| Composition | Space group | Number of atoms | Temperatures (K) [$D$ (Å$^2$/fs)] |
|---|---|---|---|
| $Li_{10}GeS_2P_{12}$ | $P4_2mc$ | 200 | 650, 900, 1150, 1400 [0.00216,0.00541,0.00922,0.012905] |
| $Li_2B_{12}H_{12}$ | $Pa_3$ | 832 | 600 [0.00007] |
| $Li_{1.94}B_{12}H_{12}$ | $Pa_3$ | 830 | 600 [0.00025] |
| $Li_2SnS_3$ | $C2/c$ | 384 | 400, 600, 800 [0,0,0] |
| $Li_{1.97}SnS_3$ | $C2/c$ | 382 | 400, 600, 800, 1000, 1200 [0,0,0.00001,0.00005,0.00006] |
| $Li_3La_4Ti_8O_{24}$ | $Pmmm$ | 312 | 500, 700, 900, 1100 [0,0.00001,0.00033,0.00054] |
| $Li_3N$ | $P6/mmm$ | 256 | 300, 500, 700, 900, 1100 [0,0.00010,0.00095,0.00234,0.00407] |
| $Li_{2.94}N_{0.98}$ | $P6/mmm$ | 212 | 300, 500, 700, 900 [0,0.00012,0.00029,0.00547] |
| $Li_3OCl$ | $P4/mmm$ | 320 | 1250 [0] |
| $Li_{2.95}O_{0.98}Cl_{0.98}$ | $P4/mmm$ | 315 | 1000, 1250, 1500, 1650 [0.00007,0.00042,0.00671,0.01747] |
| $Li_6PS_5Br$ | $F\bar{4}3m$ | 208 | 500 [0] |
| $Li_{5.94}PS_5Br$ | $F\bar{4}3m$ | 207 | 500 [0.00001] |
| $Li_7La_3Zr_2O_{12}$ | $I4_1/acd$ | 384 | 400, 600, 800 [0,0,0.00003] |

(Continuation)

| Composition | Space group | Number of atoms | Temperatures (K) [$D$ (Å$^2$/fs)] |
|---|---|---|---|
| $Li_{6.88}La_3Zr_2O_{12}$ | $I4_1/acd$ | 382 | 400, 600, 800 [0.00005,0.00017,0.00037] |
| $Li_7P_3S_{11}$ | $P\bar{1}$ | 336 | 400, 600, 800 [0.00052,0.00270,0.00616] |
| $Li_{6.94}P_3S_{11}$ | $P\bar{1}$ | 335 | 400, 600, 800 [0.00065,0.00304,0.00660] |
| LiF | $Fm\bar{3}m$ | 200 | 800, 1000, 1250 [0,0,0] |
| $Li_{0.98}F_{0.98}$ | $Fm\bar{3}m$ | 196 | 400, 600, 800 [0,0,0.00002] |
| $LiGaO_2$ | $Pna2_1$ | 128 | 300, 500, 700, 900, 1100, 1300 [0,0,0,0,0,0] |
| $Li_{0.97}Ga_{0.97}O_{1.94}$ | $Pna2_1$ | 124 | 500, 700, 900, 1100, 1300, 1500 [0,0,0,0.00002,0.00010,0.00019] |
| $LiIO_3$ | $P6_3$ | 270 | 300, 500, 700, 900, 1100 [0,0,0,0,0.00031] |
| $LiI_{0.94}O_3$ | $P6_3$ | 267 | 300, 500, 700, 900 [0,0,0.00017,0.00022] |
| $LiMn(HCO_2)_3$ | $P2_13$ | 224 | 500 [0] |
| $Li_{0.94}Mn(HCO_2)_3$ | $P2_13$ | 223 | 500, 700, 900 [0,0,0.00001] |
| $LiNbO_3$ | $R3c$ | 180 | 500, 700, 900 [0,0,0] |
| $Li_{0.97}NbO_3$ | $R3c$ | 179 | 500, 700, 900 [0.00001,0.00003,0.00014] |

| (Continuation) | | | |
|---|---|---|---|
| Composition | Space group | Number of atoms | Temperatures (K) [$D$ (Å$^2$/fs)] |
| $LiTi_2(PO_4)_3$ | $P\bar{3}c1$ | 432 | 400, 600, 800 [0,0,0] |
| $Li_{0.92}Ti_2(PO_4)_3$ | $P\bar{3}c1$ | 430 | 400, 600, 800 [0,0,0.00001] |

**Supplementary Table** I: Li-based SSE included in the DFT-AIMD database and details on the corresponding finite-temperature simulations. Ion diffusion coefficients, $D$, listed as "0" are $< 10^{-5}$ Å$^2$/fs.

| Composition | Space group | Number of atoms | Temperatures (K) [$D$ (Å$^2$/fs)] |
|---|---|---|---|
| $Na_2B_4O_7$ | $P\bar{1}$ | 312 | 300, 450, 600, 750 [0,0,0,0] |
| $Na_3SbS_4$ | $I\bar{4}3m$ | 288 | 500, 700, 900 [0,0,0.00003] |
| $NaBH_4$ | $Pnma$ | 288 | 300, 450, 600, 750 [0,0,0,0] |
| $Na_{0.98}BH_4$ | $Pnma$ | 287 | 300, 450, 600, 750 [0.00003,0.00006,0.00018,0.00027] |
| $NaBO_2$ | $C2/c$ | 288 | 750 [0] |
| $NaMn(HCO_2)_3$ | $P2_13$ | 224 | 500 [0] |
| $Na_{0.94}Mn(HCO_2)_3$ | $P2_13$ | 223 | 500, 700, 900 [0,0,0.00001] |
| $NaZr_2(PO_4)_3$ | $C2/c$ | 216 | 500 [0] |
| $Na_{0.92}Zr_2(PO_4)_3$ | $C2/c$ | 215 | 500 [0] |
| $AgCrSe_2$ | $P3m1$ | 192 | 500 [0] |
| $Ag_{0.98}CrSe_2$ | $P3m1$ | 191 | 500 [0.00001] |
| AgI | $F\bar{4}3m$ | 250 | 200, 350, 500, 650 [0,0.00154,0.00408,0.00692] |

(Continuation)

| Composition | Space group | Number of atoms | Temperatures (K) [$D$ (Å$^2$/fs)] |
|---|---|---|---|
| $Cu_{1.99}Se$ | $Fm\bar{3}m$ | 215 | 300, 500, 700, 900 [0.00002,0.00061,0.00281,0.00555] |
| CuI | $F\bar{4}3m$ | 216 | 500, 700, 900 [0.00072,0.00271,0.00746] |

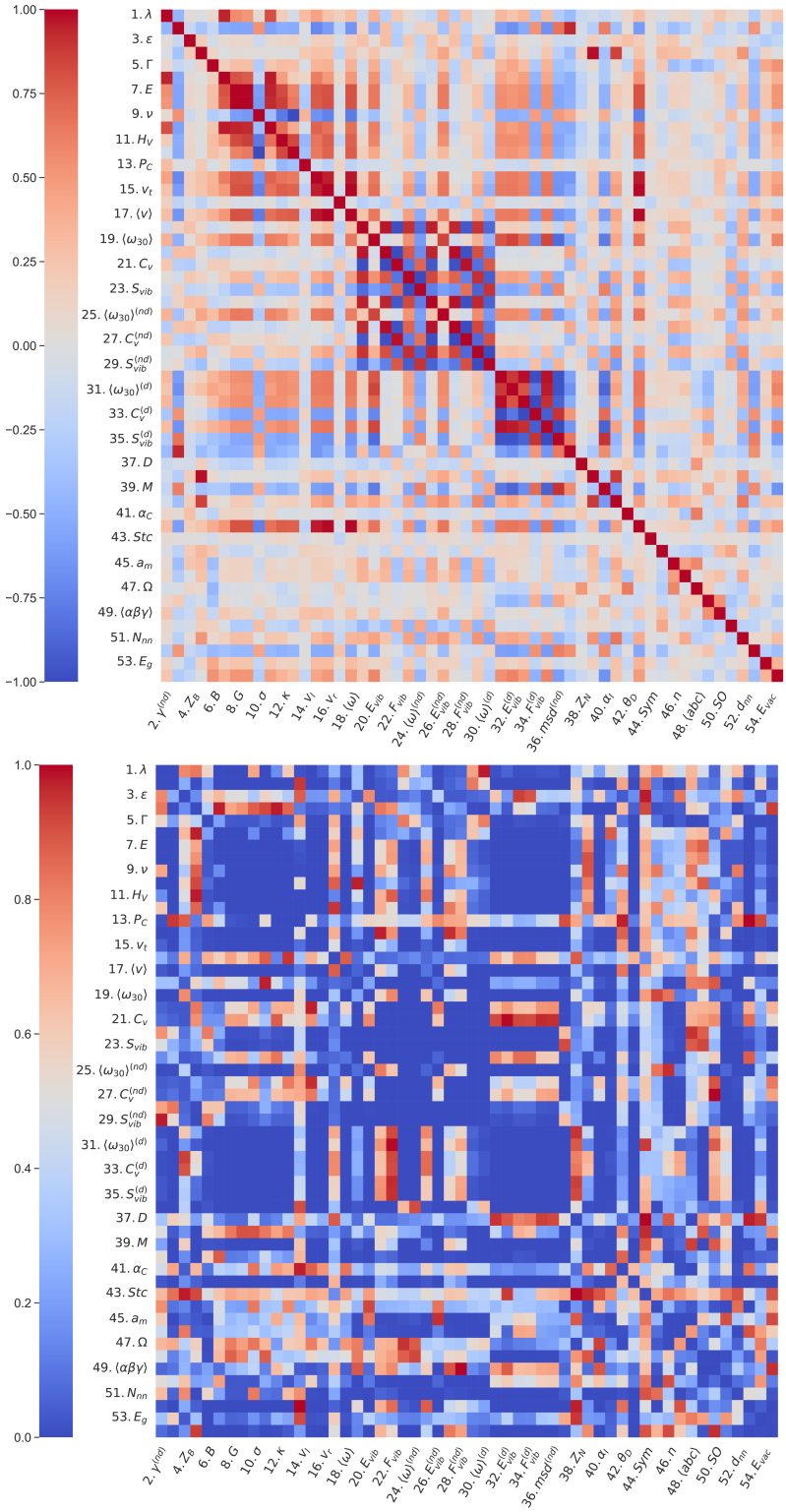
**Supplementary Table** II: Na-, Ag- and Cu-based SSE included in the DFT-AIMD database and details on the corresponding finite-temperature simulations. Ion diffusion coefficients, $D$, listed as "0" are $< 10^{-5}$ Å$^2$/fs.

| Composition | Space group | Number of atoms | Temperatures (K) [$D$ (Å$^2$/fs)] |
|---|---|---|---|
| $BaF_2$ | $Fm\bar{3}m$ | 192 | 1000, 1250, 1375, 1500 [0,0.00005,0.00406,0.00688] |
| $CaF_2$ | $Fm\bar{3}m$ | 192 | 500, 1000, 1250, 1500 [0,0,0.00006,0.00322] |
| $CH_3NH_3PbBr_3$ | $Pm$ | 324 | 250, 400 [0,0] |
| $CH_3NH_3PbBr_{2.96}$ | $Pm$ | 323 | 400, 600, 800 [0,0.00001,0.00366] |
| $CH_3NH_3PbIBr_2$ | $Pm$ | 324 | 250, 400 [0,0] |
| $CH_3NH_3PbI_{0.93}Br_2$ | $Pm$ | 322 | 400 [0] |
| $CH_3NH_3PbI_2Br_1$ | $Pm$ | 324 | 100, 250, 400 [0] |
| $CH_3NH_3PbI_2Br_{0.93}$ | $Pm$ | 322 | 400 [0] |
| $CH_3NH_3PbI_3$ | $Pm$ | 324 | 100, 250, 400 [0,0,0] |
| $CH_3NH_3PbI_{2.96}$ | $Pm$ | 323 | 400, 600, 800 [0,0.00016,0.00407] |
| $CsPbBr_3$ | $Pm\bar{3}m$ | 180 | 500, 700, 850 [0,0.00018,0.00047] |
| $CsPbBr_{2.97}$ | $Pm\bar{3}m$ | 179 | 500, 700, 900 [0.00009,0.00016,0.00028] |
| $SrCl_2$ | $C2/m$ | 192 | 1500, 2000, 2500 [0, 0.00001,0.00873] |
| $SrF_2$ | $Fm\bar{3}m$ | 192 | 1250, 1375, 1500, 1750 [0,0,0.00111,0.00856] |

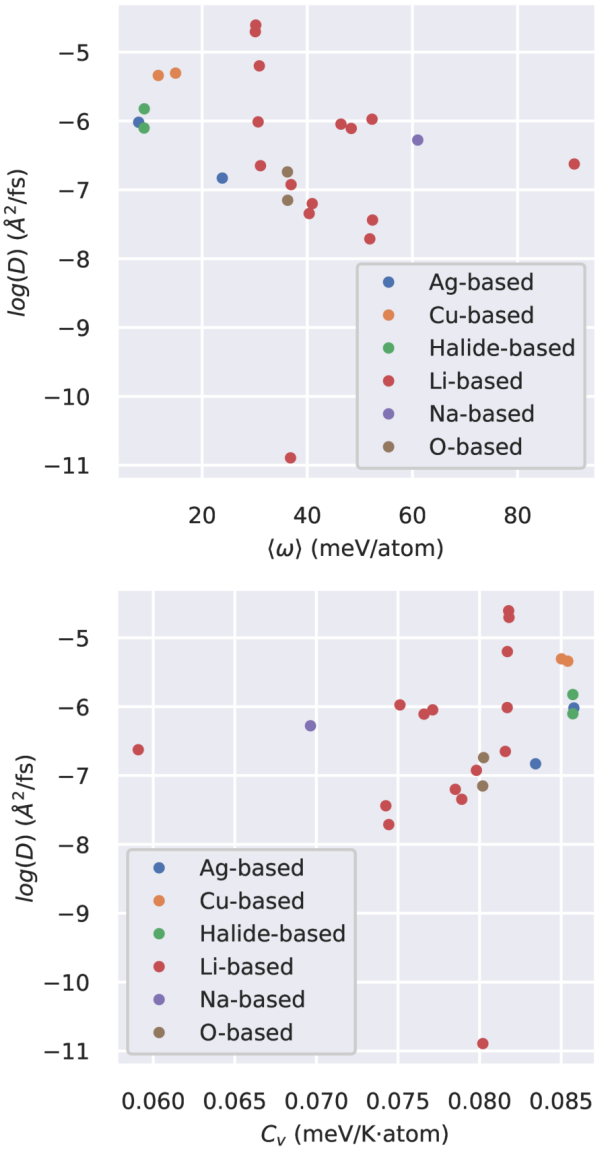
(Continuation)

| Composition | Space group | Number of atoms | Temperatures (K) [$D$ (Å$^2$/fs)] |
|---|---|---|---|
| $Bi_2O_3$ | $Pn\bar{3}m$ | 180 | 500 |
| | | | [0] |
| $LaGaO_{2.98}$ | $C2/c$ | 239 | 500, 700, 900 |
| | | | [0,0,0.00001] |
| $SrCoO_3$ | $Pbam$ | 160 | 200, 400 |
| | | | [0,0] |
| $SrCoO_{2.75}$ | $Pbam$ | 152 | 200, 400 |
| | | | [0.00001,0.00002] |
| $SrTiO_{2.95}$ | $Pc$ | 317 | 1000, 1500, 2000 |
| | | | [0,0.00002,0.00003] |

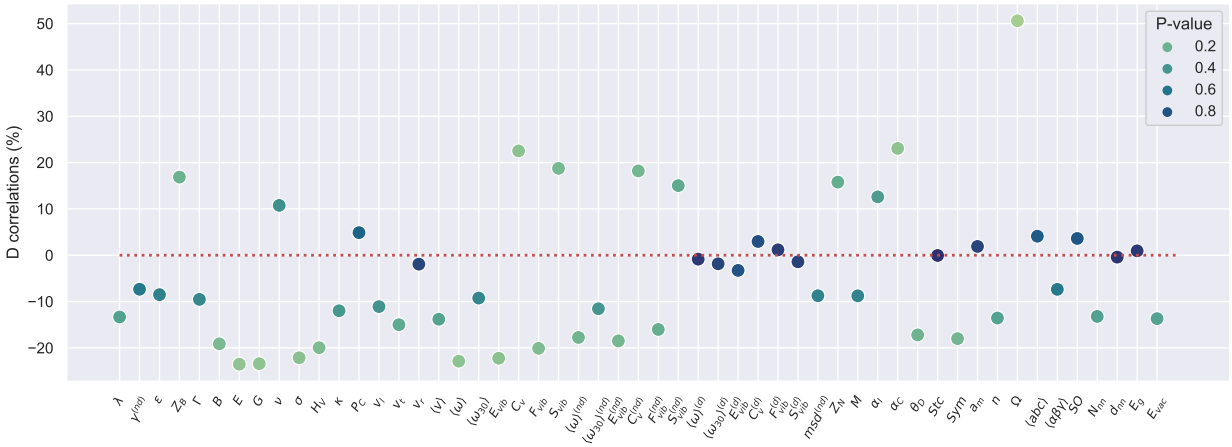
**Supplementary Table** III: Halide- and O-based SSE included in the DFT-AIMD database and details on the corresponding finite-temperature simulations. Ion diffusion coefficients, $D$, listed as "0" are $< 10^{-5}$ Å$^2$/fs.



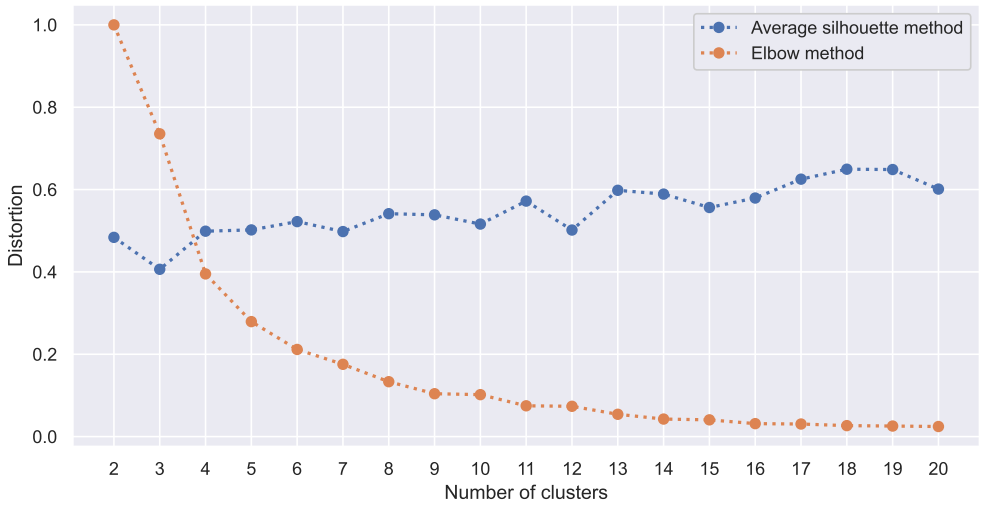

**Supplementary Figure** 1: Pearson correlogram calculated for all SSE (Top) and the corresponding $p$-value matrix (Bottom).



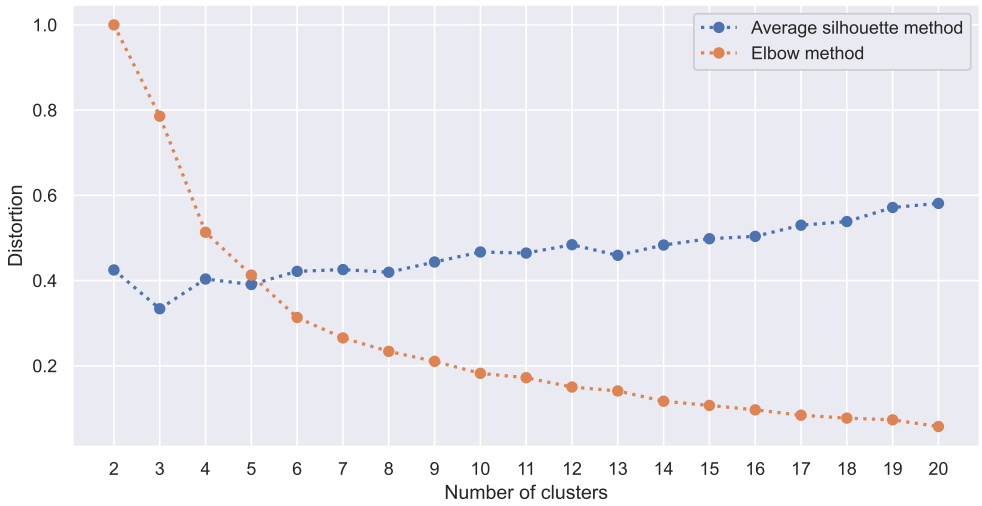

**Supplementary Figure** 2: Representation of the ion diffusion coefficient and other materials descriptors calculated at $T = 500 \pm 100$ K for all the materials contained in the SSE DFT-AIMD database. The diffusion coefficient is represented in logarithmic scale.



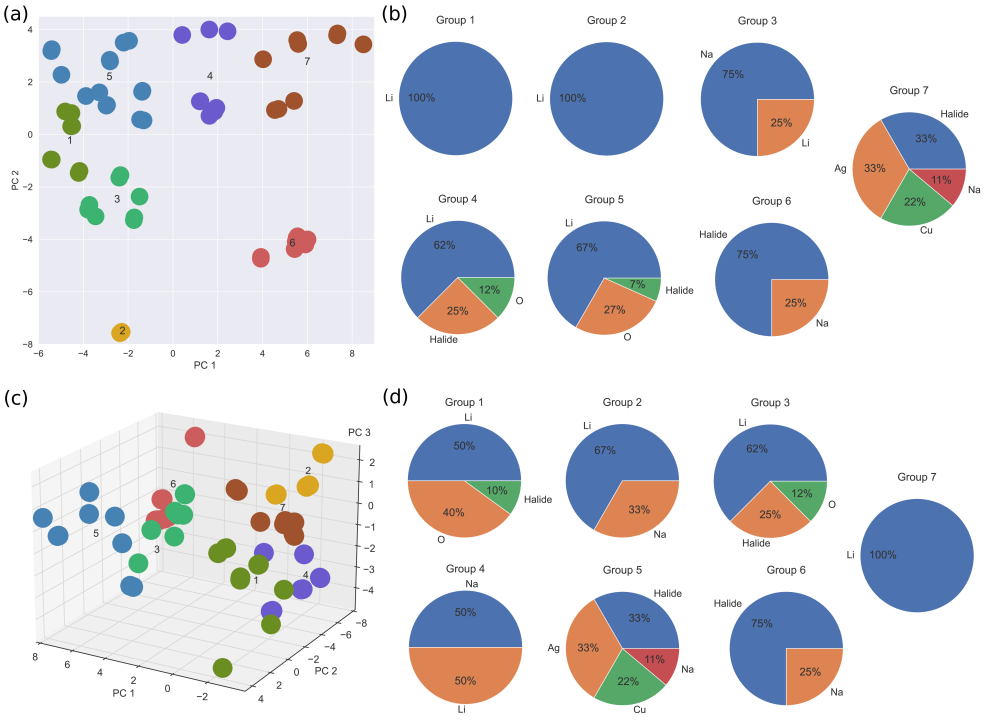

**Supplementary Figure** 3: Pearson $D$ correlations estimated for all SSE.



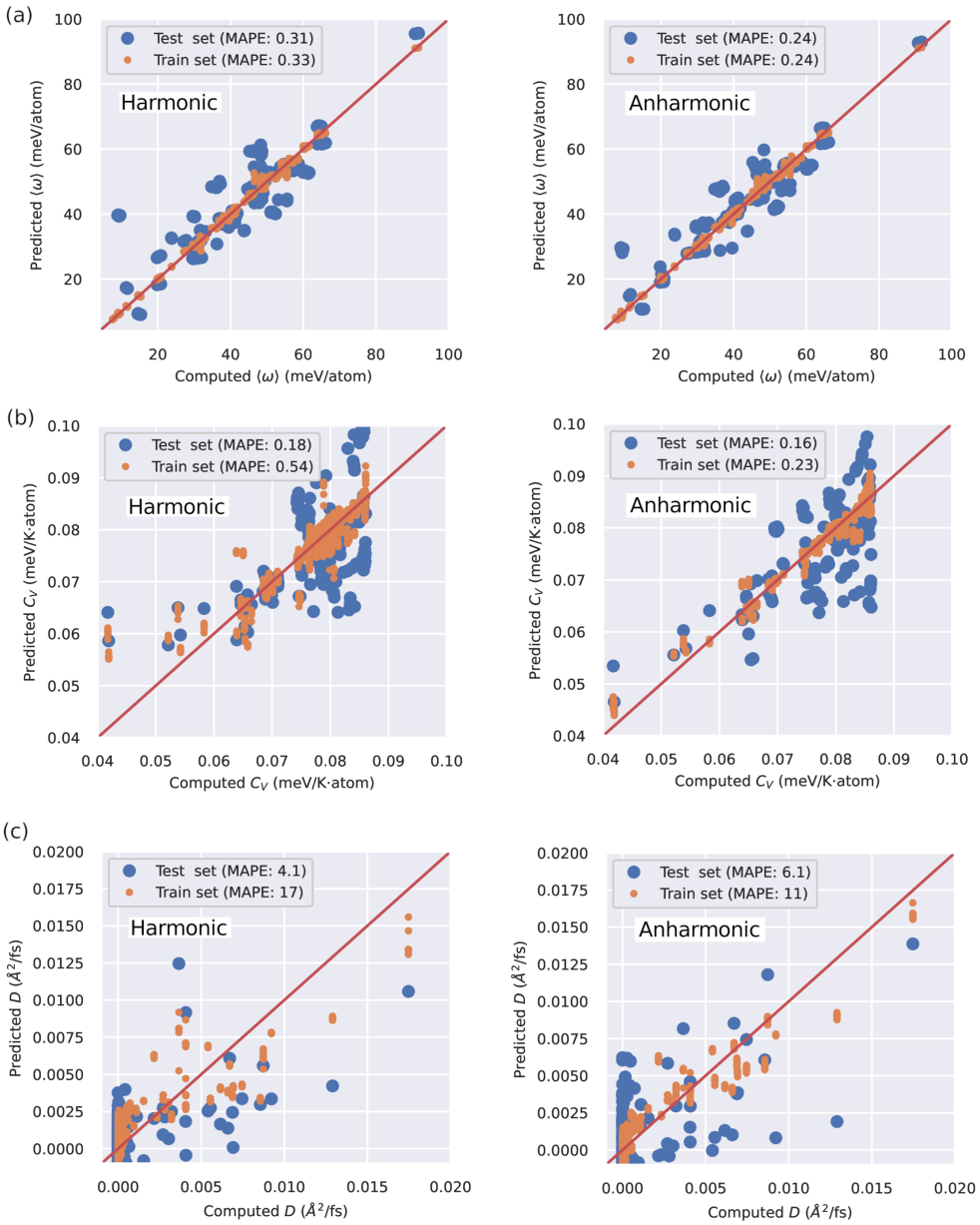

**Supplementary Figure** 4: K-means clustering analysis convergence for the two-dimensional PC1–PC2 representation. The elbow method is used to find the "elbow" point where adding additional data samples does not change cluster membership much. Silhouette score determines whether there are large gaps between each sample and all other samples within the same cluster or across different clusters. A subminimal number of 7 clusters was selected to account for the SSE database variance.




**Supplementary Figure** 5: K-means clustering analysis convergence for the three-dimensional PC1–PC2–PC3 representation. The elbow method is used to find the "elbow" point where adding additional data samples does not change cluster membership much. Silhouette score determines whether there are large gaps between each sample and all other samples within the same cluster or across different clusters. A subminimal number of 7 clusters was selected to account for the SSE database variance.




**Supplementary Figure** 6: K-means clustering analysis results obtained for the SSE DFT-AIMD database considering 7 clusters. (a)-(b) Classification of the analyzed materials in the orthogonal bidimensional space PC1–PC2. (c)-(d) Classification of the analyzed materials in the orthogonal tridimensional space PC1–PC2–PC3. The position of the cluster numbers in the PC plots coincide with the clusters centroids.




**Supplementary Figure** 7: Machine learning (ML) models trained in our DFT-AIMD database for prediction of different SSE $T$-dependent quantities, considering a kernel-ridge regression algorithm. The ML models were trained by considering and neglecting AIMD-based vibrational descriptors that explicitly incorporate anharmonic effects, labelled as "anharmonic" and "harmonic", respectively. K-fold validation results obtained for (a) $\langle\omega\rangle$, (b) constant volume heat capacity and (c) ionic diffusion coefficient obtained from AIMD simulations. "MAPE" stands for the mean absolute percentage error of the ML predictions (Methods).