature seems to increase the tolerance. These studies have indicated that both, the loss due to thermal evaporation and leaching, decreased with increasing Cs content [21, 22]. Recent increased research effort to understand these potential nuclear waste form host materials is therefore clearly evident.

It is apparent however, that neither any clear consensus has been reached on the issue, nor is our scientific understanding of the mechanism for Cs incorporation complete. However, it is clear that both the individual and coupling effects from multiple cations in the B site on the Cs incorporation are important in designing the composition of Cs-hollandite waste form. The effect of the concentration of the elements at the A and B sites for Cs incorporation is expected to be non-linear. There is no doubt that our knowledge on this ceramic phase design in particular and in general with regard to the waste form

development have greatly expanded, mainly due to the recent advancement by the experimental community. But, our current knowledge about how Cs incorporation in hollandite is affected by different cations is still limited, because, the available literature encompasses only a fraction of possibilities in the B-site composition space. With such immensity of possibilities of different B site cations, performing experimentation over the vast composition space is by all means is not realistic. Computer assisted experimentation may however be able to address some of the aforesaid challenges. An effort to that end has very recently been undertaken by Utlak et al. [24], who have developed a thermodynamic database on the $\text{BaO-Cs}_2\text{O-TiO}_2\text{-Cr}_2\text{O}_3\text{-Al}_2\text{O}_3\text{-Fe}_2\text{O}_3\text{-FeO-Ga}_2\text{O}_3$ system by using computer coupling of phase diagrams and thermochemistry (CAPHAD) methodology and available experimental data. Their assessment of the thermodynamic stability of the predicted compositions formed by combinations of five B-site cations provide valuable input to the experimental community. The possible B-site composition space is, however, far wider than that have been assessed. A concerted effort using computational and experimental methods is expected to accelerate the advancement in the field. To this end, we carried out Artificial Neural Network (ANN) simulations on the hollandite system to explore the B-site compositional space of Cs-bearing hollandite. As discussed earlier, the optimization of the tunnel/channel size has been one of the key challenges to the use of hollandite as the Cs nuclear waste-form. We focus on this particular aspect, i.e., assess channel size over the B-site composition space while restricting the A-site compositions only to Cs and Ba to narrow the working scope. The B-site composition space in our study encompasses eighteen 3+ cations and fifteen 4+ cations. After ascertaining the reliability of the predictive power of the trained network, we provide detailed profiles of the tunnel size in the aforesaid composition space.

Methods:

Artificial Neural Network (ANN) methodology, as the name suggests, is inspired and framed on the basis of complex biological neural networks, but with a simplistic viewpoint [25]. In a typical ANN approach, the network consists of: (i) an input layer, (ii) one or more hidden layers, and (iii) an output layer. During the learning/training process, the network establishes the relation between the input layer and the output layer via the intermediate hidden layer(s). The hidden layer(s) continually adjusts the weight and bias of the input layer until the desired outcome is reached or exceeds the iteration count. Given the desirable outcome is achieved in the learning process, the trained network is put through a testing cycle with a dataset that is completely unknown to the trained network. A well trained and tested network is considered to be reliable and expected to have a predictive power with high degree of accuracy. ANN, as implemented in the Matlab program package, was used for the simulation and analysis in this study

[26, 27]. Three (Levenberg-Marquardt, scaled conjugate gradient, and Bayesian regularization) backpropagation training algorithms were tested and Bayesian regularization algorithm provides the best performance. We refrain from any more details on ANN, but interested readers are referred to Wang [28] for a recent study on its application for incorporation of iodine in apatite-structured materials and the review by Mueller et al. [29] for details on the applications of machine learning in materials science.

A dataset of 91 hollandite compositions sourced from the inorganic crystal structure database (ICSD) [30], and crystallography open database (COD) [31] was used for the study. The dataset was divided into ~70% and ~30% for the training and testing purposes, respectively. The dataset consists of tetragonal (space group I4/m) phase hollandite, with octahedral B-sites and 8-coordinated A-sites. In order to reliably predict the characteristics (such as lattice parameters, channel size, etc.) in unknown compositions, the more diverse the compositions (accounting the elements in the periodic table) of the training dataset, the greater accuracy one would or can expect from the predictions. From that perspective, the mined dataset is very reasonable. For example, for A-site elements with charge of +1 and/or +2, it includes 12 elements covering alkaline, alkaline earth, some transition metal elements among others (Ba, Cs, K, Rb, Na, Tl, Ag, Pb, Bi, Sr, Pr, Li). For B-site, the dataset includes 20 elements including light alkaline earth, transition metal, and group 13-15 elements (Mo, Rh, Al, Si, Fe, Mn, V, Ru, Sb, Ti, Cr, Ga, Zr, Co, Mg, Ni, Ir, Sn, Nb, Sc) with a charge of 2+, 3+ and/or 4+. Considering that several of the B-site elements are multivalent, the extent of composition space is even larger.

There are two input parameters that were fed to the neural network – ionic radius [32, 33] and the electronegativity. While the ionic radius depends on both the charge and coordination state, the electronegativity of any element is constant. However, for elements with similar ionic radius and charge, electronegativity plays an important role on their difference in chemical bonding. Therefore, considering electronegativity as an input parameter may provide better constraints to the predicted properties. In fact, the inclusion of electronegativity did provide better constraints on the predicted properties in the ANN study on apatite system [28]. Thus, we stick to this consideration in the present study. For consistency of the ionic radius, 8-coordinated A-site has been considered throughout. Both of the input parameters were determined by the weighted (molar fraction) average of the individual atoms in any specific site. The output of the ANN simulations in the present study are lattice parameters and the channel size.

In a nutshell, during the supervised training phase, the ANN established the input (ionic radius and electronegativity) – output (lattice parameter, channel size) correlation. The efficacy of the trained

network was subsequently put to the test (testing phase). The trained network was subsequently utilized for predicting channel sizes in unknown hollandite compositions encompassing a wide B-site composition space. A trained network is considered reliable for the prediction purposes only when the output on both the training and the testing phases fulfills the desired outcome.

Results:

Training the ANN and Testing of the trained ANN:

The network was trained with a set of 64 data. Correspondingly, the trained network was tested with 27 data. The dataset was strategically divided so that each of the training and the testing dataset encompasses the widest variation of compositions and channel sizes possible with the total of 91 data. This strategy is appropriate for a small set of data for the present problem. The neural network considered has a total of four (Fig. 2) input parameters, two (ionic radius and electronegativity) coming from each of the A-site and B-site. Since, the third crystallographic site is occupied by only oxygen atoms and is always fully occupied, the ionic radius and electronegativity for this site do not change with compositional variance and thus need not be included for the simulation. The hidden layer contains 4 neurons and there is one output layer.

Fig. 3 displays the correlation between the actual and predicted lattice parameter, a (in Å). The result indicates very high reliability of the predicted outcome from this trained ANN based on the correlation coefficient value of $R > 0.97$, together with the very low average error values (0.34%, 0.54% and 0.40% for the training, testing and entire dataset, respectively) between the predicted and actual data for either of the training and testing data. The maximum error is less than 3%. This high degree of correspondence between the actual and the predicted outcome is also evident for the lattice parameter, c (in Å) (Fig. 4). The R is > 0.98 , and the maximum error remains within 2.2% (the average errors for the training, testing and entire dataset are 0.36%, 0.33% and 0.35%, respectively). In addition, the residual distributions for all the fits of lattice parameters a and c are approximately normally distributed around zero error (the insets of Figs. 3 and 4), indicating a statistically reasonable prediction of the neural network. These results provide further confidence to utilize the trained network for prediction purposes.

To put some perspective in terms of reliability of the results of the ANN in the current study, in the early days of ANN based predictions, the fitting coefficient (R) value of greater than 0.60 was considered reliable. Although we have come a relatively long way from those initial days of ANN, the value of R depends significantly, among others, on the number of available data, the spread of that data and subtle

partition of the data between the training and testing data set. This is especially true where the available data is relatively small. Because both our training and testing data have been strategically chosen to encompass the entire regime of the lattice parameters (a or c), the network is likely to be fairly reliable for predicting the aforesaid parameters.

As the channel size is linked to the incorporation of Cs at the A-site, the channel size was assessed by employing the ANN. The assumption is that, due to the large size of Cs⁺ ion, the channel of hollandite structure needs to be large enough, but not too large, to host the ion. Because oxygen atoms are the 1st nearest neighbor of the channel cation in the hollandite structure, the average Cs–O bond distance of the 8-coordinated Cs is a good parameter to assess the channel size. The results of the ANN simulations for the channel size are shown in Fig. 5. The value of R remains greater than 0.93 in all cases. The errors are also small and approximately normally distributed, with the average error (for the training, testing and entire dataset) varying from 0.91 – 0.94%, and the maximum error remaining < 4.1%. It is thus very likely that predictions of the channel sizes for unknown compositions from this trained network will be fairly reliable.

Predictions of the channel size from the trained ANN:

Having assessed the reliability of the trained network for the channel size, we performed further ANN modeling, scanning the entire B-site composition space for a given A-site composition. Out of the 20 B-site candidate elements considered here, we selectively constructed a 15 x 18 matrix with a combination of M³⁺ and M⁴⁺ cations in order to compensate the charge for the A-site occupancy. For simplicity, we restricted the B-site composition to consist of only two (one 3+ cation and one 4+ cation) types of cation. Since the main emphasis is Cs at the A-site, we only considered Cs and Ba for the A-site. For A-site occupancy, we have varied Ba from 0.01 to 1.33. For each Ba content, we have varied the Cs content from 0.0 to 1.00. It is noted here that the maximum A-site occupancy is 2.0. The variation of the occupancy is only for the sake of the numerical modeling. The feasibility of the resulting hollandite compositions will be evaluated based on the channel size and the structural tolerance factor, as will be seen later in the text.

The left panel of Fig. 6 displays the dependence of the channel size as a function of the average A-site and B-site cation radius. The channel size profile does not show any apparent trend as a function of A-site radius, but has a strong dependence on the B-site radius where it displays a decreasing channel size from 3.3 Å to 2.5 Å with decreasing B-site radius from 0.9 Å to 0.4 Å. We note here that the channel

size range below 2.8 Å may be too small to incorporate Cs, especially for medium to high load conditions. In any case, as we want to understand how channel size depends on the A- and B-site radius, we filtered the data for selected Cs content (Fig. 6, right panel). For a given B-site radius, the channel size variation is not negligible, but it is clear that the effect of B-site radius is much more significant than that of A-site radius and Cs content. This suggests strong influence of the B-site composition on the channel size. This in fact is in agreement with experimental data where the dominance of B-site cations in controlling the overall cell volume and the channel size has been reported [34].

Assessment of predicted hollandite compositions:

While the effect of B-site composition on the channel size is apparent, the ANN simulations only provide predictions based purely on the crystal chemistry consideration and the trained network from known hollandites. In order to ascertain whether the predicted compositions are realistically plausible, it is critically important to further assess the data including both those from literature and those from the prediction using any available models. The most used and generally accepted criterion to assess the stability of a crystal structure in this regard is called the tolerance factor, t , and for hollandite it is given by [35]

$$t_H = \frac{[(r_A+r_O)^2 - \frac{1}{2}(r_B+r_O)^2]^{\frac{1}{2}}}{\sqrt{\frac{3}{2}}(r_B+r_O)} \quad (1)$$

where, average radius of the A- and B-site cations are r_A and r_B , respectively and r_O being the oxygen ionic radius. Proximity of the t_H value towards unity (~ 1) is generally considered to be indicative of increasing structural stability of the hollandite composition and experimental studies have demonstrated using this to assess the structural stability [19, 20].

In order to find if any correlation between the channel size and the stability of the composition exists, we have explored the tolerance factor, t_H , for the ANN predicted data. It is clear from Fig. 7 (left) that numerous data points on the predicted compositional space deviate significantly from the $t_H \sim 1$, suggesting not all predicted compositions may be feasible. It is interesting however to see the correlation between t_H and the channel size. An overall non-linear increase in the channel size with decreasing t_H is apparent along with a non-monotonous behavior around $t_H \sim 1$. The effects of Ba and Cs content on the value of t_H for any given channel size are however clearly visible (Fig. 7, left). It is noted here that a desired tunnel site of fixed channel radius can in principle be achieved by suitably choosing the B-site compositions even in case of varying A-site compositions. Since the ionic radius of Ba (1.42 Å) is smaller

than that of the Cs (1.74 Å), for a given Cs (Ba) content, the A-site average radius should decrease (increase) with increasing Ba (Cs) content. This effect is enhanced in the tolerance factor because the numerator in Eq. (1) contains a quadratic term involving the A-site radius. The total (Ba and Cs combined) occupation of A-site and Cs occupation seems to affect both the channel size and t_H (Fig. 7, left), which is not negligible with respect to the channel size dependence on the B-site composition, a characteristic we have also seen in Fig. 6.

Using $t_H \sim 1$ as an indicator of the structural stability, we consider the data lying within the range $0.90 \leq t_H \leq 1.10$ for the purpose of this study (symmetric deviation from the unity though chosen arbitrarily is expected to be reasonable). We note here that Kesson and White [35] attributed a t_H value range of 0.93 – 1.16 for the stability of hollandites and values reported in recent experimental studies remain well within that range [14, 19, 20]. Predicted channel sizes lying within this t_H range can be considered as potential compositions with hollandite structure for Cs immobilization. In addition, from the ionic radius consideration and given that the first near-neighbor is oxygen, a straightforward estimate suggests that the channel size can vary from 2.8 Å (for Cs free with full A-site occupancy, Ba radius 1.42Å, and O radius 1.38Å) to 3.12 Å (Ba free with full occupancy of A-site, Cs radius 1.74 Å and O radius 1.38 Å) in hollandite. Therefore, hollandite compositions with a channel size > 3.12 Å are less likely to be relevant, but just for visual purpose we plot the data up to 3.15 Å.

B-site compositional control:

In any case, at this point it is apparent that the channel size is primarily dictated by the B-site radius. Therefore, mapping of B-site composition space can provide essential structural information and serve as a guide to the future exploration of suitable hollandite compositions that incorporate Cs. In Fig. 7 (right panel), we provide such mapping with both the constraints discussed above, applied to the dataset, as indicated by the rectangle in the left panel in Fig. 7. The heatmap shows the dependence of channel size on the B-site 3+ and 4+ cations. While there is no standard technique to constrain the channel size, in the high Cs loading limit, based on ionic radius consideration, a channel size value of 3.0 (± 0.10) Å is likely a fair assumption. The composition space shown here (Fig. 7, right) qualitatively depicts that, apart from the widely studied Ti-based (Ti^{4+} as major B-site cation) hollandites, there can be several other 4+ cations which can also result in the channel size range and tolerance factor, t_H , that is apt for Cs incorporation. For instance, Zr^{4+} or Sn^{4+} based hollandites at B-site mixing with Fe^{3+} , Cr^{3+} or Ti^{3+} cations

are possible considerations (Fig. 7, right) with their channel size varying from 3.07 – 3.12 Å for $\text{Ba}_x\text{Cs}_y\text{M}^{3+}\text{Zr}^{4+}$ ($\text{M} = \text{Fe}, \text{Cr}$ and $x+y = 1.0 - 2.0$) compositions and that for $\text{Ba}_x\text{Cs}_y\text{Ti}^{3+}\text{Sn}^{4+}$ varying from 3.02 – 3.07 Å (marked by black squares in Fig. 7, right).

The radioparagenesis [36] of Cs to Ba induces chemical changes and may cause instability in case of insufficient charge compensation. So, for any potential composition, the ability to compensate the charge imbalance that occurs by radioparagenesis process should also be assessed. From that perspective, although the M^{4+} cations mentioned above cannot be readily reduced to M^{3+} for charge compensation, all of the M^{3+} cations discussed above are in principle capable of maintaining the overall charge balance and redox state of the system by consuming the electron produced in the beta decay process (i.e., $^{137}\text{Cs}^{1+} \rightarrow ^{137}\text{Ba}^{2+} + \beta(e^-)$) and $\text{M}^{3+} + \beta(e^-) = \text{M}^{2+}$). Also, the relatively greater difference in the ionic radius between the M^{3+} and M^{2+} (than the difference between M^{4+} and M^{3+}) makes more sense in terms of compensating the decrease in tunnel size due to $\text{Cs} \rightarrow \text{Ba}$ conversion. The reason is that as only a fraction of total (eight) B-site cations participates in this charge compensation process as compared to one-to-one for $\text{Cs} \rightarrow \text{Ba}$ (ionic radius 1.74 vs. 1.42 Å, respectively), the effective increase of the average ionic radius is suppressed in case of B-site cations. Further, with regard to long term durability, Zr in particular is also well known for high resistance to heat and corrosion. It may therefore be worthy of exploring Zr-based hollandite phases.

Because of their relatively high stability, Ti-based hollandite compositions are the most studied ones with regard to radionuclide Cs incorporation. Given the importance of Ti-based hollandites, we further explored the composition space for Ti- based hollandites (Fig. 8 & S1). For comparison purposes, we plot Ti^{3+} and Ti^{4+} with the other B-site cation being M^{4+} and M^{3+} , respectively, for the t_H range 0.90 – 1.10 and channel size range 2.80 – 3.15 Å. For any hollandite composition displayed in Fig. 8 (and Fig. S1), Ba contents of 0.34 and 1.00 are shown, and for any given Ba content, Cs content varies from 0.0 – 1.00 with an interval of 0.2.

As expected, for any given Ba content, increasing Cs content increases the average A-site radius and the channel size (except for Zr^{4+} -based hollandites, which display an anomalous decreasing trend in channel size with increasing average A-site radius and remains inconclusive to us). This increased channel size is likely to facilitate the stability of the Ti-based hollandites as has been demonstrated for the $[\text{Cs}_x\text{Ba}_y][(\text{Al}^{3+}, \text{Ti}^{3+})_{(2y+x)}\text{Ti}^{4+}_{(8-2y-x)}]\text{O}_{16}$ composition [19, 20]. An $\text{M}^{4+}\text{Ti}^{3+}$ combination displays a larger variation in the channel size in comparison to the $\text{M}^{3+}\text{Ti}^{4+}$ combination. For example, several $\text{M}^{4+}\text{Ti}^{3+}$

(M=Mn, Rh, Ir, Mo) pairs display a wide range ($\sim 2.85 \text{ \AA} - \sim 3.05 \text{ \AA}$). In contrast, for $M^{3+}Ti^{4+}$ combination, most of the channel sizes remain in between 2.925 to 3.05 \AA , except for Pr, Ce, La, Bi compositions. What is interesting is that, despite this large variation in the channel size with B-site compositional difference, the variation in the average B-site radius is relatively small (Fig. 8, left). On the other hand, as expected, the effect of Cs content is quite clear from Fig. 8 (right).

Based on this composition-space – channel-size profile, it can be said that several unexplored Ti-based compositions are likely to be competitive in terms of their potential to Cs incorporation. There also exists a number of compositions that display channel sizes $> 3.10 \text{ \AA}$. For Ti^{3+} (Ti^{4+}) based compositions, the general trend is that the upper bound of the channel size is pushed to higher values with the increase in ionic radius of the M^{4+} (M^{3+}) cations. While this study cannot address the actual stability of these possible candidate compositions from any thermodynamic point of view, or for that matter, shed light on the issue relating to A- and B-site ordering and channel size, given that the neural network was trained successfully and the dataset in Fig. 8 (and Fig. S1) has been filtered with t_H values remaining around 1.00 (± 0.10), the wide spectrum of compositions provided here may be worthy of further exploration for Ti-based hollandite for Cs immobilization.

Concluding remarks:

ANN simulations were employed to explore the compositional space of the B-site cations in hollandite phase. Predictions based on these simulations suggest that a wide variety of compositions that meet the tolerance factor (t_H) criterion and also display channel size in the range of 2.80 – 3.15 \AA are possible. Being the most studied hollandite compositions, some Ti-based hollandites show relatively low leaching rates under different environmental conditions. However, how their channel size influence their chemical durability is not fully understood, especially for higher Cs loading compositions. In addition to Ti-based hollandites, it would be interesting to explore the feasibility of other M^{4+} cation-based hollandite compositions. The predicted dataset thus may serve as a starting point for studying unexplored Ti-based and non-Ti-based hollandite compositions for Cs incorporation.

While the present study provides valuable data with regard to the B-site composition and the channel size, it still remains within the realm of data science. Whether the proposed hollandite compositions are feasible in terms of synthesis, thermal stability and long term durability with regard to decay of ^{137}Cs or not remains unknown at present. Utilization and application of the data in materials research are, therefore, needed to gain insights on the materials characteristics. It will be of interest to assess the

thermodynamic stability of the predicted compositions by experimental synthesis and/or computational thermodynamics calculations.

Acknowledgments

This work was supported as part of the Center for Performance and Design of Nuclear Waste Forms and Containers, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award DE-SC0016584. The computation used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References:

1. Weber W. J., Navrotsky A., Stefanovsky S., Vance E. R., and Vernaz E., *Materials Science of High-Level Nuclear Waste Immobilization*, MRS Bulletin, 34, 46 – 53 (2009).
2. Ringwood A. E., Kesson S. E., Ware N. G., Hibberson W., and Major A., *Immobilization of high level nuclear reactor wastes in SYNROC*, *Nature* 278, 219 – 223 (1979).
3. Ringwood A. E., Kesson S. E., Ware N. G., Hibberson W., and Major A., *The SYNROC process: A geochemical approach to nuclear waste immobilization*, *Geochem. Journal* 13, 141 – 165 (1979a).
4. Ringwood, A. E. *Disposal of high-level nuclear wastes: a geological perspective* *Mineral. Mag.* 49, 159 – 176 (1985).
5. Stefanovsky S. V., Yudinsev S.V., Giere R., and Lumpkin G. R., *Nuclear waste forms, Energy, Waste and the Environment: A Geochemical perspective* (Special Publication 236, Geological Society of London, London, 2004), p. 37.
6. Wang L., and Liang T., *Ceramics for high level radioactive waste solidification*. *J. Adv. Ceram.* 1, 194 – 203 (2012).
7. Roy R., Vance E. R., and Alamo J., [NZP], *A new radiophase for ceramic nuclear waste forms*. *Mater. Res. Bull* 17, 585 (1982).
8. Hess N. J., Espinosa F. J., Conradson S. D., Weber W. J., *Beta radiation effects in ¹³⁷Cs-substituted pollucite*. *J. Nucl. Mater.* 281, 22 (2000).
9. Kesson S.E., White T. J., $[Ba_xCs_y][(Ti, Al)_{(2x+y)}Ti_{(8-2x-y)}]O_{16}$ Synroc-type hollandites I. *Phase chemistry*. *Proc. R. Soc. London A* 405, 73 (1986).
10. Bart F., Leturcq G., Rabiller H., in: *Environmental Issues and Waste Management Technologies in the Ceramic and Nuclear Industries IX: Proceedings of the symposium held at the 105th Annual Meeting of The American Ceramic Society, April 27–30, Nashville, Tennessee*, *Ceram. Trans.* 155, 11 (2004).
11. Kesson S. E., “*The Immobilization of Cesium in SYNROC Hollandite*,” *Radioac. Waste Manag.* 4[3], 53–72 (1983).
12. Cheary R. W., *A structural analysis of Potassium, Rubidium and Caesium substitution in Barium hollandite*. *Acta Cryst.* B43, 28 – 34 (1987).
13. Aubin-Chevaldonnet V., Caurant D., Dannoux A., Gourier D., Charpentier T., Mazerolles L., Advocat T., *Preparation and characterization of (Ba, Cs)(M, Ti)₈O₁₆ (M = Al³⁺, Fe³⁺, Ga³⁺, Cr³⁺, Sc³⁺, Mg²⁺) hollandite ceramics developed for radioactive cesium immobilization*, *J. Nucl. Mater.* 366, 137–160 (2007).

14. Costa G. C. C, Xu H., and Navrotsky A., Thermochemistry of Barium Hollandites. *J. Am. Ceram. Soc.*, 96[5], 1554 – 1561 (2013).
15. Amoroso, J., Marra, J., Conradson, S. D., Tang, M., and Brinkman, K. S., Melt processed single phase hollandite waste forms for nuclear waste immobilization: $Ba_{1.0}Cs_{(0.3)}A_{2.3}Ti_{5.7}O_{16}$; A = Cr, Fe, Al. *J Alloy Compd* 584, 590 – 599 (2014).
16. Tumurugoti P., Sundaram S. K., Brinkman K. S., Amoroso J. W., Fox K. M., Melt-processed multiphase ceramic waste forms. *Ceramic Transactions Series 250*, 205 – 212 (2014).
17. Tumurugoti P., Sundaram S. K., Misture S. T., Cesium immobilization in (Ba,Cr)-hollandites: Effects on structure *J. Solid State Chem.* 258, 72 – 78 (2018).
18. Xu Y., Wen Y., Grote R., Amoroso J., Shuller-Nickles L., Brinkman K. S., A-site compositional effects in Ga-doped hollandite materials of the form $Ba_xCs_yGa_{2x+y}Ti_{8-2x-y}O_{16}$: implications for Cs immobilization in crystalline ceramic waste forms, *Sci. Rep* 6, 27412 (2016).
19. Yang Y., Wang X., Luo S., Yang X., Ma J., Stability studies of $[Cs_xBa_y][(Al^{3+}, Ti^{3+})_{(2x+y)}Ti_{(8-2y-x)}^{4+}]O_{16}$ ceramics for radioactive caesium immobilization. *Ceramics International* 45, 7865 – 7870 (2019).
20. Yang Y., Yang X., Wang X., Luo S., Ma J., Huang Y., Chemical evolution effects on phase and microstructure of $[Cs_xBa_y][Ti_{(2y+x)}^{3+}Ti_{(8-2y-x)}^{4+}]O_{16}$ ceramic waste forms for radioactive cesium immobilization. *Journal of Nuclear Materials.* 517, 57–62 (2019a).
21. Grote, R., Zhao M., S-Nickles L., Amoroso J., Gong W., Lilova K., Navrotsky A., Tang M., Brinkman K. S., Compositional control of tunnel features in hollandite-based ceramics: structure and stability of $(Ba,Cs)_{1.33}(Zn,Ti)_8O_{16}$. *Journal of Materials Science.* 54(2), 1112–1125 (2019).
22. Grote, R., Hong T., S-Nickles L., Amoroso J., Tang M., Brinkman K. S., Radiation tolerant ceramics for nuclear waste immobilization: Structure and stability of cesium containing hollandite of the form $(Ba,Cs)_{1.33}(Zn,Ti)_8O_{16}$ and $(Ba,Cs)_{1.33}(Ga,Ti)_8O_{16}$. *Journal of Nuclear Materials.* 518, 166–176 (2019a).
23. Dandeneau C. S., Hong T., Brinkman K. S., Vance E. R., Amoroso J. W., Comparison of structure, morphology, and leach characteristics of multi-phase ceramics produced via melt processing and hot isostatic pressing. *Journal of Nuclear Materials.* 502, 113–122 (2018).
24. Utlak, S. A., Besmann T. M., Brinkman K. S., Amoroso J. W., Thermodynamic assessment of the hollandite high level radioactive waste form. *Journal of American Ceramic Society* (2019) (DOI: 10.1111/jace.16438).
25. Hassoun M. H., *Fundamentals of Artificial Neural Networks.* (1995) MIT Press Cambridge, MA. USA.
26. The MathWorks, I. (2017). *MATLAB and Artificial Neural Network Toolbox Release R2017a.* Massachusetts: The MathWorks. Inc.

27. Demuth, H., and Beale, M. (2013). *Neural Network Toolbox for use with MATLAB, User's Guide*. Massachusetts: The MathWorks, Inc.
28. Wang, J., Incorporation of iodine into apatite structure: a crystal chemistry approach using Artificial Neural Network. *Frontiers in Earth Science* 3, 20 (2015).
29. Mueller T., Kusne A. G., and Ramprasad R., Machine learning in materials science: recent progress and emerging applications. *Rev. in Comput. Chem.* 29, 186 – 273 (2016).
30. ICSD (2010) *Inorganic Crystal Structure Database. Fachinformationszentrum Karlsruhe. FIZ Karlsruhe*: National Institute of Standards and Technology.
31. COD, Crystallography Open Database (<http://www.crystallography.net/cod/index.php>)
32. Shannon R. D., Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, *Acta Cryst. A* 32, 751–767 (1976).
33. Henderson W., *Main Group Chemistry*. (2000) Hoboken, NJ: Royal Society of Chemistry.
34. Cheary R. W., An Analysis of the Structural Characteristics of Hollandite Compounds. *Acta Crystallogr.* 42[3], 229–36 (1986).
35. Kesson S. E., and White T. J., Radius Ratio Tolerance Factors and the Stability of Hollandites. *J. Solid State Chem.*, 63[1], 122 – 5 (1986a).
36. Jiang C., Uberuaga B. P., Sickafus K. E., Nortier F. M., Kitten J. J., Marks N. A., Stanek C. R., Using “radioparagenesis” to design robust nuclear waste forms. *Energ Environ Sci.* 3, 130 – 135 (2010).

Fig 1: Representative hollandite ($A_2B_8O_{16}$) structure. Small red spheres correspond to oxygen atoms, B-site octahedral cages (pale yellow) are connected either by edge or corner sharing oxygen atoms and the large green spheres correspond to the A- site cations that includes Cs.

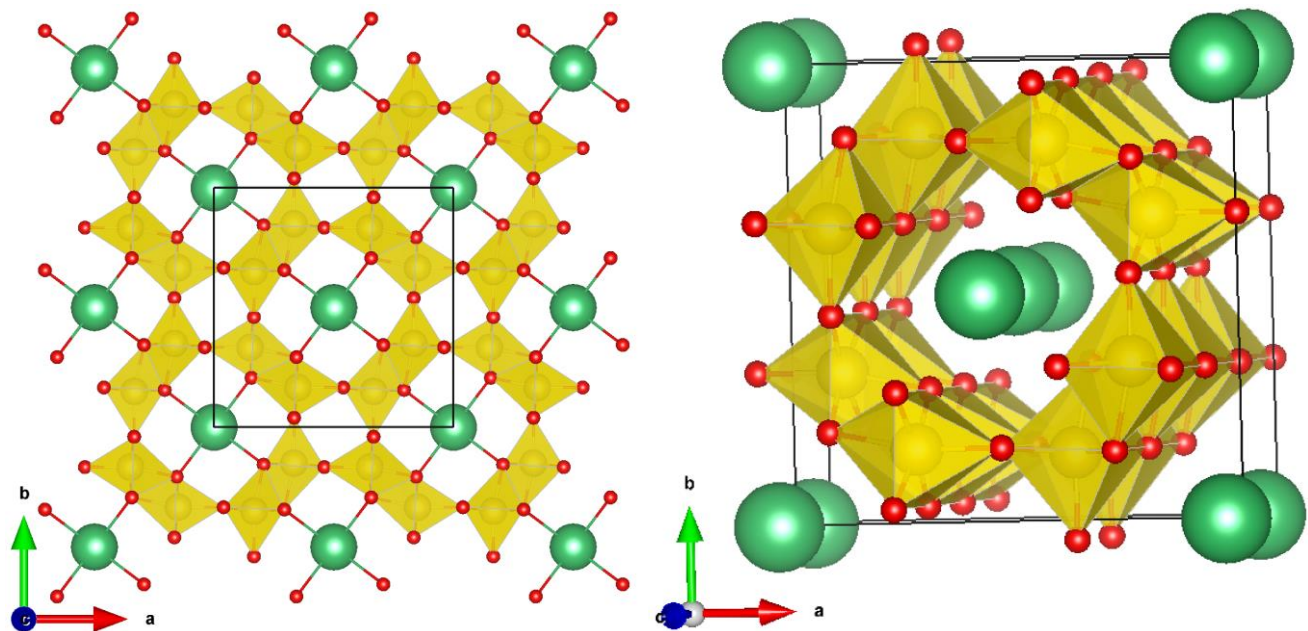


Fig 2: Schematic for the Artificial Neural Network used in the study. One input layer with four input parameters (ideally there would be 6 input parameters, two each for A, B and oxygen sites. But, since oxygen sites remain fully occupied with only oxygen atoms it need not be considered as input), one hidden layer with 4 neurons, one output layer and one output parameter.

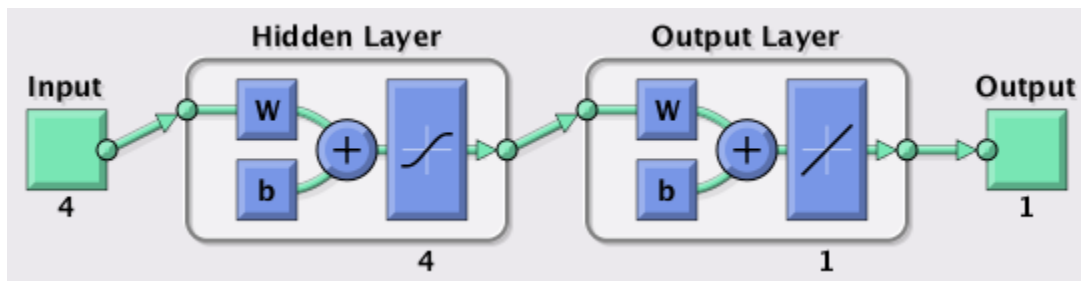


Fig 3: ANN simulation results for the unit cell parameter, a (in Å). Correlation between the actual and predicted values are shown. Circles represent actual data points, solid magenta lines are ANN fit to the data and the black dashed lines represent correlation coefficient value of $R=1$. The corresponding error distributions around *zero-error* (shown by red solid triangles) are shown in the inset.

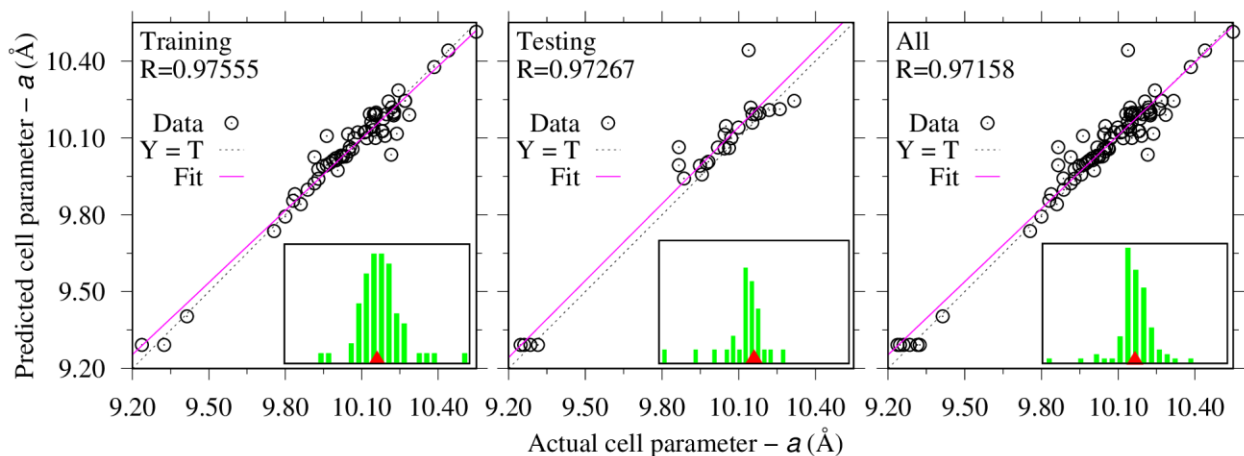


Fig 4: ANN simulation results for the unit cell parameter, c (in Å). All other details are same as described in the caption of Fig. 3.

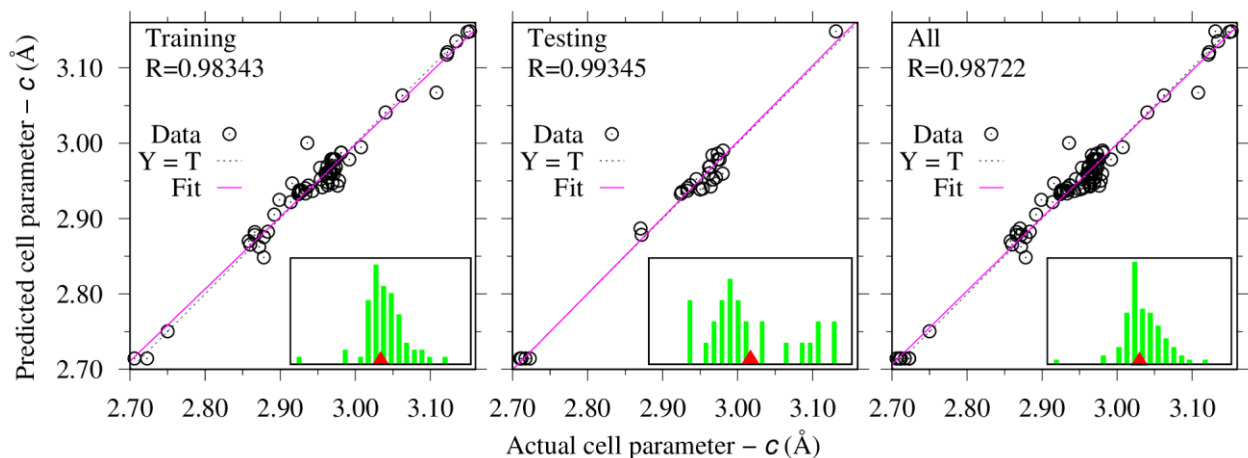


Fig 5: ANN simulation results for the channel size (in Å). All other details are same as described in the caption of Fig. 3.

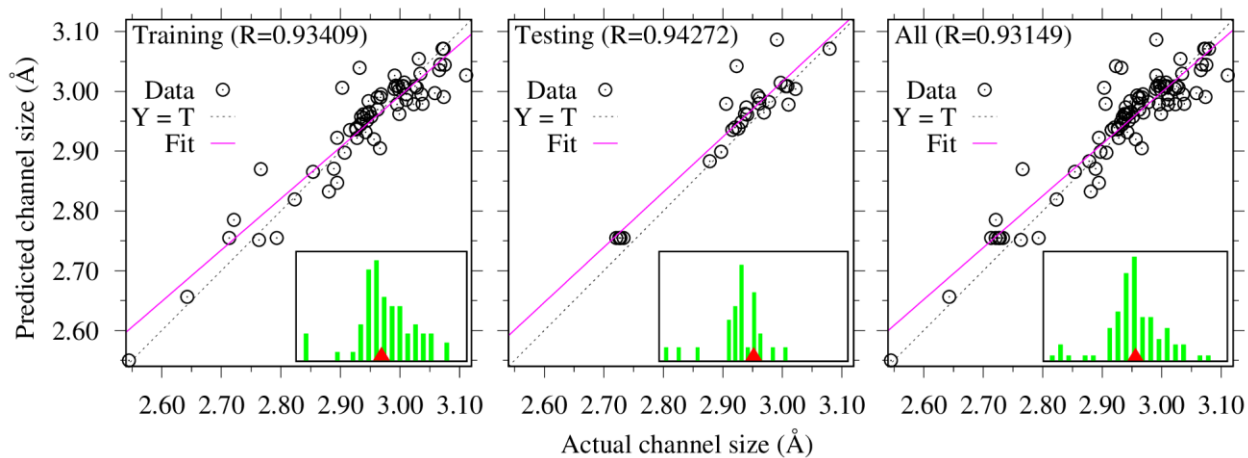


Fig 6: **Left:** Surface plot showing ANN predicted channel size dependence on the A- and B- site radius. The plot considers entire predicted data set. The channel size value is shown by the color coded vertical scale. **Right:** Scatter plot showing channel size variation with A- and B-site radius, obtained with selected Cs content (0.0, 0.2, 0.4, 0.6, 0.8, 1.0). Cs content is color coded for visual purpose and shown by the vertical color coded scale.

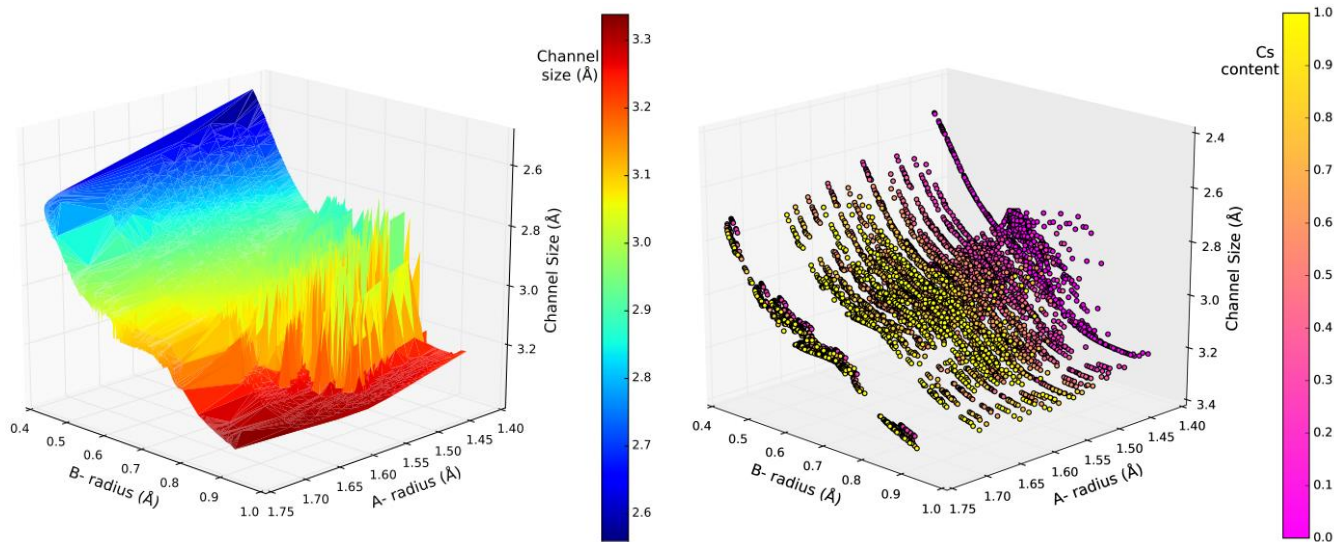


Fig 7: **Left:** ANN predicted channel size – tolerance factor (t_H) profile for some A-site compositions as shown in the legends. B-site compositions comprises of combinations of 3+ and 4+ cations from the entire data set considered here. **Right:** Heat map showing channel size dependence as functions of B-site 3+ and 4+ cations for the selected region ($0.90 \leq t_H \leq 1.10$ and $2.80 \text{ \AA} \leq \text{channel size} \leq 3.15 \text{ \AA}$) as marked by rectangle in the left panel. We note here that for visual purposes we have considered 3.15 \AA as the upper limit of channel size in the plot, instead of 3.12 \AA (from full Cs occupancy consideration) mentioned in the main text. Note the heatmap plot consists of data corresponding to the A-site compositions shown in the left panel and further filtered to data only within the rectangular box.

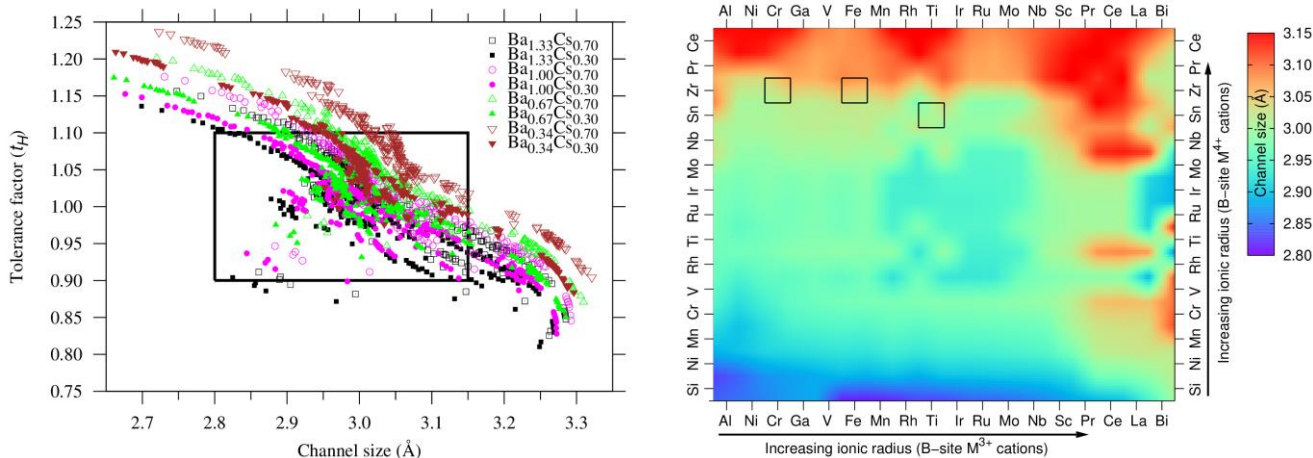


Fig 8: Predicted channel size – B-site composition profile for Ti-based hollandites. The color coded bars represent the B-site radius in Å (*left panel*) and Cs content of A-site (*right panel*). Data with tolerance factor range of 0.90 – 1.10 are considered. For clarity data with Cs content of 0.0, 0.2, 0.4, 0.6, 0.8, 1.00 and Ba content of 0.34 (small symbols) and 1.00 (large symbols) are shown. As indicated by the axis labels for either of the plots, the left y-axis corresponds to $Ti^{3+}M^{4+}$ compositions (solid circles) and right y-axis corresponds to $M^{3+}Ti^{4+}$ compositions (hollow diamonds). The compositions in the shaded region between 2.90 – 3.10 Å may be considered as potential candidates for Ti-based hollandites.

