A machine learning interatomic potential for the structural properties of hematite: supplemental material

IMMEDIATE

1. EMPIRICAL POTENTIALS

A. Interface

The Interface Force Field (IFF) is a set of parameters and a methodology that is compatible with the functional forms of major existing force fields [1]. In this sense, there exist many variants for the IFF and in this work we have used the "IFF-CHARMM", which refers to the use of IFF-derived parameters within a simulation engine that uses the CHARMM force field functional forms [2]. The interaction between species i and j is composed by a van der Waals term (Lennard-Jones) plus an electrostatic contribution, namely,

$$U_{ij} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{e^2}{4\pi\epsilon_0}\frac{z_i z_j}{r_{ij}^2}, \tag{S1}$$

where $1/4\pi\epsilon_0$ is the usual Coulomb constant and r_{ij} is the center-to-center separation between atoms i and j. Charges and Lennard-Jones (LJ) parameters were derived by Kanhaiya *et al.* [3]. Charge numbers are $z_{\mathrm{Fe}} = 1.74$ and $z_{\mathrm{O}} = -1.16$ and $e = 1.602 \times 10^{-19}$ C is the elementary charge. The other parameters are summarized in Table S1.

B. Clay

The functional form of the Clay potential is identical to IFF's: it is composed by non-bonded interactions only, a sum of a LJ term and a Coulombic contribution [see Eq. (S1)] [4]. Charge numbers for Fe and O atoms in the Clay parametrization are set to 1.575 and -1.05, respectively. The LJ parameters are mistyped in its original reference (Ref. [4]) and in this work we have used their corrected values [5], shown in Table S2.

C. Tersoff

The potential we refer as Tersoff belongs to a class of analytical bond-order potentials (ABOP), with origins in Abell's [6] discussion of bond order. Tersoff [7, 8] first developed these ideas into an empirical potential for covalent materials [9, 10]. The ABOP formalism closely follows Tersoff's work but is extended in a style similar to Brenner's [11]. Notably, this entire family of potentials for covalent systems is fundamentally connected to the embedded atom method (EAM) potentials [12, 13], used for metals [14].

In this work, we used a slightly modified version of Tersoff equations, fully described in Byggmästar *et al.* work, were parameters for the Fe-O interaction were also derived [15]. Parameters for Fe-Fe were developed by Müller *et al.* [16] and Bjorkas and Nordlund [17], while the O-O interaction parameters are from Erhart *et. al* [18].

Table S1. Parameters for the Interface Force Field, Eq. (S1), used in this work.

Pair	ε_{ij}(kcal/mol)	σ_{ij}(Å)
Fe-Fe	0.2	1.8085
O-O	0.2	3.1003
Fe-O	0.2	2.4544

Table S2. Parameters for Clay potential, Eq. (S1), used in this work.

Pair	ε_{ij} (kcal/mol)	σ_{ij} (Å)
Fe-Fe	9.0298×10^{-7}	4.9062
O-O	0.1554	3.1650
Fe-O	3.7459×10^{-8}	4.0356

D. Reax

The Reactive Force Field, in this work denoted as Reax, was developed by Adri van Duin, William A. Goddard III, and collaborators. It is a bond-order-based interatomic interaction designed to model chemical reactions within classical molecular dynamics simulations [19].

The Reax methodology is founded on a general relationship between bond distance and bond order, and subsequently between bond order and bond energy, which ensures the correct dissociation of bonds to separated atoms. The contributions from other valence terms, such as bond angles and torsions, are also formulated as a function of bond order, causing them to decay smoothly to zero during dissociation events. Non-bonded interactions are handled by Coulomb and Morse (van der Waals) potentials applied between all atom pairs, with no exclusions. To avoid nonphysical behavior at close proximity, these non-bonded interactions are shielded at short ranges, causing them to plateau at a constant value as the interatomic distance approaches zero.

For the complete mathematical description of the Reax Force Field, we invite the reader to check the aforementioned original references. For the Reax parameters used in this work we used the Shin *et. al* derivations [20].

E. Core-Shell

The core-shell model is a classical approach that extends the rigid ion Born model of solids [21] to include the effects of electronic polarizability [22]. A polarizable ion, typically the anion (O^{2-}), is modeled as a massive core (representing the nucleus and core electrons) connected to a massless shell (representing the valence electrons) by a harmonic spring [23]. The polarizability of the ion is thus determined by the charges of the core and shell and the stiffness of the spring constant connecting them. Cations are generally treated as unpolarizable point charges [23]. Table S3 details the specific charges assigned to each particle type.

Table S3. Particle charges for the Core-Shell model [23].

Particle	Charge (e)
Fe	+3.00
O_{core}	+0.21
O_{shell}	-2.21

Interatomic forces in the model are described by a combination of long-range and short-range potentials. The long-range electrostatic interactions are calculated between all cores and shells using a Coulomb potential. Short-range interactions are modeled with a pair-potential function acting between the shells of neighboring ions. For this model, a Buckingham potential is used, with the functional form:

$$U_{ij}(r_{ij}) = A_{ij} \exp(-r_{ij}/\rho_{ij}) - \frac{C_{ij}}{r_{ij}^6}. \tag{S2}$$

These short-range potentials are defined only for interactions involving shells, as detailed in Table S4. Core-core interactions are purely Coulombic.

Finally, the intramolecular interaction between an oxygen core and its own shell is represented by the aforementioned harmonic spring [23]. The spring constant for this core-shell (cs) bond is k_{cs} = 27.4 eV/Å^2. Note that for implementation in the simulation package LAMMPS (lmp)

Table S4. Core-shell parameters for the Buckingham contribution (Eq. S2) [23].

Pair	A_{ij} (eV)	ρ_{ij} (Å)	C_{ij} (eV·Å^6)
Fe-Fe	0.0	–	0.0
Fe–O_{core}	0.0	–	0.0
Fe-O_{shell}	1102.4	0.3299	0.0
O_{core}-O_{core}	0.0	–	0.0
O_{core}-O_{shell}	0.0	–	0.0
O_{shell}-O_{shell}	22764.0	0.1490	27.88

(through the `harmonic` bond style), this corresponds to a bond coefficient $k_{\mathrm{lmp}} = 13.7$ eV/Å^2, since $k_{\mathrm{lmp}} = k_{\mathrm{cs}}/2$.

F. Surface Energy Calculation

The surface energies (γ) were determined following an established methodology based on slab models. The procedure first required establishing a reference bulk energy per formula unit (E_{b}). This was calculated from a $3 \times 3 \times 1$ supercell containing $n = 54$ Fe_2O_3 units, constructed from a fully relaxed primitive cell. Subsequently, symmetric slabs were cleaved for the (0001) and (01$\bar{1}$2) surfaces, with respective surface areas (A) of 208.7 Å^2 and 242.6 Å^2. The slabs were modeled with a thickness of approximately 40.0 Å to ensure convergence of the surface properties. To prevent interactions between periodic images, a vacuum gap of 100 Å was added. After performing a full geometric relaxation of the slab atoms, the resulting total energy (Es) was used to derive the surface energy as:

$$\gamma = \frac{E_{\mathbf{s}} - nE_{\mathrm{b}}}{2A}.$$

G. Relative Deviations for Calculated Quantities

Here we present tables containing the relative deviation for the structural and elastic properties calculated in this work, shown in tables 1-6 of the main text.

Table S5. Average deviations from experimental values (%) for lattice parameters of hematite (see Tab. 1 of main text). $\langle\sigma\rangle$ is the absolute average deviation [see Eq. (1) of the main text]

Parameter	Tersoff	Reax	CS	IFF	Clay	Fe-MLIP	DFT
a	2.70	-0.89	0.50	0.14	0.93	2.44	2.52
c	-7.01	-13.27	-2.98	-0.90	0.29	1.92	0.49
c/a	-9.44	-12.48	-3.44	-1.02	-0.62	-0.48	-1.28
$\langle\sigma\rangle$	6.38	8.88	2.31	0.69	0.61	1.61	1.43

2. ELASTIC TENSOR AND ELASTIC MODULI OF HEMATITE

The primitive unit cell of hematite belongs to the trigonal crystal system and to the $R\bar{3}c$ (nr. 167) space group. For this symmetry, the elastic tensor **C** can be expressed in matrix form (in Voigt notation) as

Table S6. Average deviations from experimental values (%) for elastic constants of hematite (see Tab. 2 of main text).

Parameter	Tersoff	Reax	CS	IFF	Clay	Fe-MLIP	DFT
C_{11}	2.02	68.67	43.64	17.45	-11.18	-16.35	-5.14
C_{33}	27.14	97.09	34.36	28.88	0.90	-11.60	-6.32
C_{44}	35.00	102.07	110.17	144.31	71.03	31.55	24.31
C_{12}	-1.68	-4.00	11.30	1.95	-27.89	-24.65	-8.59
C_{13}	-21.63	-55.35	-29.37	-8.50	-36.29	-44.75	-31.81
$\lvert C_{14}\rvert$	142.31	865.38	325.00	-87.50	-43.27	11.54	-45.19
$\langle\sigma\rangle$	38.30	198.76	92.31	48.10	31.76	23.41	20.23

Table S7. Average deviations from experimental values (%) for elastic moduli and related quantities for hematite (see Tab. S10).

Parameter	Tersoff	Reax	CS	IFF	Clay	Fe-MLIP	DFT
Y	0.04	83.52	51.04	34.14	2.16	-12.75	-11.27
K	14.65	35.52	29.26	27.41	-6.50	-14.02	-0.24
G	-2.64	91.76	53.41	33.96	2.64	-13.30	-13.52
β	-14.29	-26.53	-22.45	-22.45	6.12	14.29	0.00
ν	-13.16	-36.84	-26.32	-21.05	-23.68	-18.42	-13.16
K/G	18.14	-29.20	-15.49	-4.87	-8.85	-0.44	15.49
H	-14.06	148.27	73.66	37.43	19.55	-3.81	-18.95
$\langle\sigma\rangle$	11.00	64.52	38.80	25.90	9.93	11.00	10.38

Table S8. Average deviations from experimental values (%) for lattice and elastic parameters of maghemite (γ-Fe_2O_3) (see Tab. 5 of main text).

Parameter	Tersoff	CS	IFF	Clay	Fe-MLIP
a	0.29	1.05	6.04	4.08	2.59
c	0.68	-0.60	-12.0	-1.65	3.19
c/a	0.40	-1.64	-17.0	-5.52	0.60
C_{11}	-59.7	8.88	34.91	-20.6	-28.2
C_{12}	-43.3	26.90	48.23	3.36	-4.06
C_{13}	-46.0	3.85	-16.5	-13.1	-8.83
C_{33}	-49.6	-42.6	-61.9	-74.6	-36.3
C_{44}	-76.2	-30.7	-30.8	-41.8	-34.8
C_{66}	-76.7	27.54	-15.8	-28.3	-30.0
K	-45.1	-6.62	-6.65	-44.3	-9.86
$\langle\sigma\rangle$	39.8	15.0	25.0	23.7	15.8

Table S9. Average deviations from experimental values (%) for lattice and elastic parameters of magnetite (Fe_3O_4) (see Tab. 6 of main text).

Parameter	Tersoff	Reax	Fe-MLIP
a	2.12	-3.31	2.82
C_{11}	-51.6	-6.63	-13.5
C_{12}	-36.4	-53.2	-13.1
C_{44}	-74.4	8.12	1.87
Y	-66.4	19.22	2.81
K	-42.7	-27.5	-11.1
G	-63.7	49.09	19.73
β	87.27	49.09	20.00
ν	25.81	-32.3	-6.45
K/G	163.7	-18.5	23.81
H	-76.0	147.6	31.11
$\langle\sigma\rangle$	53.6	37.7	13.3

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & 0 & 0 \\ C_{12} & C_{11} & C_{13} & -C_{14} & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ C_{14} & -C_{14} & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & C_{14} \\ 0 & 0 & 0 & 0 & C_{14} & \frac{1}{2}(C_{11} - C_{12}) \end{pmatrix}.$$

The relationship between stress (σ_j) and strain (ε_i) follows Hooke's law, $\sigma_i = C_{ij}\varepsilon_j$, where C_{ij} are the elements of $\mathbf{C}$ [24]. Indices refer to the standard directions in strain and stress theory, as defined in Voigt notation [25], specifically $xx = 1$, $yy = 2$, $zz = 3$, $yz = 4$, $xz = 5$, and $xy = 6$. The inverse relationship is given by $\varepsilon_i = S_{ij}\sigma_j$, yields the compliance tensor. The explicit equations for S_{ij} in terms of C_{ij} are

$$\begin{aligned} S_{11} &= (C_{22}C_{33} - C_{23}^2)/\Delta \\ S_{22} &= (C_{11}C_{33} - C_{13}^2)/\Delta \\ S_{33} &= (C_{11}C_{22} - C_{12}^2)/\Delta \\ S_{12} &= (C_{13}C_{23} - C_{12}C_{33})/\Delta \\ S_{13} &= (C_{12}C_{23} - C_{13}C_{22})/\Delta \\ S_{23} &= (C_{12}C_{13} - C_{11}C_{23})/\Delta \\ S_{44} &= \frac{1}{C_{44}}, \quad S_{55} = \frac{1}{C_{55}}, \quad S_{66} = \frac{1}{C_{66}}, \end{aligned} \tag{S3}$$

where $\Delta = C_{11}C_{22}C_{33} + 2C_{12}C_{13}C_{23} - C_{11}C_{23}^2 - C_{22}C_{13}^2 - C_{33}C_{12}^2$.

For obtaining the Bulk (K) and Shear (G) moduli we used the Voigt-Reuss-Hill averaging scheme [26].

$$\begin{aligned} K &= \frac{K_\mathrm{v} + K_\mathrm{r}}{2} \\ G &= \frac{G_\mathrm{v} + G_\mathrm{r}}{2}, \end{aligned} \tag{S4}$$

where

$$K_{\mathrm{v}} = \frac{f_1 + 2f_2}{9}$$
$$K_{\mathrm{r}} = \frac{1}{\tilde{f}_1 + 2\tilde{f}_2}$$
$$G_{\mathrm{v}} = \frac{f_1 - f_2 + 3f_3}{15}$$
$$G_{\mathrm{r}} = \frac{15}{4\tilde{f}_1 - 4\tilde{f}_2 - 3\tilde{f}_3}$$

are their corresponding Voigt (v) [25] and Reuss (r) [27] approximations, with auxiliary functions f_k and $\tilde{f}_k$ given by

$$f_1 = C_{11} + C_{22} + C_{33}$$
$$f_2 = C_{12} + C_{13} + C_{23}$$
$$f_3 = C_{44} + C_{55} + C_{66}$$
$$\tilde{f}_1 = S_{11} + S_{22} + S_{33}$$
$$\tilde{f}_2 = S_{12} + S_{13} + S_{23}$$
$$\tilde{f}_3 = S_{44} + S_{55} + S_{66}.$$

The elements C_{ij} were obtained from simulations and S_{ij} are calculated through Eqs. (S3).

The isotropic Young's modulus (Y), Poisson's ratio (ν), and compressibility (β) can be derived from the bulk modulus (K) and shear modulus (G) – as shown in Eq. (S4) – using the standard relations:

$$Y = \frac{9KG}{3K + G}$$
$$\nu = \frac{3K - 2G}{6K + 2G}$$
$$\beta = \frac{1}{K}.$$

Table S10. Elastic moduli and related quantities for empirical potentials and our trained ML potential (Fe-MLIP) for hematite. DFT data are from Ref. [28] and experimental data is a calculated consensus from literature (see the Suppl. Mat. for details). Young's (Y), Bulk (K) and Shear (G) moduli and Vickers hardness (H) are given in GPa units, while the compressibility (β) in GPa^{-1}. Poisson (ν) and Pugh (K/G) ratios are dimensionless. Colors help to visualize overestimation (red) and underestimation (blue) in relation to experiments.

Parameter	Tersoff	Reax	CS	IFF	Clay	Fe-MLIP	DFT	Exp.
Y	236.2	433.3	356.6	316.7	241.2	206.0	209.5	236.1
K	236.3	279.3	266.4	262.6	192.7	177.2	205.6	206.1
G	88.6	174.5	139.6	121.9	93.4	78.9	78.7	91.0
β	0.0042	0.0036	0.0038	0.0038	0.0052	0.0056	0.0049	0.0049
ν	0.33	0.24	0.28	0.30	0.29	0.31	0.33	0.38
K/G	2.67	1.60	1.91	2.15	2.06	2.25	2.61	2.26
H	7.21	20.83	14.57	11.53	10.03	8.07	6.80	8.39

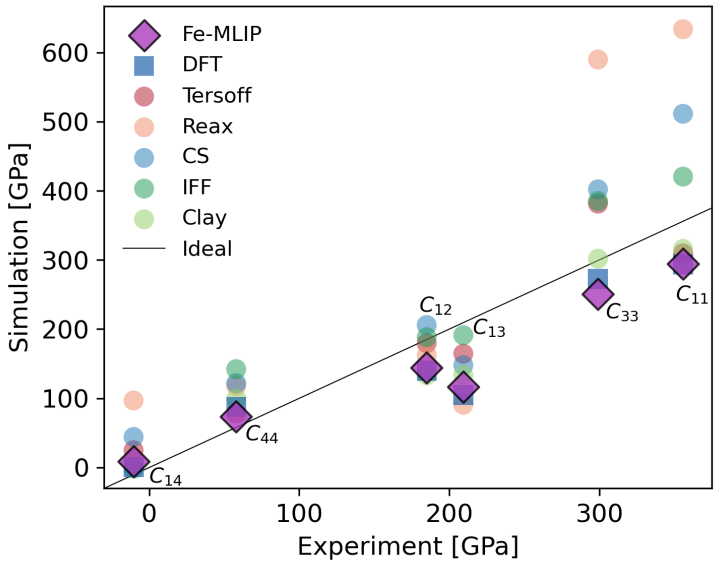

Fig. S1. Elastic tensor constants, C_{ij} for hematite calculated with different potentials. The continuous line (ideal) is given by Liebermann's experimental results [29].

Table S10 shows the bulk, K, Young, Y, and shear, G, moduli, the compressibility, β, Poisson's ratio, ν, Pugh ratio, G/K, and the Vickers hardness, H, for hematite.

Figure S1 shows a parity plot of the elastic tensor elements, C_{ij}, comparing the results of the potentials to experimental data (Ref. [29]).

3. EXPERIMENTAL ELASTIC PROPERTIES CONSENSUS

A. Hematite

Surprisingly, considering the economic importance of hematite, we have found relatively few works in the literature addressing its elastic properties. In Table S11 we summarize the experimental results for the lattice constants and elastic moduli found in the literature. As can be seen, the property that has been measured the most is the isotropic bulk modulus (K), while the Young (E) and shear (G) moduli have been scarcely investigated. The experimental results shown in Table S11 were measured over a range of different conditions and samples. Thus, we use the bootstrap median estimation of the properties reported in the literature as a consensus for the reported experimental data, and use these estimates as a reference. We performed the median estimation by sampling the median with replacement ($n = 10000$) over the experimental data, and consider the uncertainty as two standard deviations, approximating a 95% confidence interval. This procedure for the median estimation is robust against outliers, and allows us to estimate the uncertainty associated with the different experiments as well.

Table S11. Experimental lattice constants, a and c (in Å), the elastic moduli for hematite (in GPa): namely, the bulk modulus and its volume derivative, K and K', the Young modulus, E, the shear modulus, G, and the Poisson ratio, ν (adimensional). Values indicated with apostrophes (′) were considered fixed by the authors. [a]) Calculated by us from the elastic tensor constants c_{ij}. [b]) Obtained by the relationships between elastic moduli.

Ref.	Material % purity	Conditions	a	c	K	K'	E	G	ν
[30][a]					99.7		216.3	95.0	0.139
[31]		25.0°C, 1 bar			202.66	4.53	237.46[b]	0.386	
[32]		298 K					260		
[33]	Synthetic	<30 kbar			231	4.0′			
		>30 kbar			178	4.0′			
[34]		25°C, 1 bar			199(6)	4.0′			

Table S11 – Continued from previous page

Ref.	Material % purity	Conditions	a	c	K	K'	E	G	ν
[35]		298 K, 1 bar	5.0347(4)	13.7473(15)	225(4)	4.0′			
[36]	High purity	298 K, 1 bar			246.1(4)	4.5(2)	179.98[b]	65.3(1.5)	0.378[b]
[37]					212		250.2[b]	96	0.303[b]
[38]	Powder	P=0	5.031(7)	13.765(7)	230(5)	3.5(6)			
[39]	Bulk	298 K, 1 bar	5.034(5)	13.747(2)	205(5)	4.0′			
	Nanocrystal				211(10)	4.0′			
[40]	Synthetic 99.998%				202.66(20)				
[41]	99.99%		5.049(5)	13.73(2)	206.1(4.8)	4.0′			
					202.1(3.8)	4.3(3)			
[42]	Bulk	298 K, 1 bar			300(30)				
	7 nm nanopart.				336(5)				
[43]		298 K, 1 bar	5.060(5)	13.75(2)	268(14)	4.0′			
[44]	Powder, 99.99%				201(4)	4.3(2)			
[45]	Powder, 99.99%				201	4.3′			
[46]	3% magnetite, ∼3.2% poros.	298 K	5.0350(1)	13.7511(3)	184.0(9)		214.8(3)	82.00(7)	0.306
					201		228	87	0.311
[47]	99.999%	$P < 25.3$ GPa	5.0354(17)	13.7477(48)	207(3)	4.0′			
					199(8)	5.3(9)			
[48]			5.0993	13.8040	235.4(8)	3.29(25)	234.748[b]	88.0	0.334[b]

In Figures S2 and S3 we show violin plots for the distributions of the lattice constants and elastic moduli reported in the literature (shown in Table S11) along with the *consensus* estimation.

For the elastic tensor constants, C_{ij}, only a handful of works are available in the literature, shown in Table S12. In this case, it should be noted that we have found only three experimental measurements, with one of them (Voigt [30]) being substantially softer (lower) than the other two. This disparity is also clear for the lattice constants (max. values) shown in Fig. S2, and for the elastic moduli (lower values) shown in Fig. S3. In that regard, Voigt pointed out that the sample for hematite "was more scarcer and not entirely free of defects" (free translation from Ref. [30]). Other authors agreed that Voigt's sample was questionable [31, 49]. Despite this, these results were used as a reference in later theoretical works [50, 51]. In this work we use the values obtained by Liebermann [36] as reference for the elastic tensor constants.

For the Vicker's hardness, we used the average of the results obtained by Bhaumik [53], of 8.58 GPa (at 100 g load, based on literature standards [54]), and the result of 8.2 GPa from Chicot *et al.* [55], resulting in an average value of 8.39 GPa.

B. Magnetite

Consensus data for magnetite is shown in the Tab. S13.

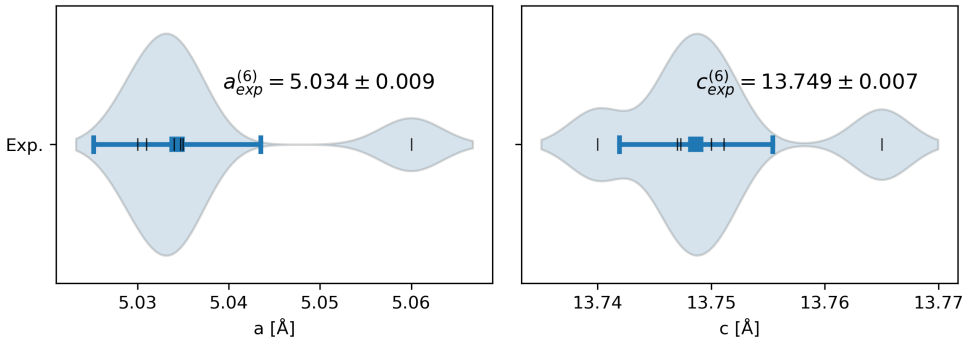

Fig. S2. Experimental lattice constants violin plot distributions for α-Fe_2O_3. The blue squares indicate the median estimation with two standard deviations error bars. The number in parenthesis in the labels reports the number of measurements.

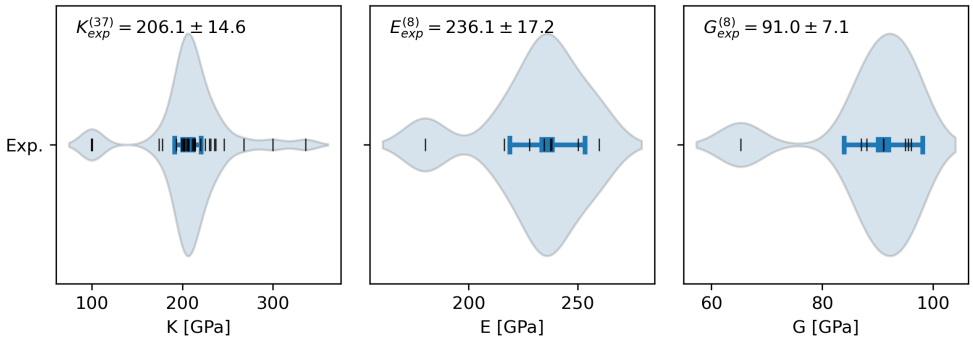

Fig. S3. Experimental elastic moduli (bulk, Young and shear modulus, K, E, and G, respectively) violin plot distributions for α-Fe_2O_3. The blue squares indicate the median estimation with two standard deviations error bars. The number in parenthesis in the labels reports the number of measurements.

Ref.	c_{11}	c_{33}	c_{44}	c_{12}	c_{13}	c_{14}
Voigt (1928) [30]*	247	232	87	56	16	-13
Birch (1966) [49]*	352.0	318.0				
Liebermann (1986) [36]*[a]	355.9	299.2	58.0	185.0	209.4	-10.4
Tomar (2004) [50]†[b]	273.54	184.05	106.6	56.8	16.52	-12.6
Tomar (2006) [51]†[b]	238.2	223.08	78.05	50.75	28.75	-13.12
Guo (2011) [52]†	319.6	294.4	80.1	125.6	104.5	5.5
Zhang (2020) [28]†	337.56	280.26	72.11	169.13	142.76	5.71

Table S12. Elastic tensor constants, C_{ij} [GPa], for hematite (Voigt notation). Symmetries: $c_{11} = c_{22}$, $c_{44} = c_{55}$, $c_{13} = c_{23}$, $c_{14} = c_{56} = -c_{24}$, $c_{66} = \frac{1}{2}(c_{11} - c_{12})$. *) Experimental results. †) Theoretical results. [a]) Based in magnetic field extrapolation $H \to 0$, except for c_{11} and c_{33}. [b]) Fitted to Voigt's data [30].

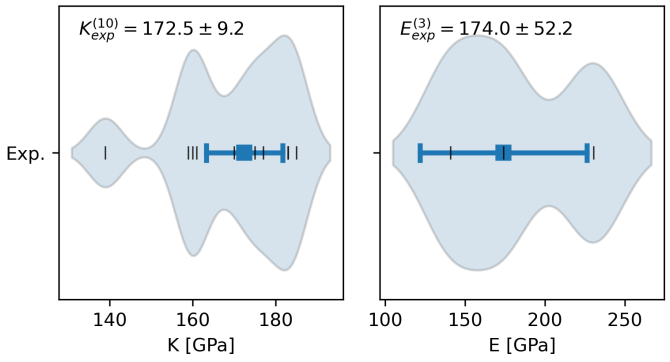

Fig. S4. Experimental elastic moduli (bulk and Young, K and E, respectively) violin plot distributions for magnetite. The blue squares indicate the median estimation with two standard deviations error bars. The number in parenthesis in the labels reports the number of measurements.

4. RAMAN AND INFRARED ACTIVE MODES' FREQUENCIES

Ref.	Conditions	a	c	K	K'	E	G	ν
Kataoka (1984) [56]						180		0.35
Smasonov (1973) [57]	100 GPa, 273 K			181.8[a]				
Straffelini (2002) [58]	32-43°C					210		0.3
Seo (2001) [59]	Room					174		
Asada (1983) [60]						141		0.33
Hearmon (1984) [61]						230.33	91.3	
Mao (1974) [62]				185				
				175				
				183				
				159				
				161				
				170				
				160				
				177				
				139				
				183				
Reichmann (2004) [63]	Ambient	8.39639(14)		180(1)	5.2(4)			
Lin (2014) [64]				186(3)	4′		66.3(1.8)	

Table S13. Experimental lattice constants and elastic moduli (units for length and pressure: Å, GPa) for hematite. [a]) Calculated by the inverse of the compressibility. Values indicated with apostrophes (′) were considered fixed by the authors.

Ref.	c_{11}	c_{44}	c_{12}
[65]	272.5	97.1	106.0
[63]	260.5(1.0)	63.3(1.5)	148.3(3.0)
[64]	240.6(4.9)	68.3(1.6)	113.8(5.9)

Table S14. Elastic tensor constants, c_{ij} [GPa], for magnetite (Voigt notation).

Fe-MLIP			Clay			Reax			Tersoff			IFF			Exp.	
n	Sym.	ω	n	Sym.	ω	n	Sym.	ω	n	Sym.	ω	n	Sym.	ω	Sym.	ω_{avg}
7	Eg	219.4	7	A1g	256.5	5	Eg	317.3	7	Eg	172.1	7	A1g	302.7	A1g	226.3 ± 2.5
9	A1g	221.8	8	Eg	272.7	12	Eg	403.5	9	Eg	215.8	8	Eg	324.7	Eg	245.7 ± 3.1
11	Eg	267.9	10	Eg	281.1	15	A1g	432.2	11	A1g	227.1	10	Eg	332.7	Eg	292.7 ± 2.5
15	Eg	286.9	14	Eg	325.0	16	Eg	446.8	16	Eg	433.1	14	Eg	375.1	Eg	299.0 ± 2.6
19	Eg	377.9	18	Eg	437.4	18	Eg	476.3	21	Eg	492.6	18	Eg	523.0	Eg	411.7 ± 3.2
26	A1g	481.4	26	A1g	612.6	25	Eg	587.0	26	Eg	628.4	26	A1g	712.8	A1g	498.7 ± 2.3
28	Eg	548.9	28	Eg	691.5	27	A1g	606.2	29	A1g	732.8	28	Eg	800.6	Eg	612.0 ± 3.0

Table S15. Calculated Raman active modes: number of the mode (n), irreducible representation symmetry, and frequency (ω [cm^{-1}]) for active Raman modes. Experimental average (ω_{avg}) taken from the data of Refs. [40, 66, 67]

Fe-MLIP			Clay			Reax			Tersoff			IFF			Exp.	
n	Sym.	ω	n	Sym.	ω	n	Sym.	ω	n	Sym.	ω	n	Sym.	ω	Sym.	ω_{avg}
5	Eu	212.2	5	Eu	236.3	10	A2u	411.5	5	Eu	169.7	5	Eu	292.3	Eu	224.0 ± 7.0
10	A2u	233.1	12	A2u	290.6	10	Eu	390.5	14	Eu	365.4	12	A2u	350.1	Eu	292.7 ± 7.0
13	Eu	286.4	16	Eu	356.0	8	Eu	345.9	13	A2u	338.1	16	Eu	426.7	A2u	300.0 ± 1.4
21	Eu	385.6	21	Eu	492.6	20	Eu	486.5	18	Eu	453.5	21	Eu	594.2	Eu	436.5 ± 2.9
23	A2u	432.0	23	A2u	517.4	23	A2u	509.2	23	A2u	519.8	23	A2u	618.3	Eu	523.0 ± 4.1
24	Eu	456.3	24	Eu	585.8	28	Eu	618.1	24	Eu	621.8	24	Eu	673.6	A2u	525.7 ± 0.6

Table S16. Calculated IR active modes: number of the mode (n), irreducible representation symmetry, and frequency (ω [cm^{-1}]) for active IR modes. Experimental average (ω_{avg}) taken from the data of Refs. [66, 68–70]

5. TRAINING

The configurations that comprised our dataset, used to train our potential, were sampled by molecular dynamics simulations with temperatures ranging from 100 to 2000 K and pressures from 1 to 100 atm. We also generated configurations by applying strains of $\pm 1\%$ and $\pm 3\%$ in each direction (x, y, and z), and combined directions (xy, yz, xz, and xyz) concomitantly with random displacements with standard deviations of 0.01, 0.02, and 0.03 Å. After sampling this first set of configurations, we trained a preliminary GNN potential from which we calculated the thermal expansion coefficient in the quasi-harmonic approximation, and random thermal displacements for the atoms at temperatures from 300 to 2100 K (with volumes corresponding to each given temperature) using the phonopy code [71]. Figure S5 shows the energy-pressure diagram of the configurations of our data.

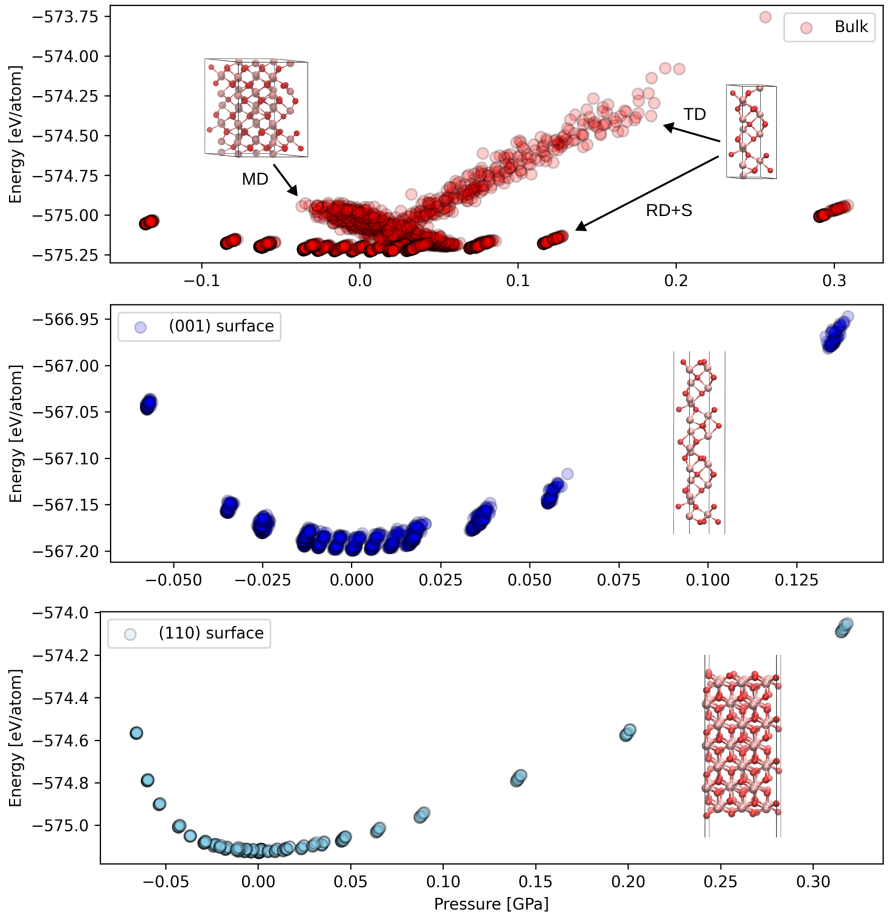

Fig. S5. Energy vs. pressure diagrams of the data sets for bulk configurations (top), (001) surface configurations (middle), and (110) surface configurations (bottom). The pictures show the supercell geometries of the configurations. The arrows indicate configurations generated by molecular dynamics (MD), random displacements and strains (RD+S), and random thermal displacements (TD).

We trained 4 different message-passing graph neural networks (MPGNN) with different initial random weights, in order to use the models in a committee by averaging them. The node states of the GNNs were represented by 32 channels up to $L = 1$ equivariant features. Two message passing interactions were used, with correlation order $\eta_c = 3$. A hybrid scheme that smoothly interpolates the Ziegler-Biersack-Littmark (ZBL) screened nuclear repulsion potential with the GNN models was employed to avoid divergences in predicted forces and energies if atoms get too close to one another [72]. A Huber loss function was used for energies, forces, and stresses in a two-phase optimization. First, we applied the exponential moving average (EMA) weight decay

with weight $\eta_{EMA} = 0.95$, and loss function component weights of $w_E^{EMA} = 1$, $w_F^{EMA} = 100$, and $w_\sigma^{EMA} = 1000$ for energies, forces and stresses, respectively. Then, we switch to stochastic weight averaging (SWA) with loss function component weights of $w_E^{SWA} = 100$, $w_F^{SWA} = 1$, and $w_\sigma^{SWA} = 100$ for another maximum 150 epochs with a learning rate of 10^{-3}, using early stopping if the models do not improve for 20 epochs. The schedulefree optimizer [73] was used during optimization.

Figure S6 shows the root mean square errors for the validation data set during training of energies, forces, and stresses. The vertical lines indicate the switching from EMA to SWA at epoch 150, and the horizontal lines are typical accuracy targets for interatomic potentials, 1 meV/atom for forces, 40 meV/Å for forces and 0.6 meV/Å^3 (̃0.1 GPa) for stresses. All models are well converged.

Figures S7-S10 show the parity graphs and the error distributions for the 4 individual trained Fe-MLIP models. All of the models display good performance on the test set, with RMS errors lesser than 1.2 meV/atom for energies, 32 meV/Å for forces, and 0.72 meV/Å^3 for stresses.

The four different models shown above can be averaged in a committee. This procedure has two main advantages: first, there is some error cancellation between the models, improving robustness and generalization; and secondly, the dispersion of the models predictions (the standard deviation, for instance) can be used as a measure of the adequacy of the potential for a given configuration, which allows the assessment of the quality of the potential for different thermodynamic conditions, for example [74]. Figure S11 shows the parity graphs and the error distributions for the committee Fe-MLIP models. The RMS errors of the committee for energies, forces and stresses are 0.75 meV/atom, 21 meV/Å, and 0.33 meV/Å^3, respectively.

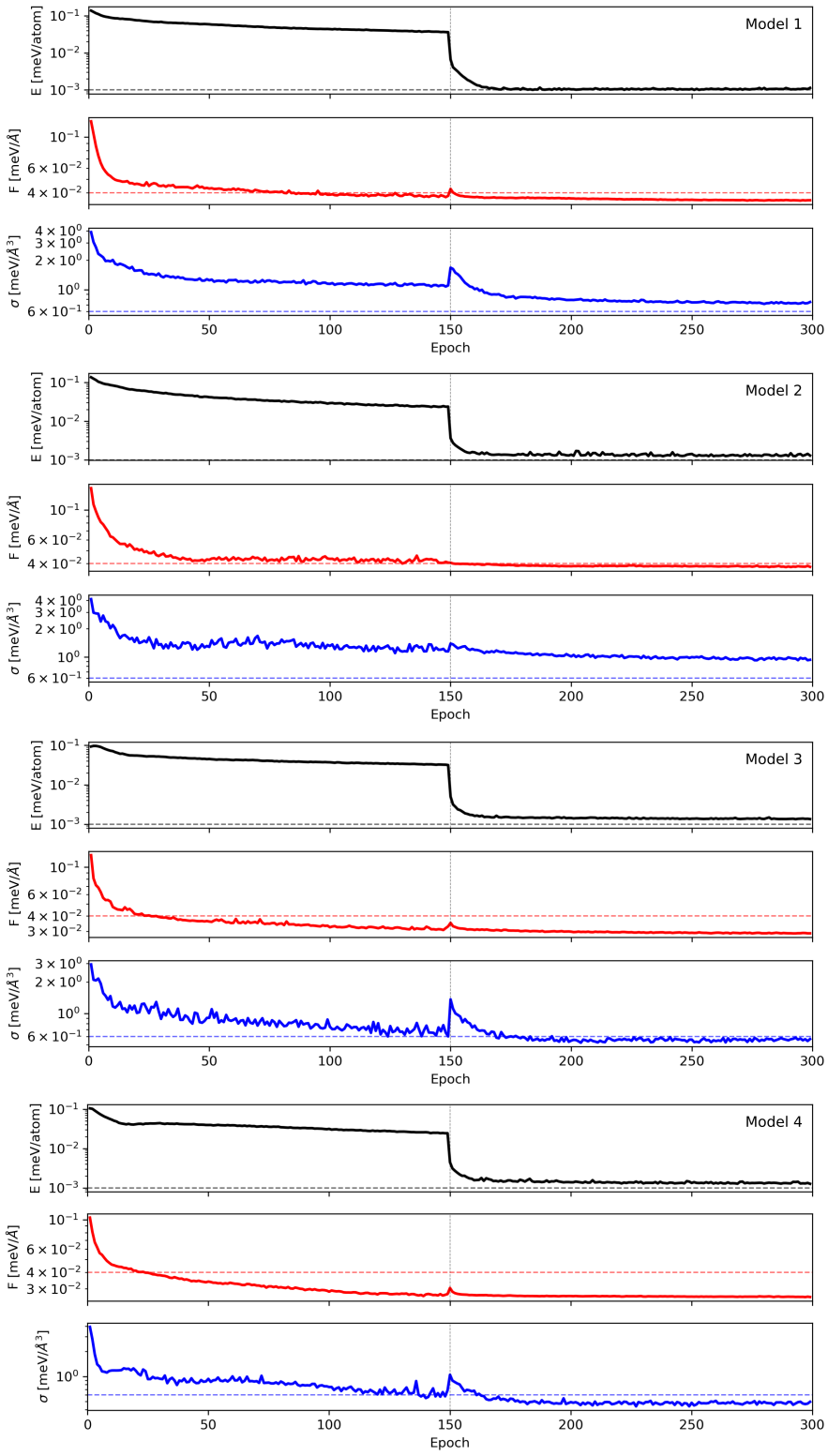

Fig. S6. Root mean square errors for the validation data set of energies E, forces F, and stresses σ, The vertical traced line indicates the switch from EMA to SWA (see text).

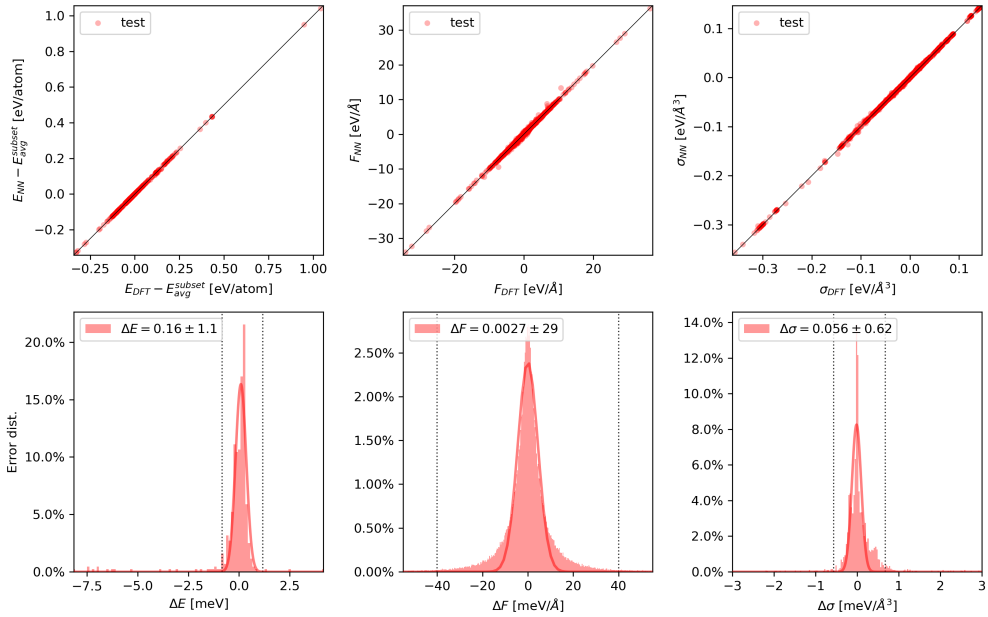

Fig. S7. Model 1 Fe-MLIP potential. Parity graph (top) and error distributions (bottom) for energies (left), forces (center), and stresses (right) calculated on the test set. The dotted lines indicate the ranges ± 1 meV/atom, ± 1 meV/Å, and ± 0.624 meV/Å^3 (0.1 GPa).

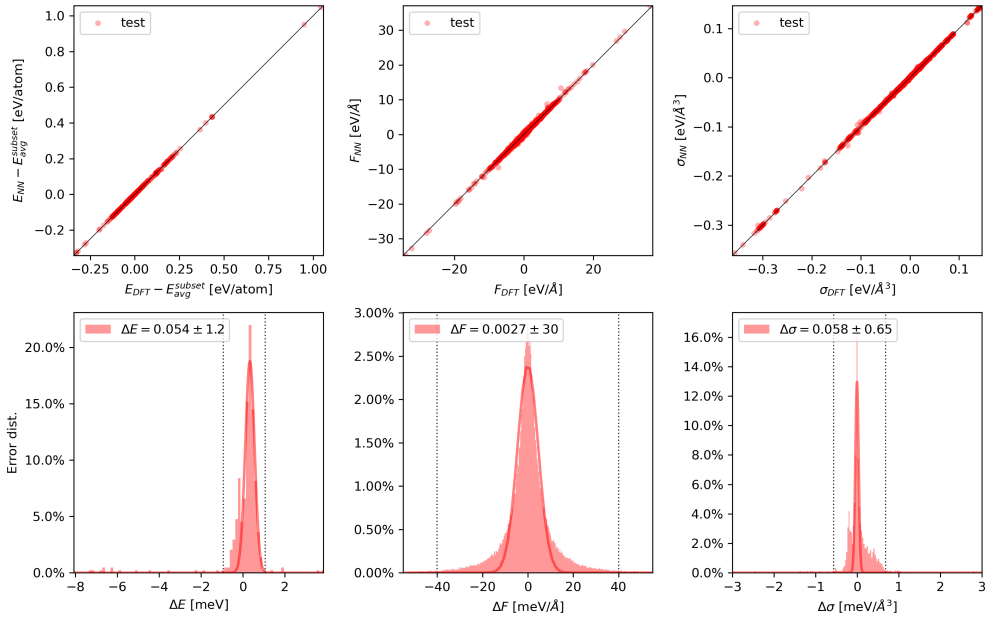

Fig. S8. Model 2 Fe-MLIP potential. Parity graph (top) and error distributions (bottom) for energies (left), forces (center), and stresses (right) calculated on the test set. The dotted lines indicate the ranges ± 1 meV/atom, ± 1 meV/Å, and ± 0.624 meV/Å^3 (0.1 GPa).

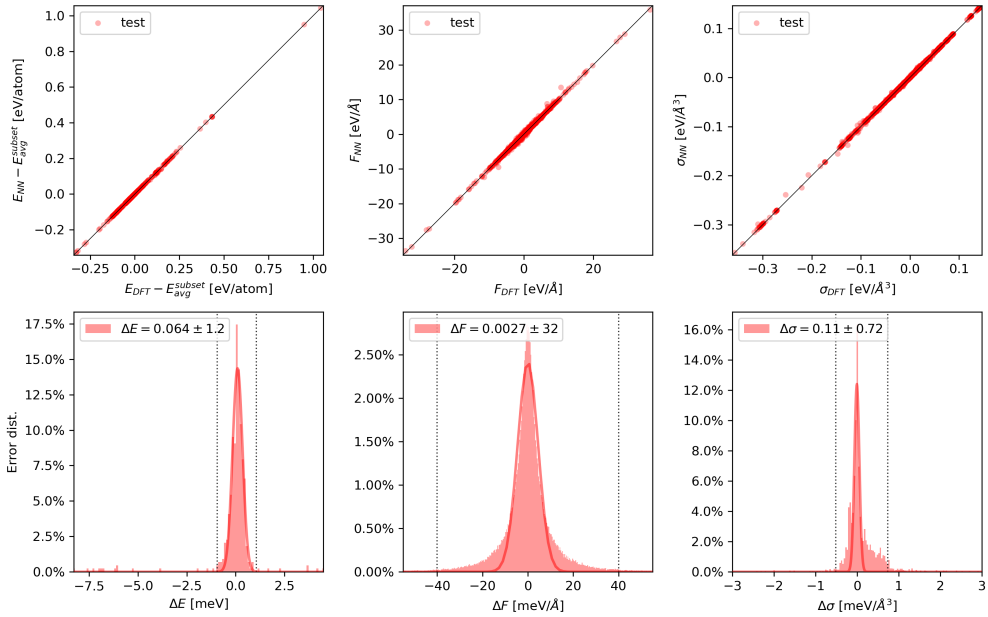

Fig. S9. Model 3 Fe-MLIP potential. Parity graph (top) and error distributions (bottom) for energies (left), forces (center), and stresses (right) calculated on the test set. The dotted lines indicate the ranges ± 1 meV/atom, ± 1 meV/Å, and ± 0.624 meV/Å^3 (0.1 GPa).

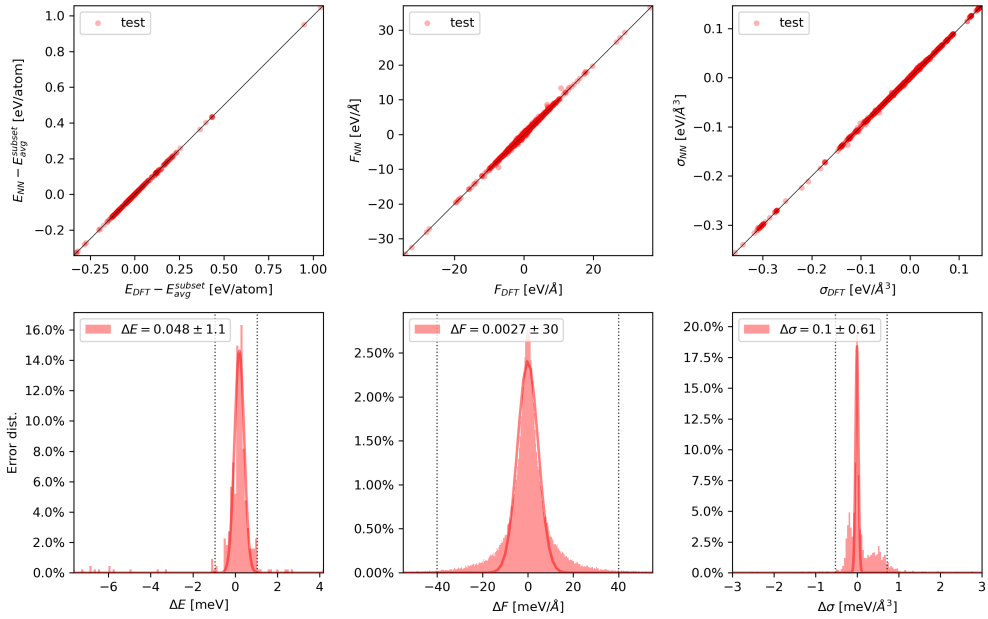

Fig. S10. Model 4 Fe-MLIP potential. Parity graph (top) and error distributions (bottom) for energies (left), forces (center), and stresses (right) calculated on the test set. The dotted lines indicate the ranges ± 1 meV/atom, ± 1 meV/Å, and ± 0.624 meV/Å^3 (0.1 GPa).

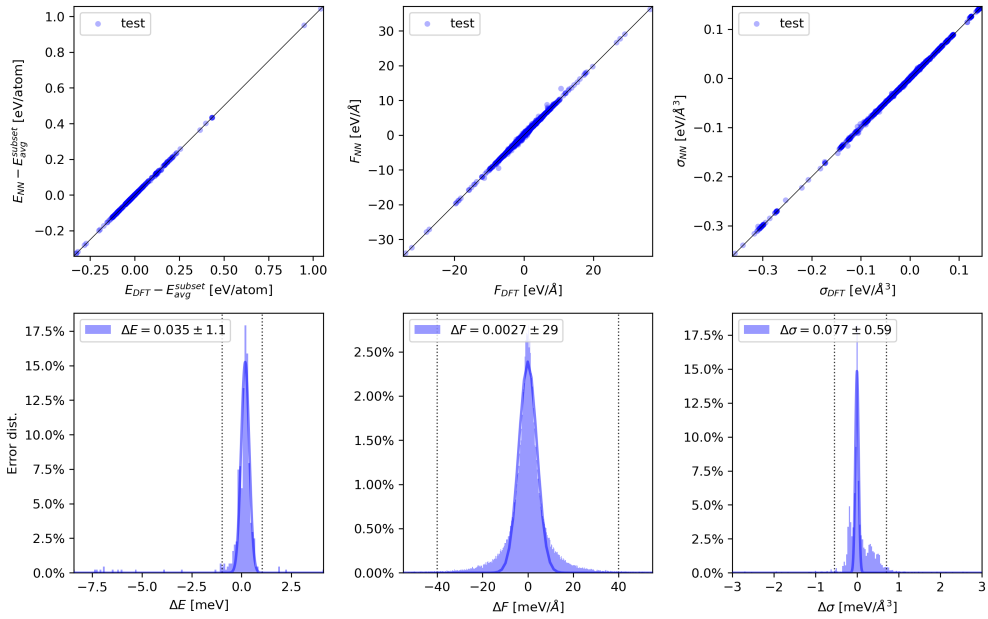

Fig. S11. Committee (4 models) Fe-MLIP potential. Parity graph (top) and error distributions (bottom) for energies (left), forces (center), and stresses (right) calculated on the test set. The dotted lines indicate the ranges ± 1 meV/atom, ± 1 meV/Å, and ± 0.624 meV/Å^3 (0.1 GPa)

.

REFERENCES

1. H. Heinz, T.-J. Lin, R. Kishore Mishra, and F. S. Emami, "Thermodynamically consistent force fields for the assembly of inorganic, organic, and biological nanostructures: The interface force field," Langmuir **29**, 1754–1765 (2013).
2. J. Huang and A. D. MacKerell, Jr., "CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data," J. Comput. Chem. **34**, 2135–2145 (2013).
3. K. Kanhaiya, M. Nathanson, P. J. in 't Veld, *et al.*, "Accurate force fields for atomistic simulations of oxides, hydroxides, and organic hybrid materials up to the micrometer scale," J. Chem. Theory Comput. **19**, 8293–8322 (2023).
4. R. T. Cygan, J.-J. Liang, and A. G. Kalinichev, "Molecular models of hydroxide, oxyhydroxide, and clay phases and the development of a general force field," The J. Phys. Chem. B **108**, 1255–1266 (2004).
5. R. T. Cygan, J. A. Greathouse, and A. G. Kalinichev, "Advances in clayff molecular simulation of layered and nanoporous materials and their aqueous interfaces," The J. Phys. Chem. C **125**, 17573–17589 (2021).
6. G. C. Abell, "Empirical chemical pseudopotential theory of cohesion for covalent and ionic solids," Phys. Rev. B **31**, 6184–6196 (1985).
7. J. Tersoff, "New empirical approach for the structure and energy of covalent systems," Phys. Rev. Lett. **56**, 632–635 (1986).
8. J. Tersoff, "New empirical approach for the structure and energy of covalent systems," Phys. Rev. B **37**, 6991–7000 (1988).
9. J. Tersoff, "Empirical interatomic potential for silicon with improved elastic properties," Phys. Rev. B **38**, 9902–9905 (1988).
10. J. Tersoff, "Empirical interatomic potential for carbon, with applications to amorphous carbon," Phys. Rev. Lett. **61**, 2879–2882 (1988).
11. D. W. Brenner, "Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films," Phys. Rev. B **42**, 9458–9471 (1990).
12. D. W. Brenner, "Relationship between the embedded-atom method and tersoff potentials," Phys. Rev. Lett. **63**, 1022 (1989).
13. M. S. Daw and M. I. Baskes, "Embedded-atom method: Derivation and application to

impurities, surfaces, and other defects in metals," Phys. Rev. B **29**, 6443–6453 (1984).
14. S. M. Foiles, M. I. Baskes, and M. S. Daw, "Embedded-atom-method functions for the fcc metals cu, ag, au, ni, pd, pt, and their alloys," Phys. Rev. B **33**, 7983–7991 (1986).
15. J. Byggmästar, M. Nagel, K. Albe, *et al.*, "Analytical interatomic bond-order potential for simulations of oxygen defects in iron," J. Physics: Condens. Matter **31**, 215401 (2019).
16. M. Müller, P. Erhart, and K. Albe, "Analytic bond-order potential for bcc and fcc iron—comparison with established embedded-atom method potentials," J. Physics: Condens. Matter **19**, 326220 (2007).
17. C. Björkas and K. Nordlund, "A comparison of the new fe–c potential to other potentials for simulation of the fe–c system," Nucl. Instruments Methods Phys. Res. Sect. B: Beam Interactions with Mater. Atoms **259**, 853–860 (2007).
18. P. Erhart, N. Juslin, O. Goy, *et al.*, "Analytic bond-order potential for atomistic simulations of zinc oxide," J. Physics: Condens. Matter **18**, 6585–6605 (2006).
19. A. C. T. van Duin, S. Dasgupta, F. Lorant, and I. Goddard, William A., "Reaxff: A reactive force field for hydrocarbons," The J. Phys. Chem. A **105**, 9396–9409 (2001).
20. Y. K. Shin, H. Kwak, A. V. Vasenkov, *et al.*, "Development of a reaxff reactive force field for fe/cr/o/s and application to oxidation of butane over a pyrite-covered cr2o3 catalyst," ACS Catal. **5**, 7226–7236 (2015).
21. M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, Oxford, UK, 1954).
22. B. G. Dick and A. W. Overhauser, "Theory of dielectric constants of alkali halide crystals," Phys. Rev. **112**, 90–103 (1958).
23. D. J. Cooke, S. E. Redfern, and S. C. Parker, "Atomistic simulation of the structure and segregation to the (0001) and (0112) surfaces of fe2o3," Phys. Chem. Miner. **31**, 507–517 (2004).
24. J. F. Nye, *Physical Properties of Crystals: Their Representation by Tensors and Matrices* (Oxford University Press, 1985).
25. W. Voigt, *Lehrbuch der Kristallphysik (mit Ausschluss der Kristalloptik)* (Teubner, Leipzig, 1910).
26. R. Hill, "The elastic behaviour of a crystalline aggregate," Proc. Phys. Soc. A **65** (1952).
27. A. Reuss, "Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle," Zeitschrift für Angewandte Math. und Mech. **9**, 49–58 (1929).
28. C. Zhang, L. Li, Z. Yuan, *et al.*, "Mechanical properties of siderite and hematite from dft calculation," Miner. Eng. **146**, 106107 (2020).
29. R. C. Liebermann and K. A. Maasch, "Acoustic and static compression experiments on the elastic behavior of hematite," J. Geophys. Res. **91**, 4651–4656 (1986).
30. W. Voigt, *Lehrbuch der Kristallphysik (mit Ausschluss der Kristalloptik)* (Teubner, 1928), p. 754, B. G. Teubner's Sammlung von Lehrbüchern auf dem Gebiete der mathematischen Wissenschaften mit Einschluß ihrer Anwendungen.
31. R. C. Liebermann and E. Schreiber, "Elastic constants of polycrystalline hematite as a function of pressure to 3 kilobars," J. Geophys. Res. (1896-1977) **73**, 6585–6590 (1968).
32. J. A. Eaton and A. H. Morrish, "Magnetic Domains in Hematite At and Above the Morin Transition," J. Appl. Phys. **40**, 3180–3185 (1969).
33. Y. Sato and S. Akimoto, "Hydrostatic compression of four corundum-type compounds: α-Al2O3, V2O3, Cr2O3, and α-Fe2O3," J. Appl. Phys. **50**, 5285–5291 (1979).
34. D. R. Wilburn, W. A. Bassett, Y. Sato, and S. Akimoto, "X ray diffraction compression studies of hematite under hydrostatic, isothermal conditions," J. Geophys. Res. Solid Earth **83**, 3509–3512 (1978).
35. L. W. Finger and R. M. Hazen, "Crystal structure and isothermal compression of Fe2O3, Cr2O3, and V2O3 to 50 kbars," J. Appl. Phys. **51**, 5362–5367 (1980).
36. R. C. Liebermann and K. A. Maasch, "Acoustic and static compression experiments on the elastic behavior of hematite," J. Geophys. Res. Solid Earth **91**, 4651–4656 (1986).
37. A. I. of Physics, *Soviet Physics: Technical physics*, no. v. 34,p. 232 in Soviet Physics: Technical Physics (American Institute of Physics., 1987).
38. J. S. Olsen, C. S. G. Cousins, L. Gerward, *et al.*, "A study of the crystal structure of fe2o3 in the pressure range up to 65 gpa using synchrotron radiation," Phys. Scripta **43**, 327 (1991).
39. J. Z. Jiang, J. S. Olsen, L. Gerward, and S. Mørup, "Enhanced bulk modulus and reduced transition pressure in γ-fe2o3 nanocrystals," Europhys. Lett. **44**, 620 (1998).
40. S.-H. Shim and T. S. Duffy, "Raman spectroscopy of fe2o3 to 62 gpa," Am. Mineral. **87**, 318–326 (2002).

41. H. Liu, W. Caldwell, L. Benedetti, *et al.*, "Static compression of α-fe 2 o 3: linear incompressibility of lattice parameters and high-pressure transformations," Phys. chemistry minerals **30**, 582–588 (2003).
42. S. M. Clark, S. G. Prilliman, C. K. Erdonmez, and A. P. Alivisatos, "Size dependence of the pressure-induced γ to α structural phase transition in iron oxide nanocrystals," Nanotechnology **16**, 2813 (2005).
43. G. Hearne and V. Pischedda, "Pressure response of vacancy ordered maghemite (γ-fe2o3) and high pressure transformed hematite (α-fe2o3)," J. Solid State Chem. **187**, 134–142 (2012).
44. E. Bykova, M. Bykov, V. Prakapenka, *et al.*, "Novel high pressure monoclinic Fe2O3 polymorph revealed by single-crystal synchrotron x-ray diffraction studies," High Press. Res. **33**, 534–545 (2013).
45. S. Klotz, T. Strässle, and T. Hansen, "Pressure dependence of morin transition in α-fe2o3 hematite," Europhys. Lett. **104**, 16001 (2013).
46. L. Oravova, Z. Zhang, N. Church, *et al.*, "Elastic and anelastic relaxations accompanying magnetic ordering and spin-flop transitions in hematite, fe2o3," J. Physics: Condens. Matter **25**, 116006 (2013).
47. P. Schouwink, L. Dubrovinsky, K. Glazyrin, *et al.*, "High-pressure structural behavior of α-fe2o3 studied by single-crystal x-ray diffraction and synchrotron radiation up to 25 gpa," Am. Mineral. **96**, 1781–1786 (2011).
48. Y. Zou, "Revealing pressure-induced anomaly in sound velocities, and new thermoelasticity of α-fe2o3 hematite at high pressure: Implications for the earth's interior," ESS Open Arch. (2022).
49. F. Birch, "SECTION 7: COMPRESSIBILITY; ELASTIC CONSTANTS (See also Section 9)," in *Handbook of Physical Constants,* (Geological Society of America, 1966).
50. V. Tomar and M. Zhou, "A molecular dynamics simulation framework for an al+ fe2o3 reactive metal powder mixture," MRS Online Proc. Libr. (OPL) **821**, P3–27 (2004).
51. V. Tomar and M. Zhou, "Classical molecular-dynamics potential for the mechanical strength of nanocrystalline composite fcc $\mathrm{Al} + \alpha - \mathrm{fe_2o_3}$," Phys. Rev. B **73**, 174116 (2006).
52. H. Guo and A. S. Barnard, "Modeling the iron oxides and oxyhydroxides for the prediction of environmentally sensitive phase transformations," Phys. Rev. B **83**, 094112 (2011).
53. N. Bhaumik, "Vickers micro hardness of magnetite, maghemite and hematite of a titaniferous vein magnetite of bonai, orissa, and its crypto-textural implications," J. Geol. Soc. India **3**, 157–163 (1962).
54. J. R. Craig and D. J. Vaughan, *Ore microscopy and ore petrography* (John Wiley & Sons, New York, 1994), 2nd ed.
55. D. Chicot, J. Mendoza, A. Zaoui, *et al.*, "Mechanical properties of magnetite (fe3o4), hematite (α-fe2o3) and goethite (α-feo.oh) by instrumented indentation and molecular dynamics analysis," Mater. Chem. Phys. **129**, 862–870 (2011).
56. K. Kataoka, T. Yamazawa, Y. J. Pyun, and T. Homma, "Sintering behavior of magnetite pellets with tio2 addition," Transactions Iron Steel Inst. Jpn. **24**, 365–373 (1984).
57. G. V. Samsonov, ed., *The Oxide Handbook* (IFI/Plenum, New York, 1973). Translated from Russian by C. Nigel Turton and Tatiana I. Turton.
58. G. Straffelini, G. Avi, and M. Pellizzari, "Dry sliding behaviour of a sintered fe–mo–cr–c alloy," Wear **252**, 870–877 (2002).
59. M. Seo and M. Chiba, "Nano-mechano-electrochemistry of passive metal surfaces," Electrochimica Acta **47**, 319–325 (2001).
60. M. Asada and Y. Omori, "Determination of young's modulus and poisson ratio of lump ores," J. Iron Steel Inst. Jpn. **69**, 739–746 (1983).
61. R. F. S. Hearmon, "The elastic constants of crystals and other anisotropic materials," in *Elastic, Piezoelectric, Pyroelectric, Piezooptic, Electrooptic Constants, and Nonlinear Dielectric Susceptibilities of Crystals*, vol. 18 of *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology - New Series, Group III* K.-H. Hellwege and A. M. Hellwege, eds. (Springer-Verlag, Berlin, Heidelberg, 1984), pp. 1–244.
62. H.-k. Mao, T. Takahashi, W. A. Bassett, *et al.*, "Isothermal compression of magnetite to 320 kbar and pressure-induced phase transformation," J. Geophys. Res. **79**, 1165–1170 (1974).
63. H. J. Reichmann and S. D. Jacobsen, "High-pressure elasticity of a natural magnetite crystal," Am. Mineral. **89**, 1061–1066 (2004).
64. J.-F. Lin, J. Wu, J. Zhu, *et al.*, "Abnormal elastic and vibrational behaviors of magnetite at high pressures," Sci. reports **4**, 6282 (2014).

65. M. S. Doraiswami, "Elastic constants of magnetite, pyrite and chromite," Proc. Indian Acad. Sci. - Sect. A **25**, 413–420 (1947).
66. A. M. Jubb and H. C. Allen, "Vibrational spectroscopic characterization of hematite, maghemite, and magnetite thin films produced by vapor deposition," ACS Appl. Mater. & Interfaces **2**, 2804–2812 (2010).
67. I. R. Beattie and T. R. Gilson, "The single-crystal raman spectra of nearly opaque materials. iron(iii) oxide and chromium(iii) oxide," J. Chem. Soc. A pp. 980–986 (1970).
68. T. D. Glotch, P. R. Christensen, and T. G. Sharp, "Fresnel modeling of hematite crystal surfaces and application to martian hematite spherules," Icarus **181**, 408–418 (2006).
69. S. Onari, T. Arai, and K. Kudo, "Infrared lattice vibrations and dielectric dispersion in $\alpha - \mathrm{fe_2o_3}$," Phys. Rev. B **16**, 1717–1721 (1977).
70. C. J. Serna, J. L. Rendon, and J. E. Iglesias, "Infrared surface modes in corundum-type microcrystalline oxides," Spectrochimica Acta Part A: Mol. Spectrosc. **38**, 797–802 (1982).
71. A. Togo, L. Chaput, T. Tadano, and I. Tanaka, "Implementation strategies in phonopy and phono3py," J. Phys. Condens. Matter **35**, 353001 (2023).
72. H. Wang, X. Guo, L. Zhang, *et al.*, "Deep learning inter-atomic potential model for accurate irradiation damage simulations," Appl. Phys. Lett. **114**, 244101 (2019).
73. A. Defazio, X. A. Yang, H. Mehta, *et al.*, "The road less scheduled," (2024).
74. J. Carrete, H. Montes-Campos, R. Wanzenböck, *et al.*, "Deep ensembles vs committees for uncertainty estimation in neural-network force fields: Comparison and application to active learning," The J. Chem. Phys. **158**, 204801 (2023).