This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2016-5629

REVISION 1

Title: A preliminary valence-multipole potential energy model: Al-Si-H-O system

Matthew C. F. Wander and Barry R. Bickmore

Department of Geological Sciences, Brigham Young University, Provo, UT 84602, U.S.A. E-mail: mcfwander@gmail.com

Abstract

Here we test the concept that a potential energy model (force field) based on an expansion of the bond-valence model can use molecular geometry to make a reasonable prediction of the thermodynamic energy. The backbone of the model is a non-standard choice of structural descriptors for the energy decomposition, which relates the energy to particular aspects of the structure. Most force fields use a many-body decomposition to describe structures (with two-, three-, and possibly four-body terms, etc.), whereas ours employs a multipole expansion of the bond valence incident to each atom. This valence multipole model separates the energy associated with each atom into terms related to total bonding (valence monopole), bonding asymmetry (valence dipole), and ellipsoidal deformation (valence quadrupole). All of these are inherently multi-body terms that are calculated by combining two-body terms (bond valences). Provided bond valence sums are satisfied to within 0.2 v.u. of the ideal for all atoms, this model can provide accuracies of ~5 kJ/mol per unique atom in the Al-Si-H-O system, at least for the equilibrium structures tested here, comparable to most quantum mechanical calculations. More development is needed to produce a fully functional force field suitable for molecular dynamics

simulations, but this work shows that the development of such a force field is likely to be feasible.

Introduction

Molecular modeling has become increasingly popular in geochemistry and mineralogy, as knowledge of molecular-scale mechanisms and structures has become progressively more important. Quantum mechanical (QM) models are the gold standard for molecular modeling, but can be prohibitively computationally expensive, especially for the larger systems of atoms sometimes needed to adequately represent natural materials and processes. Molecular mechanics (MM) potential energy models, or "force fields," are common, less computationally intensive alternatives, but are subject to a number of difficulties that have often limited their scope and effectiveness.

We suggest that some of these problems can be addressed by restructuring the basic architecture of an MM force field. In this contribution, we report on the construction, optimization, and initial testing of a potential energy model (intended to be a precursor to a fully reactive force field) for the Al-Si-H-O system, designed around the Valence Multipole Model (VMM) (Bickmore et al., 2013; Shepherd et al., 2016), which is an extension of the Bond-Valence Model (BVM) (Brown, 2002; 2009; 2014). To our knowledge, this is the first example of a potential energy model constructed completely from a bond-valence model of chemistry rather than by using the BVM to augment existing architectures. Early indications are that such models are capable of excellent accuracy, with limited computational expense, at least with respect to estimating thermodynamic energies of equilibrium structures. Considering how often bond valence has been used in quantitative structure-activity relationships (Hiemstra et al., 1989; Sverjensky, 1994; Hiemstra and Van

Riemsdijk, 1996; Hiemstra et al., 1996; Sverjensky and Sahai, 1996; Sahai and Sverjensky, 1997b; a; Sverjensky et al., 1997; Lufaso and Woodward, 2001; Sahai, 2002; Bickmore et al., 2004; Etxebarria et al., 2005; Sverjensky, 2005; Bickmore et al., 2006a; Bickmore et al., 2006b; Perez-Mato et al., 2009; Bickmore, 2014), this is a significant development in itself.

Theory

To introduce this new approach, we discuss the standard architecture of MM force fields, give a basic explanation of the VMM, and show how a potential energy model based on the VMM would differ from the standard architecture, while allowing for broader applicability and greater accuracy, at a reasonable computational cost.

Standard MM Force Fields

MM force fields mimic interactions between atoms and molecules by treating them essentially as "balls on springs" (Hinchliffe, 2003). That is, the geometry of a system of atoms is defined in terms of a set of structural descriptors (e.g., interatomic distances and bond angles), ideal values are specified for the descriptors (e.g., preferred bond lengths and angles), and energy cost functions are applied to deviations from the ideal values. The simplest and most often used of these energy cost functions is Hooke's Law for springs (Eqn. 1), in which the spring length is the structural descriptor of interest. Here, u is the potential energy, k is a constant, x_0 is the ideal value of the structural descriptor of interest, and x is the actual value.

$$u = \frac{1}{2}k(x - x_0)^2 \quad (1)$$

Each of these energy terms includes one or more parameters that can be adjusted to fit the overall model to some set of data, including crystal or molecular structures, thermodynamic data, or even structures and associated energies gleaned from QM calculations (Rappé and Casewit, 1997; Hinchliffe, 2003; Cramer, 2004; Comba et al., 2009). Other common energy cost functions have quartic or sinusoidal forms (Cramer, 2004).

One of the core assumptions of a standard MM force field is "pair-wise additivity" (Rowley, 1994), that every pair-wise energy contribution is completely independent of every other pair wise energy contribution. The geometry of the system is described primarily via pair-wise potential energy terms, u_{ij}, based on the distance between a pair of atoms (e.g., Coulomb, Van der Waals, and Morse potentials). The total potential energy of the system (U) is approximated simply by summing the pair-wise terms, as in Eqn. 2.

$$U = \sum_{i<j} u_{ij} \qquad (2)$$

This assumption is not strictly true, however. For instance, changing one bond's length (for whatever reason) changes the ideal lengths of neighboring bonds linked to the same atoms, as well as the angles between them. This fact follows from QM theory, and explains why structures optimized via MM tend to have higher symmetry than the corresponding experimental structures (Comba et al., 2009).

Such a simple approach generally only poorly describes the geometry of the system and it is frequently necessary to add multi-body terms (e.g., bond-angle bending terms involving three bodies, or out-of-plane or dihedral-angle distortion terms involving four

bodies) as corrections to the lower-order terms, decreasing in importance with the number of bodies involved. This is called many-body decomposition; it means simply that the energy terms are organized by increasing number of bodies (atoms) involved. This can be expressed as an expansion about the pair-wise energies, as in Eqn. 3, where *N* is the total number of atoms or molecules in the system (Rowley, 1994).

$$U = \sum_{i<j} u_{ij} + \sum_{i<j<k} \Delta u_{ijk} + \cdots + \sum \Delta u_{1,2,3\ldots N} \quad (3)$$

Mathematically, this is an excellent choice, because it is a limiting series in which the two-body terms will be most important, and subsequent terms decrease in importance as the number of bodies increases. Chemically, however, these terms are coupled with one another, and this coupling increases for the higher-order terms. As the number of interactions considered increases, so does the number and degree of highly coupled, adjustable model parameters. While this can produce high accuracy results for a particular set of chemical conditions, this can lead to force fields that can be over-fitted to that limited domain, and are woefully inaccurate outside that domain. Furthermore, the computational expense of tracking spatial relationships between increasing numbers of bodies can quickly become prohibitive.

Even with multi-body corrections, force fields based on pair potentials with universal ideal bond lengths and angles for interactions of a given type may not be able to model chemical reactions during which coordination-number changes take place. This drove the need to create reactive force fields, potential energy surfaces for molecules that

were capable of bond dissociation and formation. The methods for creating reactive force fields that can address bond and coordination number changes usually involve extra potential energy terms built on a foundation of more traditional terms (Liang et al., 2013). Thus, the parameterization problem is exacerbated, and the range of utility of these types of models can be further limited.

Bond-Order Potentials and the Bond-Valence Model

One method used to augment traditional MM force fields is to use bond order as a structural descriptor. Bond order generally refers to the number of electron pairs involved in a given bond; *e.g.*, single bonds involve one electron pair, double bonds involve two, and non-integral bond orders are possible, as well. Particular atoms tend to be most stable when the bond orders of incident bonds sum to some ideal value, which is determined by standard electron counting rules, e.g., the octet rule. There are various methods for calculating bond order, based on bond lengths, bond angles, calculated electron densities, and so on, but the advantage of using bond order as a structural descriptor is that the energy cost comes from deviation of the summed bond orders about each atom from an ideal value; *i.e.*, it is based on the total bonding environment of an atom. Some examples of MM force fields that include bond-order terms, called bond-order potentials (BOPs), are the Tersoff (Tersoff, 1988), Brenner (Brenner, 1990), Finnis-Sinclair (Finnis and Sinclair, 1984), and ReaxFF (van Duin et al., 2003) potentials.

One class of BOPs (Grinberg et al., 2002; Cooper et al., 2003; Grinberg et al., 2004; Shin et al., 2005; Shin et al., 2007; Shin et al., 2008; Grinberg et al., 2009; Liu et al., 2013a;

Liu et al., 2013b; Takenaka et al., 2013) uses the BVM (Brown, 2002; 2009; 2014) to estimate bond orders. The BVM is essentially a method for relating bond order to bond length. (Note: From this point forward, the terms "bond order" and "bond valence" are used interchangeably.) The principal axiom of the BVM is the valence sum rule (Eqn. 4), which predicts that in a stable structure the sum (S_i) of the valences (orders) of the bonds (s_{ij}) incident to atom *i* from neighboring atoms *j* must cancel the atomic valence of atom *i* (V_i). In the traditional forms of the BVM, the bond valences are negative in the direction of cations and positive in the direction of anions.

$$\sum_j s_{ij} + V_i = S_i + V_i = 0 \tag{4}$$

The value of V_i is usually taken as equivalent to the oxidation number, but important variants exist (O'Keeffe and Brese, 1992; Brown, 2002; 2014; Wander et al., 2015a). The individual bond valences are calculated using an equation relating bond length to bond valence, which is calibrated on known equilibrium structures by applying some bond valence-length relationship and enforcing the valence sum rule. The most commonly used bond valence-length relationship is shown in Eqn. 5, where R_{ij} is the bond length, R_0 is the length of a single (1 valence unit or v.u.) bond, and B is a constant that dictates the curvature.

$$s_{ij} = e^{(R_0 - R_{ij})/B} \tag{5}$$

The individual bond valences are computationally comparable to simple pair-wise terms. But if a BOP applies an energy cost function like Eqn. 1 to the deviation of S_i from V_i,

153 rather than deviation of individual bond valences from ideal values, then we have a true
154 multi-body term that is calculated simply by summing pair-wise terms. This is why BOPs
155 are attractive, *i.e.*, they employ multi-body terms that are relatively computationally
156 inexpensive, and require only two-body parameterization.
157
158 Expanding the Bond-Valence Model
159 As mentioned above, however, the multi-body terms in BOPs have more or less been
160 treated as corrections to a model framework based on standard many-body terms that
161 assume pair-wise additivity, the independence of the individual bond potentials. Our goal
162 is to discover what it would take to create a potential energy model based, as completely as
163 possible, on the concept of bond valence. This requires some expansion of the BVM, which
164 we have addressed under four headings: 1) introducing directionality via the Valence
165 Multipole Model (VMM), 2) introducing fully covalent interactions, 3) redefining the shape
166 of bond valence-length curves, and 4) optimization of the curves to produce appropriate
167 potential energy surfaces.
168 *The Valence Multipole Model.* Bond valence alone addresses only bond lengths,
169 rather than the complete spatial distribution of bonds, and so is incapable of describing all
170 aspects of a structure. The VMM addresses this problem by describing full structures via a
171 multipole expansion of the valence bonds incident to each atom, made possible by treating
172 bond valence as a vector quantity (Harvey et al., 2006; Bickmore et al., 2013). In this way,
173 we can consider both scalar magnitudes of bonding, as well as directionality.
174 Multipole expansions are commonly used to describe spatial distributions of things
175 like electric charge or mass, using a succession of terms that describe progressively finer

details. Typical multipole terms include the monopole (a scalar describing the total quantity of interest within a defined region), the dipole moment (a vector describing the lopsidedness of the distribution around some point), the quadrupole moment (a rank-2 tensor describing the spherical symmetry of the distribution), and so on (octupole moment, etc.) In the VMM, the bond monopole (S_i) is the total bonding reaching an atom, the dipole moment is the vector sum of the bonds reaching that atom, and the quadrupole moment is a tensor describing the ellipsoidal distribution of bond valence. The bond dipole can describe non-centrosymmetric distortions of the coordination sphere, while the bond quadrupole moment can describe centrosymmetric distortions. (To be clear, we should emphasize the fact that bond valence does not equate to charge transfer, although this is a common misconception. If that were the case, then a multipole expansion of bond valence would essentially be the same as a multipole expansion of the charge distribution, which is more familiar to scientists. Rather, we are using the same mathematical formalism to describe the spatial distribution of "bonding" incident to an atom, which is physically a more nebulous concept. As explained below, we will likely add a Coulomb term to our model in the future.)

We find it convenient to normalize (Eqn. 6) the individual bond-valence vectors $\vec{s}_{ij}$ to S_i, resulting in fractional bond-valence vectors $\left(\vec{p}_{ij}\right)$ that are used to calculate the valence dipole $\left(\vec{P}_i\right)$ and valence quadrupole (Θ_i) terms for each atom, as in Eqns. 7-9. Here N is the number of bonds from neighboring atoms j incident to the central atom i, and p_{ij_α} or p_{ij_β} represents the magnitude of $\vec{p}_{ij}$ projected onto one of the Cartesian directions, where $\alpha, \beta = x, y, z$. (E.g., $\vec{p}_{ij_x} = \vec{x} \bullet \vec{p}_{ij}$, where $\vec{x}$ is the unit vector in the x direction.) $\delta_{\alpha\beta}$ is the Kronecker delta, which is one in the case that direction α equals direction β, but otherwise

zero. The $\vec{P}_i$ and Θ_i values calculated using fractional bond-valence vectors $(\vec{p}_{ij})$ thus depend only upon the shape of the bond-valence distribution, and not on the atomic valence of the central atom, or the over- or under-bondedness of the atom (Shepherd et al., 2016).

$$\vec{p}_{ij} = \begin{bmatrix} p_{ij_x} \\ p_{ij_y} \\ p_{ij_z} \end{bmatrix} = {\vec{s}_{ij}}/{S_i} \tag{6}$$

$$\vec{P}_i = \sum_j \vec{p}_{ij} \tag{7}$$

$$\boldsymbol{\Theta}_i = \frac{1}{2}\left[\sum_{j=1}^{N}\sum_{\alpha=x,y,z}\sum_{\beta=x,y,z}\left(3p_{ij_\alpha}p_{ij_\beta} - \delta_{\alpha\beta}\left\|\vec{p}_{ij}\right\|^2\right)\right] \tag{8}$$

$$\boldsymbol{\Theta}_i = \frac{1}{2}\begin{pmatrix} \sum_{j=1}^{\mathrm{N}}\left(3\left(p_{ij_x}\right)^2 - \left\|\vec{p}_{ij}\right\|^2\right) & 3\sum_{j=1}^{\mathrm{N}} p_{ij_x}p_{ij_y} & 3\sum_{j=1}^{\mathrm{N}} p_{ij_x}p_{ij_z} \\ 3\sum_{\mathrm{j}=1}^{\mathrm{N}} p_{ij_y}\,p_{ij_x} & \sum_{j=1}^{\mathrm{N}}\left(3\left(p_{ij_y}\right)^2 - \left\|\vec{p}_{ij}\right\|^2\right) & 3\sum_{j=1}^{\mathrm{N}} p_{ij_y}\,p_{ij_z} \\ 3\sum_{j=1}^{\mathrm{N}} p_{ij_z}p_{ij_x} & 3\sum_{j=1}^{\mathrm{N}} p_{ij_z}p_{ij_y} & \sum_{j=1}^{\mathrm{N}}\left(3\left(p_{ij_z}\right)^2 - \left\|\vec{p}_{ij}\right\|^2\right) \end{pmatrix} \tag{9}$$

Finally, we discard some of the directional information contained in $\vec{P}_i$ and $\boldsymbol{\Theta}_i$ to condense them into single scalar values, by calculating their magnitudes (norms). Eqn. 10 shows the equation for the norm of $\vec{P}_i$, and Eqn. 11 shows the equation for the Frobenius

norm of $\mathbf{\Theta}_i$, where θ_{mn} represents the nine elements of $\mathbf{\Theta}_i$ in rows *m* and columns *n*. The Frobenius norm does not depend on the rotation of the axes (Horn and Johnson, 2012; Golub, 2013).

$$\left\|\vec{P_i}\right\| = \sqrt{P_{i_x}^2 + P_{i_y}^2 + P_{i_z}^2} \tag{10}$$

$$\|\mathbf{\Theta}_i\| = \sqrt{\sum_{m=1}^{3}\sum_{n=1}^{3}|\theta_{mn}|^2} \tag{11}$$

As a result, we now have three scalar quantities (S_i, $\left\|\vec{P_i}\right\|$, and $\|\mathbf{\Theta}_i\|$), capable of describing different aspects of the total bonding environment of each atom in a structure. Previous work (Bickmore et al., 2013; Shepherd et al., 2016) has shown that $\left\|\vec{P_i}\right\|$ and $\|\mathbf{\Theta}_i\|$ vary systematically as a function of the bonding environment. Therefore, it should be possible to model how their ideal values change under various circumstances. However, it should be noted that it may be necessary, in some circumstances, to recover some of the directional information lost when taking the Frobenius norm of the quadrupole. For instance, $\|\mathbf{\Theta}_i\|$ does not distinguish between oblate and prolate spheroidal deformation.

Bond Valence for Fully Covalent Bonds. The standard formulation of the BVM effectively addresses ionic and polar-covalent, but not fully covalent (i.e., between atoms of the same element), bonds. The reason for this is that if atomic valence (V_i) is taken as equivalent to the oxidation number, fully covalent (*i.e.*, same-element) bonds must be ignored for calculating the total bond valence. For example, C in diamond has an oxidation

state of zero, but forms four single bonds with its neighbors. Likewise, C in ethane (C_2H_6) has an oxidation state of +3, but still forms four single bonds to the three surrounding H atoms, and the neighboring C. One must simply decide on the correct value of V_i for covalent systems. In both of the above cases, it is clearly 4, rather than 0 or 3, as the classic model suggests.

Fully covalent bonds can easily be included in a bond-valence treatment, however, if certain adjustments are made to the model. For instance, if we define atomic valence as the maximum number of single bonds incident to an atom (always a positive number), it would suffice for most typical situations involving minerals. In this case, we would also need to define the bond valence as positive in both directions, and recast the valence sum rule (Eqn. 4) as in Eqn. 12.

$$\sum_j s_{ij} = S_i \approx V_i \qquad (12)$$

O'Keeffe and Brese (1992) showed that a bond-valence approach could be used to describe anion-anion bonds in, for example, persulfides, sulfosalts, and peroxides, but ignored bonds weaker than 0.25 v.u. However, Wander et. al. (2015a) showed, for the Al-Si-H-O and Al-Si-K-O systems, that calculated bond-valence sums could actually be improved by including even very weak anion-anion bonds. Including cation-cation bonds made little difference in these systems, but certainly would in others, particularly in cases with partially and fully reduced metals.

The utility of including fully covalent bonding goes beyond improving bond-valence sums, however. It widens our definition of bonding, allowing for more types of two-body

interactions to be included in our bonds. First, it provides a means, within a strict bond-valence framework, of incorporating attractive and repulsive potentials between co-ions, depending on whether the atoms involved are under- or over-bonded, respectively. It also allows for the incorporation of Van der Waals interactions within the bond-valence framework. Second, if we acknowledge that the energy (*i.e.*, well depth) of a given bond depends on both the bond valence and factors such as the bond character, we can qualitatively predict a very broad range of chemical behavior (Wander et al., 2015a).

Bond Valence-Length Curve Shapes. To define bond valence-length relationships for individual atom pairs, one usually assumes a simple model form similar to Eqn. 5, and then adjusts one or two of the parameters to enforce the valence sum rule (Eqn. 4) within a collection of crystal structures. One potential problem with this is that a given atom pair typically exhibits a fairly narrow range of bond lengths in crystal structures, over which many simple decay functions can reasonably represent the relationship. Wander et al. (2015b) used both molecular and crystal structures, however, to show that over a wider range of bond lengths, simple one- or two-parameter decay functions are not always flexible enough to adequately capture the relationship. This failure is problematic, given that our goal is to produce a valence-based potential energy model for use in molecular dynamics simulations, which would require accurate representation of a wider range of bond valences.

Eqns. 13 and 14 show more flexible forms of the bond valence-length relationship, which we will use here in addition to Eqn. 5. Eqn. 13 is a weighted, geometric mean of the exponential decay function in Eqn. 5 and a power-law decay function. It has three

parameters: R_0, B, and w, which is a weighting coefficient with possible values between 0 and 1.

$$s_{ij} = e^{(R_0 - R_{ij})w/B} \left(\frac{R_0}{R_{ij}}\right)^{(1-w)/B} \quad (13)$$

Eqn. 14 is a weighted, arithmetic mean of two exponential decay functions, with five parameters, including distinct values of R_0 and B.

$$s_{ij} = we^{(R_{0,1} - R_{ij})/B_1} + (1-w)e^{(R_{0,2} - R_{ij})/B_2} \quad (14)$$

In Eqns. 5 and 13, R_0 corresponds to the length of the single bond, but this is not the case for Eqn 14. There is still one and only one R_{ij} value that does correspond to the single bond length, however. In all cases where we have applied Eqn. 14, the curvature break occurs at around s_{ij} = 1 v.u.

Bond Valence-Length Curve Optimization. Bond valence-length relationships can be estimated based on known structures for which the individual bond valences, or at least the total bond valence incident to each atom, can be unambiguously assigned. In some cases, however, the available structures are too narrowly focused to sufficiently constrain the curve shape. In other cases, a wide distribution of bond lengths is available, but it is difficult or impossible to unambiguously assign bond valence values. This is particularly problematic for optimizing B values in Eqns. 5, 13, and 14, because it can demonstrate enormous flexibility to compensate for poor structural data (*e.g.*, poorly characterized O-H bond lengths). In such cases, it may be possible to use supplementary methods.

We assume that bond valence is essentially identical to bond order, and use prior work (Badger, 1934; Johnston, 1966) on the relationship between vibrational force

295 constants and bond order to leverage available vibrational spectroscopic data. Johnston
296 (1966) posited the relationship in Eqn. 15, where k_{ij} is the vibrational force constant for a
297 bond between atoms *i* and *j*, of order s_{ij}, and k_1 is the vibrational force constant for a single
298 bond of the same type.

299 $$s_{ij} = \frac{k_{ij}}{k_1} \quad (15)$$

300 In cases for which we know the vibrational force constants for triple, double and single
301 bonds, they conform closely to a 3:2:1 ratio. Making this assumption that bond valence is
302 essentially equivalent to other bond order estimates allows us to use vibrational data from
303 a much broader array of molecules, including radicals, to constrain the shape of bond
304 valence-length curves. This is particularly important for estimating the *B* parameters with
305 confidence (Eqns. 5, 13, and 14), because we can now relate this value to the changes in the
306 experimental vibrational force constant with distance. Whether one uses the traditional
307 structural data, or the method we suggest here using experimental vibrational data, one
308 should get the same answer.
309
310 Potential Energy Model
311 Our expanded bonding model provides structural descriptors and rough estimates,
312 at least, of ideal values for those descriptors under various circumstances. To produce a
313 complete potential energy model, however, we must create energy cost functions for
314 deviations from the ideal values. Here we describe the basic form of our preliminary
315 model, along with the associated energy cost functions.
316 *Energy Decomposition.* The essential form of our model is shown in Eqn. 16, where *E*
317 denotes energy, and the subscripts VM, VD, and VQ denote the valence monopole, dipole,

and quadrupole terms for each atom *i*. (Again, these are calculated from bond vectors, rather than the distribution of electric charge.) These terms are then summed over all atoms in the system.

$$E_{Total} = \sum_i \left(E_{VM,i} + E_{VD,i} + E_{VQ,i}\right) \quad (16)$$

As model development progresses, we anticipate that we may need to add other terms, such as Coulomb interactions at long range, dipole-dipole terms on neighboring atoms, and pressure corrections. However, the simple model shown in Eqn. 16 is sufficient to show the utility of the valence-multipole energy decomposition.

Valence Monopole Energy. The valence monopole energy for each atom *i* (Eqn. 17) is a function of the total incident bond valence (S_i) and the atomic valence (V_i). It is scaled to the a quantity ($D_{E1,i}$) analogous to the well depth term in a Morse function, and includes an exponent (α_i). This form is valid only if $S_i \approx V_i$, but it is not known how close they need to be.

$$E_{VM,i} = \frac{1}{2} V_i D_{E1,i} \left(\left(\frac{S_i}{V_i} \right)^{\alpha_i} - 1 \right)^2 \quad (17)$$

As shown in Eqn. 18, the value of α_i is related to the softness parameters (B) in the bond valence-length relationships used (Eqns. 5, 13, 14), and both the force constant (k_1) and well depth ($D_{E1,i}$) for a single bond of the type incident to the atom.

$$\alpha_i = B_i \sqrt{\frac{k_{1,i}}{2D_{E1,i}}} \quad (18)$$

In all cases, the subscript *i* in Eqns. 16-18 denotes a particular atom, so terms that would normally be associated with a particular bond must somehow be averaged over all

338 the bonds incident to the atom in question. For each atom *i*, we use a simple weighted
339 arithmetic mean over all the incident bonds *ij*, to compute the $k_{1,i}$ and $D_{E1,i}$ (Eqns. 19-20).
340

$$k_{1,i} = \left(\frac{1}{S_i} \sum_j s_{ij} k_{1,ij} \right)$$

341 (19)

342 $$D_{E1,i} = \left(\frac{1}{S_i} \sum_j s_{ij} D_{E,ij} \right)$$

343 (20)

344 For the B_i value, we instead do a weighted arithmetic average of $1/B_{ij}$ (Eqn. 21).

$$\frac{1}{B_i} = \left(\frac{1}{S_i} \sum_j \frac{s_{ij}}{B_{ij}} \right)$$

345 (21)

346 However, this could not be applied directly in cases where the double exponential form
347 (Eqn. 14) was used for the bond valence-length relationship. In such cases, however, we
348 treat the two exponential terms as describing two separate bonds between the same two
349 atoms, in which case Eqn. 21 still applies.
350 What remains, at this point, is to describe how D_E values for individual bonds are
351 estimated. D_E values for bonds involving each atom pair are represented by a polynomial
352 function of s_{ij} (Eqn. 22), except in the case of H-O bonds, for which Eqn. 23 is applied. In
353 these equations, *a*, *b*, *c*, etc., denote fitted constants.

$$D_{E,ij} = a s_{ij} + b s_{ij}^2 + \cdots$$

354 (22)

$$D_{E,ij} = as_{ij}^{1/2} + bs_{ij}^{c} \quad (23)$$

Valence Dipole Energy. For the valence dipole term, we used a simple harmonic energy penalty function (Eqn. 24), where $k_{VD,i}$ is a type of spring constant. This form is unusual, because the units of $k_{VD,i}$ are not the typical energy/distance^2, but rather energy/valence^2.

$$E_{VD,i} = k_{VD,i}\left(\left\|\vec{P_i}\right\| - \left\|\vec{P_i}\right\|_{Ideal}\right)^2 \quad (24)$$

It is because of this unit difference that we have chosen the designation "spring constant" rather than the more traditional "force constant" for both the valence dipole and quadrupole moment restraint equations.

The key to this term lies in the determination of the ideal value. In many cases, especially normal cations not subject to lone-pair or Second-Order Jahn-Teller effects, and bonded to a single type of anion, $\left\|\vec{P_i}\right\|_{Ideal} = 0$ v.u. In the Al-Si-H-O system, only O (due to lone-pair effects) and H (due to its small size and asymmetric bonding) had to be treated differently. For such elements there is likely to be not one, but multiple values of $\left\|\vec{P_i}\right\|_{Ideal}$, each of which corresponds to a particular arrangement of bonds. These will depend on both the coordination number and the identity of the bound atoms.

We have only a very limited ability to guess what ideal values might be. In general, they will depend strongly on the coordination number and the row number of the bound atoms. In many cases transitions between the known points are impossible to determine.

As the coordination number increases past four, it is expected that oxygen will act as a spherical anion, and $\left\|\vec{P}_i\right\|_{Ideal}$ should approach zero (Bickmore et al., 2013).

Quadrupole Valence

We assumed a similar energy cost function for the valence quadrupole terms (Eqn. 25).

$$E_{VQ} = k_{VQ,i}(\|\mathbf{\Theta}_i\| - \|\mathbf{\Theta}_i\|_{Ideal})^2 \quad (25)$$

We found that valence quadrupole terms were not needed for the Al-Si-H-O system, so we mention them here only to note that we made the attempt. However, we are certain that valence quadrupole terms will be needed for transition metals subject to centrosymmetric First-Order Jahn-Teller distortions (Shepherd et al., 2016), and many actinide ions (e.g., uranyl).

METHODS

Given the basic form of the potential energy model, optimizing the model for the Al-Si-H-O system involved 1) constructing appropriate sets of data for calibrating and testing the model, and 2) the optimization procedure itself.

Data Sets

To calibrate the model, we needed both structural and thermodynamic data for a wide range of species, but finding such data proved to be a substantial difficulty. Even in

cases where structural data was available, it was often not of sufficient quality. The VMM is an extremely sensitive model, and errors in bond lengths as small as 0.02 Å can cause significant problems. Therefore, we used the most recent, suitable structural data available, and only used structures determined by neutron diffraction where H was present. Similarly, thermodynamic data were not always available.

We employed three data sets, including both molecules and crystals in the Al-Si-H-O system: one for calibration of Al-O, Si-O, H-O, H-H, and O-O interactions, another for calibration of Al-Al, Si-Si, Al-Si, Al-H, and Si-H interactions, and a smaller set for checking the results of the first set. Thermodynamic data for the data sets was obtained from the online NIST JANAF thermochemical tables (http://kinetics.nist.gov/janaf/), or from Robie and Hemingway (1995).

The set used for calibrating Al-O, Si-O, H-O, H-H, and O-O interactions was well-tempered, spanning a very large range of likely bond configurations. It included the fully dissociated atoms ($Al_{(g)}$, $H_{(g)}$, $Si_{(g)}$, and $O_{(g)}$), $SiO_{2(g)}$, $HOAlO_{(g)}$, several $SiO_{2(s)}$ polymorphs (quartz, cristobalite, coesite, and stishovite), α-$Al_2O_{3(s)}$ (corundum), and the $Al_2SiO_{5(s)}$ polymorphs (kyanite, andalusite, and sillimanite).(Busing and Levy, 1965; Cox et al., 1973; Hill, 1981; Bish, 1993; Downs and Palmer, 1994; Heaney and Post, 2001; Dera et al., 2003; Balan et al., 2008) In all cases Al was in the +3 oxidation state and Si in +4.

The set used for calibrating Al-Al, Si-Si, Al-Si, Al-H, and Si-H interactions included $Al_{(s)}$, $Si_{(s)}$, $Al_{2(g)}$, $Si_{2(g)}$, $AlH_{3(s)}$, $AlH_{3(g)}$, $HAlO_{(g)}$, and $SiH_{4(g)}$. In all these systems Al and Si were either partially or fully reduced. In addition, we included $Al_2H_{6(g)}$* and Si2$H_{6(g)}$*, for which we did not have complete thermodynamic data. Therefore, we performed optimizations and frequency calculations at the CCSD/cc-pVTZ level of theory in Gaussian 09, and

computed thermodynamic values from reaction calculations involving other members of the various sets.

The check set contained several species, against which we checked the calibrated model for the fully oxidized Al and Si parameters. Some species were expected to give good results: $Al(OH)_3$ (bayerite), AlOOH (boehmite), $Al_4Si_4O_{10}(OH)_8$ (kaolinite), and SiO2 (tridymite, moganite, and seifertite). However, we had no expectations about how the model would perform for others, for various reasons: e.g., radicals like AlO•, AlOO•, HO_2•, and H_3O^+, or SiO_2 (β-quartz at 848 K). (Newnham and de Haan, 1962; Dollase, 1965; Rothbauer et al., 1967; Wardle and Brindley, 1972; Saalfeld and Wedde, 1974; Dollase and Baur, 1976; d'Amour et al., 1978; Finger and Hazen, 1978; Hill, 1979; Levien et al., 1980; Levien and Prewitt, 1981; Joswig and Drits, 1986; Taylor et al., 1986; Spackman et al., 1987; Hazen et al., 1989; Ross et al., 1990; Glinnemann et al., 1992; Miehe and Graetsch, 1992; Boisen et al., 1994; Smyth et al., 1995; Daniels and Wunder, 1996; Comodi et al., 1997; Clark et al., 1998; Schmidt et al., 1998; Dera et al., 2002; Balan et al., 2006; Ikuta et al., 2007; Antao et al., 2008; Demichelis et al., 2008) However, we did not have enough species with Al-Al, Si-Si, Al-Si, Al-H, and Si-H bonds to include any of those in the check set.

Tempering. The first calibration set was specifically selected to be tempered with respect to bond valence variations, meaning that, as far as possible, species were chosen to cover the complete range of likely bond valences, while avoiding too frequent inclusion of any particular narrow range. Since we were attempting to fit D_E vs. s_{ij}, curves as exactly as possible, it was necessary to select at least one molecule or crystal with bond valences similar to the most probable values, but heavily favoring one narrow range (e.g., most Si-O bonds are ~1 v.u.) might worsen the fit in other regions of the curve. For example, Si forms

443 bonds with O of 2 v.u., ~1 v.u., and ~2/3 v.u., so we included at least one bond in each of
444 these ranges. Al-O bonds occur over an even wider range, with values around ⅓, ½, ⅗, ¾,
445 1, and 2 v.u. H-O bonds were particularly difficult, because almost all of them are close to 1
446 v.u. This places great emphasis on the effects of small deviations from unity and very heavy
447 emphasis on the few data points we have away from 1 v.u., such as the gas-phase water
448 dimer hydrogen bond.

449 In addition, we expected a valence dipole term on oxygen to be significant for
450 tetrahedrally-coordinated Si in silicates. Therefore, we used a carefully balanced selection
451 of those, favoring the most stable minerals (under ambient conditions).

452 This level of attention to this issue was entirely warranted, because our model is
453 likely to perform relatively poorly if applied to systems that have bonding configurations
454 significantly different from those in the calibration set, even if they are bracketed by the
455 set. For example, we have no Si-O bonds of ~1.5 v.u. in the set. Given that the form of our
456 D_E vs. s_{ij}, curves (Eqns. 22-23) is potentially quite flexible, the model might produce a large
457 error in such a scenario. The reason is straightforward: a small error in D_E is cumulative
458 over all the bonds in the system, and so will produce total errors on the order of 5-10x
459 higher than those of an individual bond.

460 Another reason to restrict ourselves to small sets is due to the exponential form of
461 the bond valence length relationships (Eqns. 5, 13, and 14). This causes a systematic error
462 that affects all bond-valence fitting procedures. The details of this source of error are
463 discussed in the online Supplemental Information, since it is likely only of interest to a
464 select group of methodologists.
465

Fitting procedure

We used a three-stage fitting procedure to optimize the potential energy model. The first stage involved fitting the bond valence-length curves (Eqns. 5, 13, 14), as described above. This was done either using vibrational spectroscopy data, where bond lengths could be related to bond order based on measured force constants, or the more traditional approach of optimizing the fitting parameters to enforce the valence sum rule (Eqn. 12).

We could only use vibrational spectroscopy data for H-H, O-O, and O-H atom pairs, because they were the only sets with enough data for bonds of different lengths. For H-H we had H_2, and H_2^+ (experimental data from the NIST Computational Chemistry Comparison and Benchmark Database at http://cccbdb.nist.gov. Hereafter referenced as CCCBDB). For O-O we had O_2^+, O_2, and O_2^- (CCCBDB). For H-O we had H_2O, H_2O_2, OH, OH^-, OH^+, H_3O^+ (CCCBDB), and the $(H_2O)_2$ gas phase dimer (Kalescky et al., 2012). For this approach, it is possible to use charged species, provided they clearly follow the same trend as the others. For example, ozone shows substantial force relaxation due to the impossibility of satisfying all valences (without forming triangular ozone, which is unstable). This is quite common among strained molecules. Unfortunately, it is not yet possible to use larger molecules like H_2O_2 in this procedure, because the vibrational modes are too closely coupled.

The second approach to the bond valence-length fitting scheme was much more similar to the classic procedure (Brown and Altermatt, 1985), except that we used bond lengths from both molecular and crystal structures. Cutoff distances for calculating bond valence corresponded approximately to the minimum in the radial distribution function

between first and second shell coordination. In this case we also had to utilize other parameter sets, specifically OO and OH for both AlO and SiO and SiO for AlO fitting. The SiO set was $SiO_{2(g)}$, $HSiO_2^-{}_{(g)}$, $SiOH_{4(g)}$, quartz, and stishovite; and the AlO set was: $AlO^-{}_{(g)}$, $HAlO_{(g)}$, $Al(OH)_{3(g)}$, sillmanite, andalucite, and corundum. This approach requires a separate known reference point to establish the k_1 value (Eqn. 15), such as the vibrational force constant of a small molecule like $AlO_{(g)}$.

In both procedures (classic and force matching), we began by fitting Eqn. 5 to the data, and if the curvature of the fit was inadequate, we then progressed to Eqns. 13 and 14. For the Al-Si bonding we could find no data, so we used a mixing approximation. Our choice was the arithmetic mean for R_0 and the geometric mean for k_1/B.

Ideally, the second stage of the procedure would involve optimizing D_E vs. s_{ij}, curves (Eqns. 22-23) to match the structural and thermodynamic data. But given that the thermodynamic data was not corrected for zero-point energies, in practice we were limited to optimizing D_0 vs. s_{ij}, curves. The k_1 values were fixed at the time of the bond-valence parameter fitting, and a constrained optimization was performed simultaneously over the D_0 values for all atom pairs.

In the third stage, we re-optimized all the D_0 vs. s_{ij}, curves, sequentially including the subsequent terms of the potential energy model (valence dipole and quadrupole terms). The results of the first optimization were used as initial estimates for the second optimization, and so on.

RESULTS AND DISCUSSION

511 The form of the valence multipole energy decomposition is unique, and and has not
512 been designed from one specific theoretical framework. Instead, the only theoretical basis
513 for our model is simply the standard electron counting rules that undergird the BVM, while
514 the rest of the model is based on conjecture and empirical fitting. Our results, nevertheless,
515 show that, although the mathematical form of the model is very simple compared to
516 existing reactive force fields, it demonstrates the capability to reasonably capture the
517 chemical energy of the system. Moreover, the simplicity of the model allows us, in some
518 cases, to rationalize the model in terms of theoretical expectations.
519 In this section, we report the results of the model optimization, including parameter
520 values and overall accuracy, and discuss possible connections with theory. Finally, we
521 discuss probable necessary refinements of the model.
522
523 *Model Optimization*
524 Table 1 shows the fitted parameters for the bond valence-length relationships
525 (Eqns. 5, 13, 14). These results conform to the expectation that more metallic and ionic
526 bonds require mathematically more complex relationships (Eqns. 13 and 14) to capture the
527 high curvature around s_{ij} = 1 v.u. This can be rationalized in terms of the transition
528 between σ- and π-bonding (Wander et al., 2015b).
529 The D_0 vs. s_{ij} curves were first optimized with only monopole terms (see Table 2 and
530 Figure 1a), and then with both monopole and dipole terms (see Table 3 and Figure 1b).
531 The graphs in Figures 1a and 1b are quite similar, as we would expect if the monopole term
532 were indeed dominant. This also seems consistent with the long-standing usefulness of
533 bond- and group-additivity methods for estimating thermodynamic properties of

534 molecules (Benson and Buss, 1958; Fliszár, 2008). It should also be noted that the valence-
535 energy relationships derived here do not jibe with those one might predict based on any
536 number of BVM-based quantitative structure-activity relationships (Bickmore, 2014)

537 The curves in Figure 1b show a periodic quality. The most ionic bonds (Al-O, Si-O, H-
538 O) rise quickly as a function of s_{ij} and then drop off as the bond valence surpasses 1 v.u.,
539 while the more covalent bonds (H-H and O-O) are much slower to rise, but then gain
540 strength rapidly as the covalent bonding becomes more significant. This is consistent with
541 the fact that ionic atom pairs favor a greater number of low-valence bonds, while more
542 covalent pairs favor a smaller number of high-valence bonds (Brown and Skowron, 1990;
543 Brown, 2002).

544 Compared to other covalent bonds, HH bonds at fractional bond orders are quite
545 strong. This is probably due to the fact that hydrogen has no core electrons and any
546 substantial polarization can lead to a bare proton. Furthermore, because there are no non-
547 bonding electrons there is a substantially reduced, if not eliminated, reorganization energy,
548 which would further reduce the bonding energy.

549 Another particularly strong bond, which is unfortunately not visible on the scale of
550 our graph is the O-H hydrogen bond. Despite being only a small fraction of a single bond, it
551 still has an energy of ~24 kJ/mol. The only way to explain this is to suggest that the OH
552 hydrogen bond is substantially more ionic than the OH molecular bond.

553 The fact that these curves lend themselves to physical interpretation, despite the
554 fact that they are purely empirical fits, with no *a priori* considerations about the form,
555 suggests an underlying physical basis. This seems promising in terms of the ability to
556 expand the model's predictive capability.

When we included dipole terms, we used $\left\|\overrightarrow{P_H}\right\|_{Ideal} = 0.95$ v.u. for H bound to O, and a k_{VD} value of 85.4 kJ•mol^{-1}•v.u^{-2} (Eqn. 24), and considered it unlikely to have a significant contribution to the energy, otherwise. To avoid multiplication of fitting parameters, we did not attempt to fit dipole terms for Si and Al. We have no chemical justification for their inclusion, and we considered these terms unlikely to make a substantial difference in the predicted energies of minerals.

Table 4 shows a relevant selection of likely O bonding configurations used to calculate $\left\|\overrightarrow{P_O}\right\|_{Ideal}$, with k_{VD} = 284.3 kJ•mol^{-1}•v.u^{-2} (Eqn. 24). The results generally conform to the expectation that higher-valence, and more covalent bonds, cause more distortion in the coordination shells of anions with lone pairs (Gillespie and Hargittai, 1991; Gillespie and Johnson, 1997; Bickmore et al., 2013).

The results of the final energy optimization are reported in Tables 5 and 6 for the calibration and check sets, respectively. The standard deviation of the energy error on the calibration sets was 0.8 kJ/mol per unique atom (see also Figure 4a.)

At first glance, the check set results (Figure 4b) are not as accurate as we might like. However, most of these are not typical or ideal systems. The first half are entirely composed of systems in which the valence sum rule cannot even be approximately satisfied. The second half actually does quite well with two exceptions, and both exceptions have the same basic problem over estimation of the dipole term on H, which we will explain below. Otherwise, the results are excellent, with the standard deviation error per atom of about 5 kJ/mol.

The model clearly fails when applied to molecules with significantly unsatisfied valence. We need to understand the effect on force constants when Bond Valence sums

cannot be satisfied. The classic example is ozone whose BV sums are closer to 1.5 and 2.5 v.u., rather than the expected 2. This deviation causes a proportionate weakening between the predicted (1000 N/m) and actual (570 N/m) force constant. Until we can address this, we are limited to systems where the BV sums are nearly satisfied. Since this problem is only likely to occur in the gas phase, or with certain radicals, the level of limitation is not too severe. Some part of this may be unsolvable, as the very assumptions that underlie our BV potential assume saturated or near saturated valence. In this case, single molecule force fields could be created if sufficient need warranted it. However, we suspect that there is a fundamental relationship between force constants and bond-valence saturation that will yield to analysis over time.

Two of the crystals had substantial per-atom errors on the order of 100 kJ/mol, boehmite (γ-AlOOH) and bayerite (β-$Al(OH)_3$). Both exhibited substantial underbonding of H of 20% or more. This leads to strong underestimation of the H-bond dipole and, correspondingly, a large positive deviation in energy. The geometric difference here is actually quite tiny, perhaps 0.02Å, so it may be that there are errors in the structure determinations. It may also be that the valence dipole potential energy surface for H is asymmetric, which is something we can explore in future implementations of the model.

Finally, Tables 8 and 9 report the results from the reduced calibration set. Table 8 shows the D_0 vs. s_{ij} curves and Table 9 shows the energy results. There is no way to confirm these values as we have no additional experimental data.

IMPLICATIONS

Our results show that it is possible to use an energy decomposition based on a valence multipole expansion to determine reasonable system energies, with chemically useful accuracy. As long as deviations in the bond-valence sums are less than about 0.2 v.u., this model should produce results with errors of only 5 kJ/mol per unique atom. This is a remarkable result for any black-box model, on par with comparable quantum mechanical approaches. Also, the model should have similar accuracy when applied to either molecules or crystals.

While the results are promising, some work still remains to fully implement this type of model for use in molecular modeling. The first issue is that model calibration requires enough bonds for every pair of atoms, covering a wide enough range of bond orders, which presents a challenge when considering the range of available experimental data. We will clearly need to move into computational sources for our data collection. Fortunately, since our parameterization relies largely on small molecules, it will be possible to use high-level *ab initio* approaches like coupled-cluster techniques. Even with this level of theory, however, the requisite accuracy is not a given, particularly for metallic systems.

We also anticipate the necessity of making adjustments in the shapes of the D_0 vs. s_{ij} curves, especially at high s_{ij} values, to accommodate non-equilibrium configurations encountered in molecular dynamics simulations, as well as high-pressure phases. Additionally, changes to the bond energy expression are also likely to better match changes in energy far away from equilibrium geometries and better reproduce dissociation curves.

Finally, all the interactions discussed so far only deal with very short-range interactions. To model systems in which long-range interactions play a larger role (e.g.,

aqueous solutions), we will have to consider Coulombic interactions, at least outside of the bond cutoffs.

Acknowledgements

We gratefully acknowledge the National Science Foundation (EAR-1227215 and EAR-1424682) and the Brigham Young University College of Physical and Mathematical Sciences undergraduate mentoring program for funding this project. Professor I. D. Brown and an anonymous reviewer provided helpful comments on the original manuscript.

References

Antao, S.M., Hassan, I., Wang, J., Lee, P.L., and Toby, B.H. (2008) State-of-the-art high-resolution powder x-ray diffraction (HRPXRD) illustrated with Rietveld structure refinement of quartz, sodalite, tremolite, and meionite. The Canadian Mineralogist, 46, 1501-1509.

Badger, R.M. (1934) A relation between internuclear distances and bond force constants. Journal of Chemical Physics, 2, 128-131.

Balan, E., Lazzeri, M., Morin, G., and Mauri, F. (2006) First-principles study of the OH-stretching modes of gibbsite. American Mineralogist, 91, 115-119.

Balan, E., Blanchard, M., Hochepied, J.-F., and Lazzeri, M. (2008) Surface modes in the infrared spectrum of hydrous minerals: The OH stretching modes of bayerite. Physics and Chemistry of Minerals, 35, 279-285.

Benson, S.W. and Buss, J.H. (1958) Additivity rules for the estimation of molecular properties. Thermodynamic properties. Journal of Chemical Physics, 29, 546-572.

Bickmore, B.R. (2014) Structure and acidity in aqueous solutions and oxide-water interfaces. Structure and Bonding, 158, 191-203.

Bickmore, B.R., Rosso, K.M., and Mitchell, S.C. (2006a) Is there hope for multisite complexation modeling? In J. Lützenkirchen, Ed., Surface Complexation Modelling. Elsevier, Amsterdam.

Bickmore, B.R., Tadanier, C.J., Rosso, K.M., Monn, W.D., and Eggett, D.L. (2004) Bond-valence methods for pK_a prediction: Critical reanalysis and a new approach. Geochimica et Cosmochimica Acta, 68, 2025-2042.

Bickmore, B.R., Rosso, K.M., Tadanier, C.J., Bylaska, E.J., and Doud, D. (2006b) Bond-valence methods for pKa prediction. II. Bond-valence, electrostatic, molecular geometry, and solvation effects. Geochimica et Cosmochimica Acta, 70, 4057-4071.

Bickmore, B.R., Wander, M.C.F., Edwards, J., Maurer, J., Shepherd, K., Meyer, E., Johansen, W.J., Frank, R.A., Andros, C., and Davis, M. (2013) Electronic structure effects in the vectorial bond-valence model. American Mineralogist, 98, 340-349.

Bish, D.L. (1993) Rietveld refinement of the kaolinite structure at 1.5 K. Clays and Clay Minerals, 41, 738-744.

Boisen, M.B., Gibbs, G.V., and Bukowinski, M.S.T. (1994) Framework silica structures generated using simulated annealing with a potential energy function based on an $H_6Si_2O_7$ molecule. Physics and Chemistry of Minerals, 21, 269-284.

Brenner, D.W. (1990) Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. Physical Review B, 42, 9458-9471.

Brown, I.D. (2002) The Chemical Bond in Inorganic Chemistry: The bond valence model, 278 p. Oxford University Press, New York.

Brown, I.D. (2009) Recent developments in the methods and applications of the bond valence model. Chemical Reviews, 109, 6858-6919.

Brown, I.D. (2014) Bond valence theory. Structure and Bonding, 158, 11-58.

Brown, I.D. and Altermatt, D. (1985) Bond-valence parameters obtained from a systematic analysis of the inorganic crystal structure database. Acta Crystallographica, B41, 244-247.

Brown, I.D. and Skowron, A. (1990) Electronegativity and Lewis acid strength. Journal of the American Chemical Society, 112, 3401-3403.

Busing, W. and Levy, A. (1965) Crystal and molecular structure of hydrogen peroxide. A neutron-diffraction study. Journal of Chemical Physics, 42, 3054-3059.

Clark, G.R., Rodgers, K.A., and Henderson, G.S. (1998) The crystal chemistry of doyleite, $Al(OH)_3$. Zeitschrift fur Kristallographie, 213, 96-100.

Comba, P., Hambley, T.W., and Martin, B. (2009) Molecular modeling of inorganic compounds, 326 p. Wiley-VCH, Weinheim.

Comodi, P., Zanazzi, P.F., and Poli, S. (1997) High-pressure behavior of kyanite: Compressibility and structural deformation. American Mineralogist, 82, 452-459.

Cooper, V.R., Grinberg, I., and Rappe, A.M. (2003) Extending first principles modeling with crystal chemistry: A bond-valence based classical potential. In P.K. Davies and D.J. Singh, Eds., Fundamental Physics of Ferroelectrics. American Institute of Physics, Melville, New York.

Cox, D.E., Samuelsen, E.J., and Ceckurts, K.H. (1973) Neutron-Diffraction determination of the crystal structure and magnetic form factor of gamma-oxygen. Physical Review, B7, 3102-3111.

Cramer, C.J. (2004) Essentials of Computational Chemistry, 596 p. Wiley, Chichester, UK.

d'Amour, H., Schiferl, D., Denner, W., Schulz, H., and Holzapfel, W.B. (1978) High-pressure single-crystal structure determinations for ruby up to 90 kbar using an automatic diffractometer. Journal of Applied Physics, 49, 4411-4416.

Daniels, P. and Wunder, B. (1996) $Al_3Si_2O_7(OH)_3$, phase Pi (formerly piezotite): Crystal structure of a synthetic high-pressure silicate rediscovered. European Journal of Mineralogy, 8, 1283-1292.

Demichelis, R., Noel, Y., Zicovich-Wilson, C.M., Roetti, C., Valenzano, L., and Dovesi, R. (2008) Ab-initio quantum mechanical study of akdalaite ($5Al_2O_3*H_2O$): Structure and vibrational spectrum. Journal of Physics: Conference Series, 117, 012013-012013.

Dera, P., Prewitt, C.T., Boctor, N.Z., and Hemley, R.J. (2002) Characterization of a high-pressure phase of silica from the Martian meteorite Shergotty. American Mineralogist, 87, 1018-1023.

Dera, P., Prewitt, C.T., Japel, S., Bish, D.L., and Johnston, C.T. (2003) Pressure-controlled polytypism in hydrous layered materials. American Mineralogist, 88, 1428-1435.

Dollase, W.A. (1965) Reinvestigation of the structure of low cristobalite. Zeitschrift fur Kristallographie, 121, 369-377.

Dollase, W.A. and Baur, W.H. (1976) The superstructure of meteoritic low tridymite solved by computer simulation. American Mineralogist, 61, 971-978.

Downs, R.T. and Palmer, D.C. (1994) The pressure behavior of alpha cristobalite. American Mineralogist, 79, 9-14.

Etxebarria, I., Perez-Mato, J.M., García, A., Blaha, P., Schwartz, K., and Rodriguez-Carvajal, J. (2005) Comparison of empirical bond-valence and first-principles energy calculations for a complex structural instability. Physical Review B, 72, 174108.

Finger, L.W. and Hazen, R.M. (1978) Crystal structure and compression of ruby to 46 kbar. Journal of Applied Physics, 49, 5823-5826.

Finnis, M.W. and Sinclair, J.E. (1984) A simple empirical N-body potential for transition metals. Philosophical Magazine A, 50, 45-55.

Fliszár, S. (2008) Atomic Charges, Bond Properties, and Molecular Energies, 234 p. Wiley, Hoboken, New Jersey.

Gillespie, R.J. and Hargittai, I. (1991) The VSEPR Model of Molecular Geometry, 248 p. Allyn and Bacon, Boston.

Gillespie, R.J. and Johnson, S.A. (1997) Study of bond angles and bond lengths in disiloxane and related molecules in terms of topology of the electron density and its Laplacian. Inorganic Chemistry, 36, 3031-3039.

Glinnemann, J., King, H.E., Schulz, H., Hahn, T., La Placa, S.J., and Dacol, F. (1992) Crystal structures of the low-temperature quartz-type phases of SiO_2 and GeO_2 at elevated pressure. Zeitschrift fur Kristallographie, 198, 177-212.

Golub, G.H. (2013) Matrix Computations, 728 p. Johns Hopkins University Press, Baltimore.

Grinberg, I., Cooper, V.R., and Rappe, A.M. (2002) Relationship between local structure and phase transitions of a disordered solid solution. Nature, 419, 909-911.

Grinberg, I., Cooper, V.R., and Rappe, A.M. (2004) Oxide chemistry and local structure of $PbZr_xTi_{1-x}O_3$ studied by density-functional theory supercell calculations. Physical Review B, 69, 144118.

Grinberg, I., Shin, Y.-H., and Rappé, A.M. (2009) Molecular dynamics study of dielectric response in a relaxor ferroelectric. Physical Review Letters, 103, 197601.

Harvey, M.A., Baggio, S., and Baggio, R. (2006) A new simplifying approach to molecular geometry description: the vectorial bond-valence model. Acta Crystallographica, B62, 1038-1042.

Hazen, R.M., Finger, L.W., Hemley, R.J., and Mao, H.K. (1989) High-pressure crystal chemistry and amorphization of alpha-quartz. Solid State Communications, 72, 507-511.

Heaney, P.J. and Post, J.E. (2001) Evidence for an I2/a to Imab phase transition in the silica polymorph moganite at ~570 K. American Mineralogist, 86, 1358-1366.

Hiemstra, T. and Van Riemsdijk, W.H. (1996) A surface structural approach to ion adsorption: The charge distribution (CD) model. Journal of Colloid and Interface Science, 179, 488-508.

Hiemstra, T., Van Riemsdijk, W.H., and Bolt, G.H. (1989) Multisite proton adsorption modeling at the solid/solution interface of (hydr)oxides: A new approach. I. Model description and evaluation of intrinsic reaction constants. Journal of Colloid and Interface Science, 133, 91-104.

Hiemstra, T., Venema, P., and Van Riemsdijk, W.H. (1996) Intrinsic proton affinity of reactive surface groups of metal (hydr)oxides: The bond valence principle. Journal of Colloid and Interface Science, 184, 680-692.

Hill, R.J. (1979) Crystal structure refinement and electron density distribution in diaspore. Physics and Chemistry of Minerals, 5, 179-200.

Hill, R.J. (1981) Hydrogen atoms in boehmite: A single crystal X-ray diffraction and molecular orbital study. Clays and Clay Minerals, 29, 435-445.

Hinchliffe, A. (2003) Molecular Modelling for Beginners, 410 p. Wiley, Chichester, UK.

Horn, R.A. and Johnson, C.R. (2012) Matrix Analysis, 655 p. Cambridge University Press, Cambridge.

Ikuta, D., Kawame, N., Banno, S., Hirajima, T., Ito, K., Rakovan, J.F., Downs, R.T., and Tamada, O. (2007) First in situ X-ray diffraction identification of coesite and retrograde quartz on a glass thin section of an ultrahigh-pressure metamorphic rock and their crystal structure details. American Mineralogist, 92, 57-63.

Johnston, V.H.S. (1966) Gas Phase Reaction Rate Theory, 372 p. Ronald Press, New York.

Joswig, W. and Drits, V.A. (1986) The orientation of the hydroxyl groups in dickite by X-ray diffraction. Neues Jahrbuch fur Mineralogie, Monatshefte, 1986, 19-22.

Kalescky, R., Zou, W., Kraka, E., and Cremer, D. (2012) Local vibrational modes of the water dimer—Comparison of theory and experiment. Chemical Physics Letters, 554, 243-247.

Levien, L. and Prewitt, C.T. (1981) High-pressure crystal structure and compressibility of coesite. American Mineralogist, 66, 324-333.

Levien, L., Prewitt, C.T., and Weidner, D.J. (1980) Structure and elastic properties of quartz at pressure. American Mineralogist, 65, 920-930.

Liang, T., Shin, Y.K., Cheng, Y.-T., Yilmaz, D.E., Vishnu, K., G., Verners, O., Zou, C., Phillpot, S.R., Sinnott, S.B., and van Duin, A.C.T. (2013) Reactive potentials for advanced atomistic simulations. Annual Review of Materials Research, 43, 109-129.

Liu, S., Grinberg, I., and Rappé, A.M. (2013a) Development of a bond-valence based interatomic potential for BiFeO3 for accurate molecular dynamics simulations. Journal of Physics: Condensed Matter, 25, 102202.

Liu, S., Grinberg, I., Takenaka, H., and Rappé, A.M. (2013b) Reinterpretation of the bond-valence model with bond-order formalism: An improved bond-valence-based interatomic potential for $PbTiO_3$. Physical Review B, 88, 104102.

Lufaso, M.W. and Woodward, P.M. (2001) Prediction of the crystal structures of perovskites using the software program SPuDS. Acta Crystallographica, B57, 725-738.

Miehe, G. and Graetsch, H. (1992) Crystal structure of moganite: A new structure type of silica. European Journal of Mineralogy, 4, 693-706.

Newnham, R.E. and de Haan, Y.M. (1962) Refinement of the alpha Al_2O_3, Ti_2O_3, V_2O_3 and Cr_2O_3 structures. Zeitschrift fur Kristallographie, 117, 235-237.

O'Keeffe, M. and Brese, N.E. (1992) Bond-valence parameters for anion-anion bonds in solids. Acta Crystallographica, B48, 152-154.

Perez-Mato, J.M., Withers, R.L., Larsson, A.-K., Orobengoa, D., and Liu, Y. (2009) Distortion modes and related ferroic properties of the stuffed tridymite-type compounds $SrAl_2O_4$ and $BaAl_2O_4$. Physical Review, B79, 064111.

Rappé, A.K. and Casewit, C.J. (1997) Molecular mechanics across chemistry, 444 p. University Science Books, Sausalito, CA.

Robie, R.A. and Hemingway, B.S. (1995) Thermodynamic Properties of Minerals and Related Substances at 298.15 K and 1 bar (10^5 Pascals) Pressure and at Higher Temperatures. U.S. Geological Survey Bulletin, 2131, 1-461.

Ross, N.L., Shu, J.F., Hazen, R.M., and Gasparik, T. (1990) High-pressure crystal chemistry of stishovite. American Mineralogist, 75, 739-747.

Rothbauer, R., Zigan, F., and O'Daniel, H. (1967) Verfeinerung der struktur des bayerits, $Al(OH)_3$. Zeitschrift fur Kristallographie, 125, 317-331.

Rowley, R.L. (1994) Statistical Mechanics for Thermophysical Property Calculations, 489 p. Prentice-Hall, Englewood Cliffs, New Jersey.

Saalfeld, H. and Wedde, M. (1974) Refinement of the crystal structure of gibbsite, $Al(OH)_3$. Zeitschrift fur Kristallographie, 139, 129-135.

Sahai, N. (2002) Is silica really an anomalous oxide? Surface acidity and aqueous hydrolysis revisited. Environmental Science and Technology, 36, 445-452.

Sahai, N. and Sverjensky, D.A. (1997a) Sovation and electrostatic model for specific electrolyte adsorption. Geochimica et Cosmochimica Acta, 61, 2827-2848.

Sahai, N. and Sverjensky, D.A. (1997b) Evaluation of internally consistent parameters for the triple-layer model by the systematic analysis of oxide surface titration data. Geochimica et Cosmochimica Acta, 61, 2801-2826.

Schmidt, M.W., Finger, L.W., Angel, R.J., and Dinnebier, R.E. (1998) Synthesis, crystal structure, and phase relations of $AlSiO_3OH$, a high-pressure hydrous phase. American Mineralogist, 83, 881-888.

Shepherd, K., Johansen, W.J., Goodell, T., Bickmore, B.R., Wander, M.C.F., Davis, M., Andros, C., Lind, L., and Robertson, K. (2016) The valence quadrupole moment. American Mineralogist, 101, 362-3170.

Shin, Y.-H., Cooper, V.R., Grinberg, I., and Rappe, A.M. (2005) Development of a bond-valence molecular-dynamics model for complex oxides. Physical Review B, 71, No. 054104.

Shin, Y.-H., Grinberg, I., Chen, I.-W., and Rappé, A.M. (2007) Nucleation and growth mechanism of ferroelectric domain-wall motion. Nature, 449, 881-884.

Shin, Y.-H., Son, J.-Y., Lee, B.-J., Grinberg, I., and Rappé, A.M. (2008) Order-disorder character of $PbTiO_3$. Journal of Physics: Condensed Matter, 20, 015224.

Smyth, J.R., Swope, R.J., and Pawley, A.R. (1995) H in rutile-type compounds: II. Crystal chemistry of Al substitution in H-bearing stishovite. American Mineralogist, 80, 454-456.

Spackman, M.A., Hill, R.J., and Gibbs, G.V. (1987) Exploration of structure and bonding in stishovite with fourier and pseudoatom refinement methods using single crystal and powder X-ray diffraction data. Physics and Chemistry of Minerals, 14, 139-150.

Sverjensky, D.A. (1994) Zero-point-of-charge prediction from crystal chemistry and solvation theory. Geochimica et Cosmochimica Acta, 58, 3123-3129.

Sverjensky, D.A. (2005) Prediction of surface charge on oxides in salt solutions: Revisions for 1:1 (M+L-) electrolytes. Geochimica et Cosmochimica Acta, 69, 225-257.

Sverjensky, D.A. and Sahai, N. (1996) Theoretical prediction of single-site surface-protonation equilibrium constants for oxides and silicates in water. Geochimica et Cosmochimica Acta, 60, 3773-3797.

Sverjensky, D.A., L., S.E., and Helgeson, H.C. (1997) Prediction of the thermodynamic properties of aqueous metal complexes to 1000 °C and 5 kb. Geochimica et Cosmochimica Acta, 61, 1359-1412.

Takenaka, H., Grinberg, I., and Rappé, A.M. (2013) Anisotropic local correlations and dynamics in a relaxor ferroelectric. Physical Review Letters, 110, 147602.

Taylor, J.C., Miller, S.A., and Bibby, D.M. (1986) A study of decomposition methods for refinement of H+-ZSM5 zeolite with powder diffraction data. Zeitschrift fur Kristallographie, 176, 183-192.

Tersoff, J. (1988) New empirical approach for the structure and energy of covalent systems. Physical Review B, 37, 6991-7000.

van Duin, A.C.T., Strachan, A., Stewman, S., Zhang, Q., X., X., and Goddard, W.A.I. (2003) $ReaxFF_{Si0}$ reactive force field for silicon and silicon oxide systems. Journal of Physical Chemistry A, 107, 3803-3811.

Wander, M.C.F., Bickmore, B.R., Davis, M., Johansen, W.J., Andros, C., and Lind, L. (2015a) The use of cation-cation and anion-anion bonds to augment the bond-valence model. American Mineralogist, 100, 148-159.

Wander, M.C.F., Bickmore, B.R., Lind, L., Andros, C., Hunt, J., Checketts, H., and Goodell, T. (2015b) AIM analysis and the form of the bond-valence equation. American Mineralogist, 100, 160-161.

Wardle, R. and Brindley, G.W. (1972) The crystal structures of pyrophyllite, 1Tc, and of its dehydroxylate. American Mineralogist, 57, 732-750.

869 Table 1. Optimized bond-valence parameters.

Atom 1	Atom 1 Valence	Atom 2	Atom 2 Valence	Form	$R_{0,1}$ (Å)	B_1 (Å)	$R_{0,2}$ (Å)	B_2 (Å)	R_{cut} (Å)	w
O	2	Al	3	Eqn. 14	1.7005	0.0924	2.2096	0.9990	3	0.5179
O	2	Si	4	Eqn. 14	1.5531	0.0575	2.3547	0.9612	3	0.5065
O	2	H	1	Eqn. 13	0.9530	0.1950	0	0	6	0.4835
O	2	O	2	Eqn. 5	1.4560	0.3579	0	0	6	1
Al	3	Al	3	Eqn. 5	0	0	0	0	6	1
H	1	Al	3	Eqn. 5	0	0	0	0	3	1
Si	4	Al	3	Eqn. 5	0	0	0	0	6	1
H	1	H	1	Eqn. 5	0.7919	0.5259	0	0	2	1
H	1	Si	4	Eqn. 5	0	0	0	0	3	1
Si	4	Si	4	Eqn. 5	0	0	0	0	6	1
Al	3	H	1	Eqn. 5	0	0	0	0	3	1
Al	3	Si	4	Eqn. 5	0	0	0	0	6	1
Si	4	H	1	Eqn. 5	0	0	0	0	3	1

870

871

872

873 **Table 2.** D_0 vs. s_{ij} parameters resulting from valence monopole-only energy optimization on the

874 tempered set. Force constants were taken from experimental data and not refit.

Atom 1	Atom 2	Force Const ($kJ \cdot mol^{-1} Å^{-2}$)	Form	*a*	*b*	*c*	*d*
O	Al	1777.3150	Eqn. 22	851.8089	-1192.1804	1224.6636	-383.0072
O	Si	2762.9621	Eqn. 22	527.0360	-22.0542	-47.0318	0
O	H	4656.4438	Eqn. 23	216.6120	247.5746	4.1099	0
O	O	3607.2626	Eqn. 22	38.4168	25.6211	35.4555	0
Al	Al	0	Eqn. 22	0	0	0	0
H	Al	0	Eqn. 22	0	0	0	0
Si	Al	0	Eqn. 22	0	0	0	0
H	H	3464.6712	Eqn. 22	265.6922	166.5006	-25.6530	0
H	Si	0	Eqn. 22	0	0	0	0
Si	Si	0	Eqn. 22	0	0	0	0
Al	H	0	Eqn. 22	0	0	0	0
Al	Si	0	Eqn. 22	0	0	0	0
Si	H	0	Eqn. 22	0	0	0	0

875

876

877 **Table 3.** D_0 vs. s_{ij} parameters resulting from valence dipole energy optimization on the tempered
878 set. Force constants were taken from experimental data and not refit.
879

Atom 1	Atom 2	Force Const ($kJ \cdot mol^{-1} Å^{-2}$)	Form	*a*	*b*	*c*	*d*
O	Al	1777.3150	Eqn. 22	739.2846	-834.8911	868.1735	-277.3448
O	Si	2762.9621	Eqn. 22	481.3612	32.0345	-62.7606	0
O	H	4656.4438	Eqn. 23	160.3074	315.0453	1.4734	0
O	O	3607.2626	Eqn. 22	170.4692	-91.7876	61.1265	0
Al	Al	0	Eqn. 22	0	0	0	0
H	Al	0	Eqn. 22	0	0	0	0
Si	Al	0	Eqn. 22	0	0	0	0
H	H	3464.6712	Eqn. 22	0.0074	963.5303	-556.9658	0
H	Si	0	Eqn. 22	0	0	0	0
Si	Si	0	Eqn. 22	0	0	0	0
Al	H	0	Eqn. 22	0	0	0	0
Al	Si	0	Eqn. 22	0	0	0	0
Si	H	0	Eqn. 22	0	0	0	0

880
881

882

883 **Table 4.** Configurations used to calculate $\left\|\overrightarrow{P_O}\right\|_{Ideal}$ in a variety of environments. Nc

884 denotes the oxygen coordination number.

885

N_c	Ligands	Ideal Angle(s)
0	None	N/A (VVS_{ideal}=0)
1	Any	N/A (VVS_{ideal}=S_1)
2	H	104.5°
2	Al_2, or Al & Si	180°
2	Si_2	145.5°
2	Al,H	130.8°
2	Si,H	118.7°
3	All Al or Si	131°, 131°, 98°
3	H	111.3°(x3)
4	Any (Generally only Al)	~109° ($\left\|\overrightarrow{P_i}\right\|_{Ideal}$ = 0.05 v.u.)
5+	Any (None in set)	($\left\|\overrightarrow{P_i}\right\|_{Ideal}$ = 0 v.u.)

886

Table 5. Energy fitting results from the oxidized tempered set. The experimental energies (E_{exp}) were corrected (ΔE_{corr}) from their tabulated values to reflect a free-atom standard state. The experimental energies are then compared to the calculated energies when monopole (E_{mono}), dipole (E_{di}), and quadrupole (E_{quad}) terms were included, to calculate the misfit values (ΔE_{mono}, ΔE_{di}, ΔE_{quad}). All values are reported in kJ/mol.

Molecule/Crystal	ΔE_{corr}	E_{exp}	E_{mono}	E_{di}	E_{quad}	ΔE_{mono}	ΔE_{di}	ΔE_{quad}
$(H_2O)_{2(g)}$	1276.578	-483.16	-480.63	-482.71	-482.71	-3.36	-0.21	-0.21
$Al_{(g)}$	289.068	289.07	289.07	289.07	289.07	0.00	0.00	0.00
$AlOOH_{(g)}$	955.817	-447.01	-447.01	-447.01	-447.01	0.00	0.00	0.00
Andalusite	2142.346	-2441.80	-2441.60	-2442.00	-2442.00	-0.58	0.15	0.15
Coesite	869.001	-852.30	-855.68	-854.66	-854.66	2.79	1.95	1.95
Corundum	1273.345204	-1582.30	-1580.72	-1582.93	-1582.93	-3.73	0.55	0.55
Cristobalite	869.001	-854.60	-855.19	-856.39	-856.39	-1.37	1.00	1.00
$H_2^+{}_{(g)}$	1726.521	1491.18	1491.19	1491.19	1491.19	-0.02	0.00	0.00
$H_{2(g)}$	406.552	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$H_2O_{(g)}$	638.289	-228.58	-233.63	-229.50	-229.50	6.70	0.45	0.45
$H_2O_{2(g)}$	870.026	-105.44	-105.16	-105.36	-105.36	-0.38	-0.08	-0.08
$H_{(g)}$	203.276	203.28	203.28	203.28	203.28	0.00	0.00	0.00
Kyanite	2142.346	-2443.10	-2444.44	-2442.26	-2442.26	3.62	-0.68	-0.68
$O_2^+{}_{(g)}$	1784.905	1139.99	1140.09	1140.02	1140.02	-0.14	-0.02	-0.02
$O_{2(g)}$	463.473	0.00	-0.20	-0.06	-0.06	0.28	0.05	0.05
$O_{(g)}$	231.737	231.74	231.74	231.74	231.74	0.00	0.00	0.00
Quartz	869.001	-865.30	-860.43	-861.65	-861.65	-3.80	-2.51	-2.51
$Si_{(g)}$	405.528	405.53	405.53	405.53	405.53	0.00	0.00	0.00
Sillimanite	2142.346	-2439.10	-2439.15	-2439.03	-2439.03	0.16	-0.05	-0.05

$SiO_{2(g)}$	869.001	-306.92	-306.94	-306.90	-306.90	0.06	-0.02	-0.02
Stishovite	869.001	-802.80	-802.54	-803.01	-803.01	-0.78	0.15	0.15
Diaspore	955.817	-922.70	-923.40	-922.87	-922.87	0.94	0.08	0.08
Standard Deviation						**1.86**	**1.08**	**1.08**

Table 6. Energy fitting results from the check set. The experimental energies (E_{exp}) were corrected (ΔE_{corr}) from their tabulated values to reflect a free-atom standard state. The experimental energies are then compared to the calculated energies when monopole (E_{mono}) and dipole (E_{di}) terms were included, to calculate the misfit values (ΔE_{mono} and ΔE_{di}). The misfit values were then normalized to the number of unique atoms (N_{unique}).

Molecule/ Crystal	ΔE_{corr}	E_{exp}	E_{mono}	E_{di}	ΔE_{mono}	ΔE_{di}	N_{unique}	$\Delta E_{mono}/N_{unique}$	$\Delta E_{di}/N_{unique}$
$AlO_{(g)}$	520.804	40.803	-363.98	-409.96	-404.78	-450.76	2	-202.39	-225.38
$AlOO_{(g)}$	752.541	-91.677	-984.63	-1006.08	-892.96	-914.40	3	-297.65	-304.80
$H_3O^+_{(g)}$	329.303	198.125	-717.79	-759.64	-915.91	-957.76	3	-305.30	-319.25
$HOO_{(g)}$	666.749	14.431	30.20	26.49	15.77	12.06	4	3.94	3.01
St. Dev.					**444.5**	**455.2**		**144.0**	**148.9**
β-Quartz	869.001	-856.153	-826.00	-824.99	30.15	31.17	4	7.54	7.79
Boehmite	955.817	-915.900	-922.21	-392.41	-6.31	523.49	3	-2.10	174.50
Kaolinite	4287.926	-3797.500	-3778.20	-3870.40	19.30	-72.90	17	1.14	-4.29
Moganite	869.001	-851.300	-860.03	-861.27	-8.73	-9.97	3	-2.91	-3.32
Seifertite	869.001	-794.000	-808.50	-809.18	-14.50	-15.18	3	-4.83	-5.06
Trydmite	869.001	-853.800	-848.59	-848.95	5.21	4.85	3	1.74	1.62
Bayerite	1594.106	-1153.000	-1238.28	-1228.54	-85.28	-75.54	7	-12.18	-10.79
St. Dev.					**37.4**	**210.2**		**6.1**	**67.1**

This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2016-5629

Table 7. D_0 vs. s_{ij} parameters for the reduced calibration set.

Atom 1	Atom 2	*a*	*b*
Al	Al	205.8767738	-52.59772999
H	Al	154.231069	-21.10143446
Si	Al	100.76	-17.598
H	H	1422.742954	-1919.768723
'H	Si	144.0735109	0
Si	Si	254.0441443	-53.80073857

Table 8. Energy fitting results from the reduced calibration set. The experimental energies (E_{exp}) were corrected (ΔE_{corr}) from their tabulated values to reflect a free-atom standard state. The experimental energies are then compared to the calculated energies when monopole (E_{mono}) terms were included, to calculate the misfit values (ΔE_{mono}).

Crystal/Molecule	ΔE_{corr}	E_{exp}	E_{mono}	ΔE_{mono}
$Al_{(g)}$	289.068	0	0.000	0.000
$Al_{2(g)}$	578.135	434.251	434.251	0.000
$AlH_{3(g)}$	898.896	96.619	96.619	0.000
$Si_{2(g)}$	811.056	532.654	532.647	-0.007
$Si_2H_{6(g)}$	2030.713	105.7365	99.597	-6.139
$SiH_{4(g)}$	1218.633	56.827	66.047	9.220
$Si_{(g)}$	405.528	0	3.054	3.054
St. Dev.				**4.265**

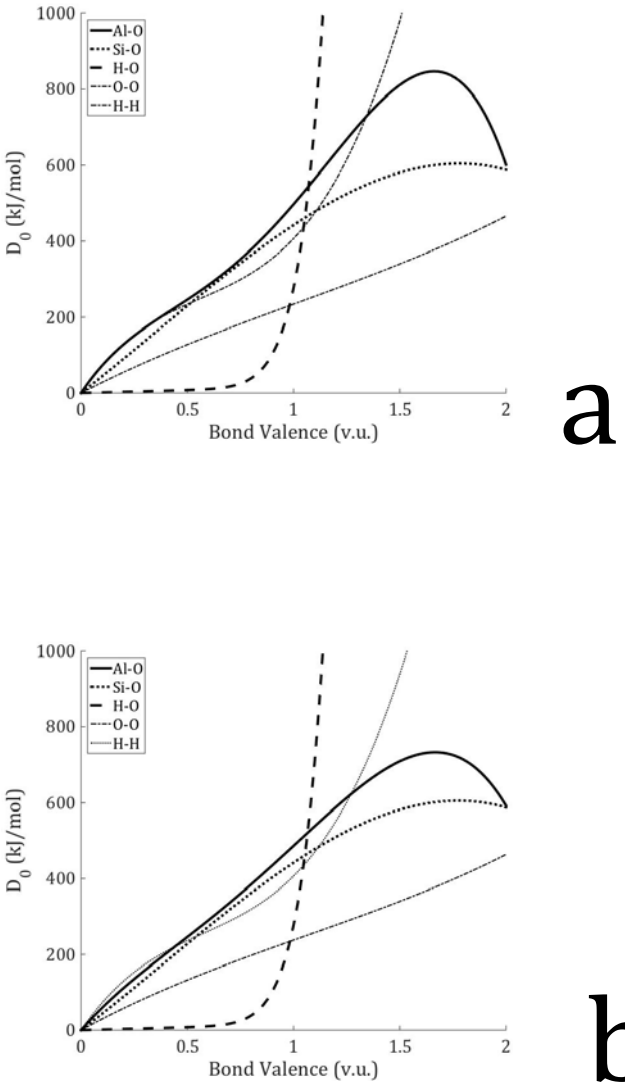
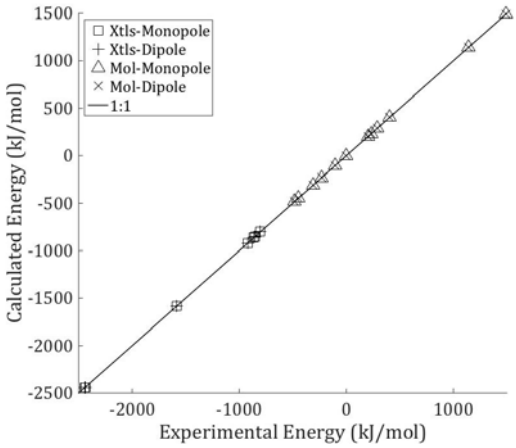
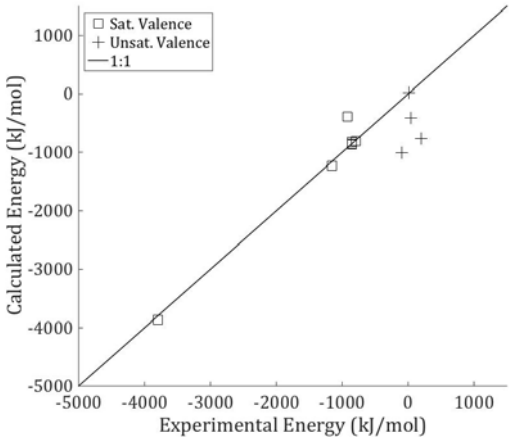


Figure 1. D_0 vs. s_{ij} curves obtained from a) the monopole-only optimization, and b) the dipole optimization. The curves show clear periodicity, with ionic bonds rising

more sharply at longer distances (lower bond valence), and multiple bonds favored for covalent bonds.



a

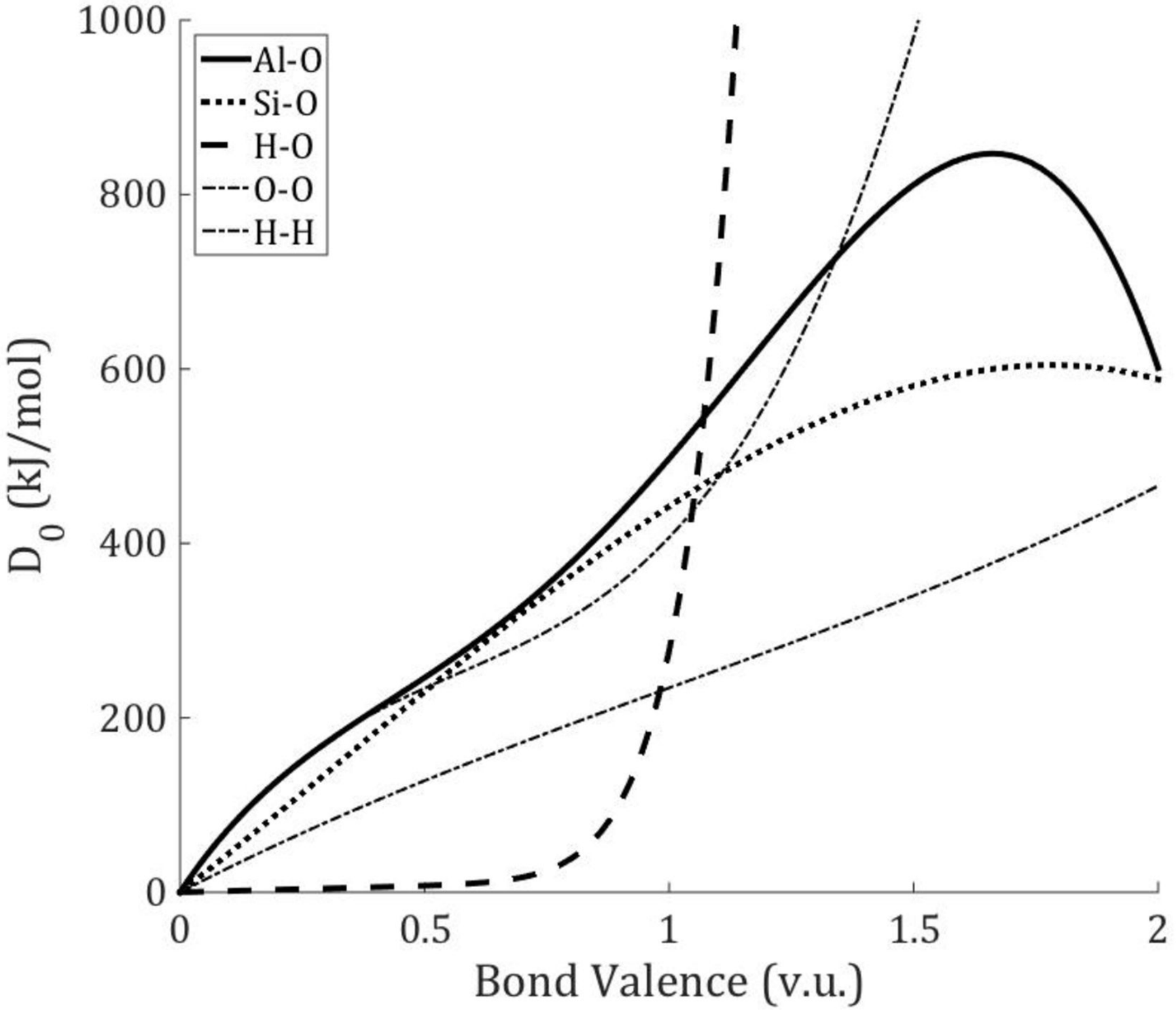


b

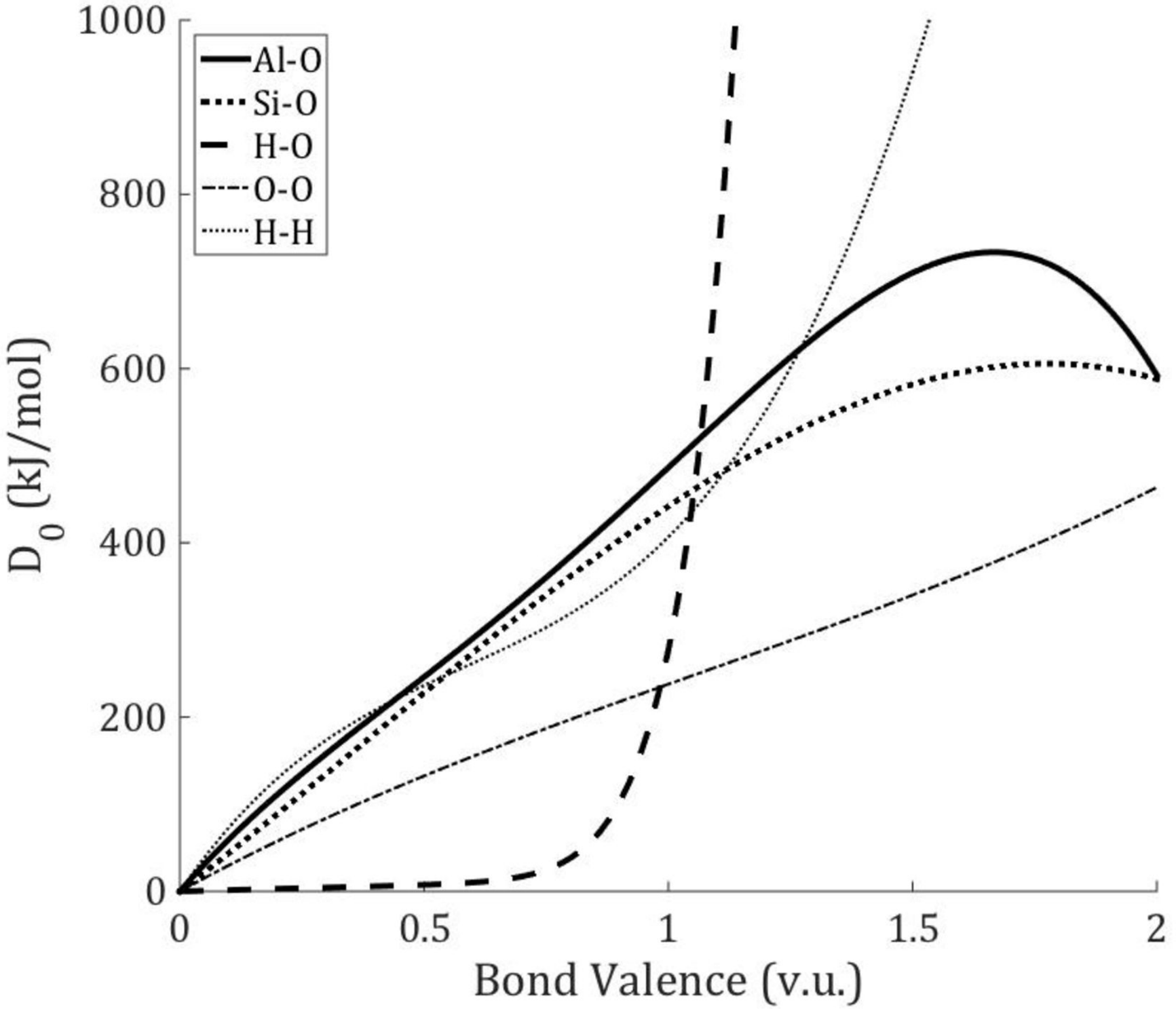
Figure 2. Comparison between experimental energy and calculated energy in the a) tempered calibration set and b) the check set. The symbols for the calibration set differentiate between crystals and molecules, and show the results for both the monopole-only optimization and the dipole optimization. The symbols for the check

set differentiate between species with fully saturated atomic valences, and those with unsaturated valences. The quality of the results is on par with the best quantum mechanical approaches, and exceeds those currently available for crystals.

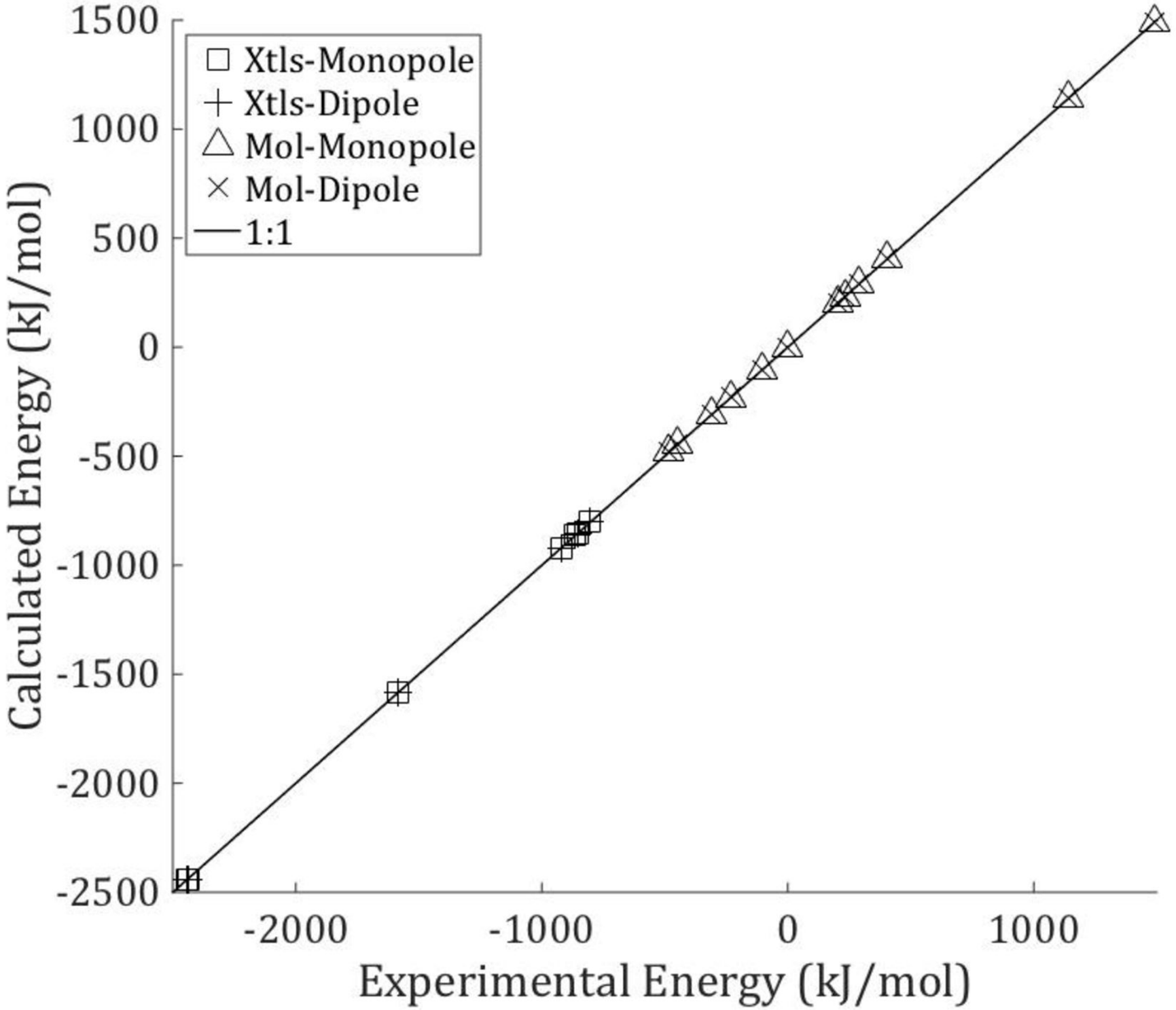
This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2016-5629



This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2016-5629



This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2016-5629



This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2016-5629