

Parameterization of a reactive force field using a Monte Carlo algorithm

Citation for published version (APA):

lype, E., Hutter, M., Jansen, A. P. J., Nedea, S. V., & Rindt, C. C. M. (2013). Parameterization of a reactive force field using a Monte Carlo algorithm. *Journal of Computational Chemistry*, 34(13), 1143-1154. DOI: 10.1002/jcc.23246

DOI:

[10.1002/jcc.23246](https://doi.org/10.1002/jcc.23246)

Document status and date:

Published: 01/01/2013

Document Version:

Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Parameterization of a Reactive Force Field using a Monte Carlo algorithm

Iype E*, Hütter M*, Jansen A P J[†], Nedea S V*, Rindt C C M^{*‡}

November 6, 2015

Abstract

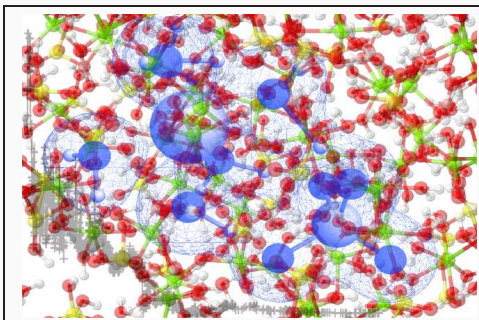
Parameterization of a Molecular Dynamics force field is essential in realistically modelling the physico-chemical processes involved in a molecular system. This step is often challenging when the equations involved in describing the force field are complicated as well as when the parameters are mostly empirical. ReaxFF is one such reactive force field which uses hundreds of parameters to describe the interactions between atoms. The optimization of the parameters in ReaxFF is done such that the the properties predicted by ReaxFF matches with a set of quantum chemical or experimental data. Usually, the optimization of the parameters is done by an inefficient single parameter parabolic-search algorithm. In this study, we use a robust Metropolis Monte-Carlo algorithm with Simulated Annealing (MMC-SA) to search for the optimum parameters for the ReaxFF force field in a high-dimensional parameter space. The optimization is done against a set of quantum chemical data for $MgSO_4$ hydrates. The optimized force field reproduced the chemical structures, the Equations of State and the water binding curves of $MgSO_4$ hydrates. The transferability test of the ReaxFF force field shows the extend of transferability for a particular molecular system. This study points out that the ReaxFF force field is not indefinitely transferable.

Keywords: ReaxFF, Optimization, Metropolis Monte Carlo, Force Field, Molecular Dynamics ■

*Department of Mechanical Engineering, Eindhoven University of Technology, Den Dolech 2, 5612 AZ Eindhoven, The Netherlands

[†]Department of Chemistry, Eindhoven University of Technology, Den Dolech 2, 5612 AZ Eindhoven, The Netherlands

[‡]c.c.m.rindt@tue.nl



Introduction

The ReaxFF force field is a new generation Molecular Dynamics(MD) force field for modelling reactive chemical systems¹. The dynamic chemical bonds in a reactive system are modelled based on the physically intuitive concept of bond order of a chemical bond. The bond order in a ReaxFF force field can take up any real value between zero and three depending on the nature of the chemical system and the distance between the atom pairs. All the connected interactions between atom pairs are included as a function of bond order and thus act as perturbation terms to the non-bonded interactions (van der Waals and Coulombic interaction). The force field parameters are optimized against quantum chemical data (structures and energies) of molecular systems¹. ReaxFF has been used to model a variety of systems such as hydrocarbons^{2,3}, Si-oxides⁴, Mg-hydride⁵, Lithium⁶, Platinum-Oxygen⁷, Iron-oxyhydroxides⁸, Gold oxides and nanoparticles^{9,10}, Cobalt¹¹, Cu-water^{12,13} etc. As such, the ability of the ReaxFF force field to accurately represent the interactions between atoms in such complex molecular systems, in terms of approximately 100 parameters per atom, is commendable.

In the present study the ReaxFF method is used to simulate *Mg*-salt hydrates which are potential candidate materials for energy storage in seasonal heat storage systems¹⁴. Magnesium sulfate hydrates can store energy of up to 2.8 GJ/m^3 ¹⁵ as heat of hydration. In addition, these hydrates are widely available and non-toxic, which make them suitable candidate materials for seasonal solar energy storage in the built environment. However, the slow hydration rate of magnesium sulfate limits the power output in such an energy storage system. In addition, the formation of cracks and pores in the material reduces the strength and re-usability of this material. Thus, we intent to model this system using ReaxFF to study the hydration and dehydration processes from a molecular point of view.

The ReaxFF force field for magnesium sulfate hydrates is parameterized using quantum chemical (Density Functional Theory) data in order to accurately simulate the dynamics of the system. The force field contains around 100 parameters per atom to describe the potential energy surface. The traditional parameterization procedure implemented in ReaxFF is inefficient since it uses a single parameter search algorithm¹⁶. The success of such a proce-

ture is highly dependent on the initial conditions as it searches only for the local minima in the parameter space. Alternatively, the Monte Carlo algorithms are well known for efficiently sampling highly irregular configurational spaces. In the present study, the Metropolis Monte Carlo Method¹⁷ is employed to randomly sample the parameter space in search for the global minimum, by utilizing the Simulated Annealing procedure of Kirkpatrick et.al.¹⁸ to optimize the ReaxFF force field. Also the transferability of the resulting force field is tested with respect to a part of the data set which was not used for optimization. In this study, a brief background on ReaxFF method is presented along with the traditionally used single parameter parabolic-search parameter optimization scheme. Then the Metropolis Monte-Carlo optimization scheme along with Simulated annealing technique for the force field parameterization is introduced. The quality of the resulting force field is analyzed by comparing the results with a set of quantum chemical data. In addition, a transferability test of the force field is also presented in this work.

Background

Molecular Dynamics (MD) simulations are increasingly being used to study complex molecular systems as the computational costs are getting lower day by day. In MD, the molecular system undergoes a systematic sampling of the configurational space defined by a particular ensemble, to generate statistical estimation of macroscopic properties including the dynamic properties. The inter atomic interaction potentials subsume all the physical properties of the system within itself and therefore play a crucial role in reproducing the real physical behavior of any system. There are several inter atomic potentials being used in MD simulations, each varying in their computational efficiencies and accuracies. One commonly used model is the Lennard-Jones (LJ) potential¹⁹. The LJ potential describes the van der Waals interactions between two atoms using a fairly simple model, mostly suitable for systems similar to inert gases, and sometimes for solid fluid interactions²⁰. There are a few reactive potentials (force fields) also available in order to simulate chemical reactions. Reactive Empirical Bond Order (REBO)²¹ is one such model which was built based on the Tersoff potential²². Attempts have been made to improve REBO force fields²³. The development of the AIREBO

potential (Adapted intermolecular REBO) is an attempt by Stuart et. al.²⁴ to generalize REBO to include long range interactions. However, it retained the fundamental problems in the shapes of the dissociation curves of REBO. Electronic structure based methods are also available such as the Embedded Atom Method (EAM)²⁵ and Car-Parinello Molecular Dynamics (CPMD)²⁶. However, these methods have the disadvantage that either their usage is limited to specific systems or they are computationally expensive. ReaxFF¹ is an empirical force field which emulates chemical reactions by introducing dynamic bonds in addition to the non-bonded long range interactions between atoms.

ReaxFF force field

The ReaxFF force field calculates the energy of the system according to Eq.1.

$$E_{system} = E_{bond} + E_{lp} + E_{over} + E_{under} + E_{val} + E_{pen} \\ + E_{3conj} + E_{tors} + E_{4conj} + E_{H-bond} + E_{vdWaal} + E_{Coulomb} \quad (1)$$

where E_{bond} represents the bond energy, E_{lp} is the energy due to the presence of lone pair and E_{over} and E_{under} are the energies arising from over and under-coordination, respectively, of atoms with respect to their valency. The valence angle (three body) energy and the torsion angle (four body) energies are accounted for in the terms E_{val} and E_{tors} respectively. E_{pen} is a penalty energy to stabilize a three body system with the center atom having two double bonds connected to it. Conjugated chemical bonds are stabilized by adding corrections for three body, E_{3conj} , and four body, E_{4conj} , conjugation terms to the respective systems. E_{H-bond} represents the hydrogen bond interactions. The long range interactions such as van der Waals interactions and Coulombic interactions are accounted for between every pair of atoms irrespective of the presence of chemical bonds. The complete expressions for each term are given by van Duin et. al.^{1,4}. The van der Waals interactions are modelled using a distance corrected Morse-potential, Eq.2a.

$$E_{vdWaal} = D_{ij} \cdot \left\{ \exp \left[\alpha_{ij} \cdot \left(1 - \frac{f_{13}(r_{ij})}{r_{vdW}} \right) \right] - 2 \cdot \exp \left[\frac{1}{2} \cdot \alpha_{ij} \cdot \left(1 - \frac{f_{13}(r_{ij})}{r_{vdW}} \right) \right] \right\}, \quad (2a)$$

$$f_{13}(r_{ij}) = \left[r_{ij}^{\lambda_{29}} + \left(\frac{1}{\lambda_w} \right)^{\lambda_{28}} \right]^{\frac{1}{\lambda_{28}}} \quad (2b)$$

Here D_{ij} is the interaction energy, r_{vdw} is the van der Waals radius and the remaining ones are empirical parameters in the force field. Eq.2b is a shielding term to limit excessive repulsive interaction between bonded atoms and atoms sharing a valence angle. Similarly, a shielded Coulomb potential as presented in Eq.3 is used to account for electro-static interactions. Corrections will be made whenever there is orbital overlap between atoms at close distances. Atomic charges are calculated using the Electron Equilibrium Method (EEM)²⁷, thus taking into account charge transfer and polarization of the molecules during chemical reactions.

$$E_{coulomb} = C \cdot \frac{q_i \cdot q_j}{[r_{ij}^3 + (1/\gamma_{ij})^3]^{1/3}}, \quad (3)$$

$$BO'_{ij} = \exp \left[P_{bo1} \cdot \left(\frac{r_{ij}}{r_o^\sigma} \right)^{P_{bo2}} \right] + \exp \left[P_{bo3} \cdot \left(\frac{r_{ij}}{r_o^\pi} \right)^{P_{bo4}} \right] \\ + \exp \left[P_{bo5} \cdot \left(\frac{r_{ij}}{r_o^{\pi\pi}} \right)^{P_{bo6}} \right] \quad (4)$$

In ReaxFF, bond order, BO'_{ij} , is calculated between every pair of atoms according to Eq.4. The parameters P_{bo1} , P_{bo3} and P_{bo5} in Eq.4 have values less than zero. The bond order, thus, takes a real value between 0 and 3 depending on the distance between atoms and the nature of the chemical bond. All the connected interactions such as bond energy, valence angle etc are made bond order dependent, thus making the bond breaking and formation a spontaneous process. All the non-bonded interactions are multiplied by a 7th order polynomial to make sure that the 1st, 2nd and 3rd derivatives are continuous and go to zero at the cutoff boundary².

Force field optimization using quantum chemical results

A typical ReaxFF force field uses approximately 100 parameters per atom to define the potential energy surface¹. Although some of these parameters may be given physical interpretations, many of them are purely empirical. Optimization of these parameters for a representative chemical system is done in such a way that the force field reproduces a set of

reliable energies and structural properties of various configurations of the system. The set of data which is used for comparison is called the training set, which include atomic charges (Mulliken charges), bond lengths, bond angles, torsion angles, potential energy changes, heat of formations etc. Such data are obtained mostly from Density Functional Theory (DFT) calculations or in some situations from experiments. The energies used for comparison in the training set are not provided as absolute energies of the geometries, instead as differences in energies between any two selected geometries, considering the fact that the reference energy states for the DFT methods are mostly different from the ReaxFF reference state. Therefore, the optimization is performed mostly for the shape of the potential energy surface and not for its absolute values. For an MD simulation, nevertheless, absolute values of energies are not relevant, only the shape of the energy surface is important.

During the force field parameter optimization, ReaxFF tries to minimize an objective function, indicated as *Error* given in Eq.5, which is the sum of the squared error between the force field obtained data and the one which is provided in the training set,

$$Error = \sum_{i=1}^n \left[\frac{(x_{i,TS} - x_{i,ReaxFF})}{\sigma_i} \right]^2 \quad (5)$$

Here, $x_{i,TS}$ is the quantum chemical or experimental data in the training set for the atomic charges or the structural parameters or the energies and $x_{i,ReaxFF}$ is the value for the corresponding data obtained from ReaxFF. The σ_i values are parameters used to scale different types of errors generated for each entry in the training set to mutually comparable magnitudes. Therefore these parameters can be used to set the relative importance between various sections within the training set. The choice of the values for these scaling parameters, σ_i , depends mostly on the relative accuracy of various data in the training set. For example, the Density Functional Theory method using various GGA functionals with $3 - 21 + G^*$ basis set is reported to have an average error in the heat of formation estimation of several protein molecules, between 8 to 30 kcal/mol²⁸, compared to the experimental values. In the case of $MgSO_4$ crystal, containing 24 atoms, we have observed a deviation in the heat of formation of approximately 25 kcal/mol (= 0.045 eV/atom) using the PW91-functional used in this study. In addition, the accuracies of the energies calculated using various functionals and the basis sets in DFT can be different from each other. Such aspects must be taken into

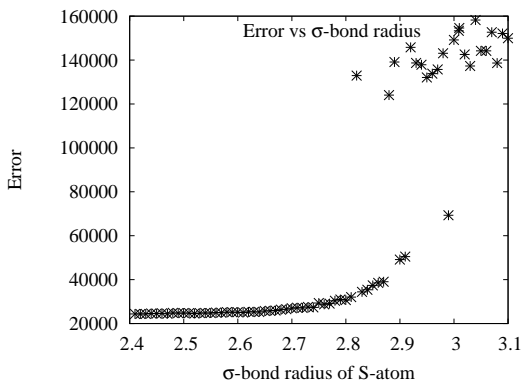
account in choosing the σ_i -s. Besides, only the relative values of the σ_i have any significance and scaling all of them uniformly will not make any difference to the optimization scheme except that it will merely change the magnitude of *Error*. The approximate values of σ_i used in the calculation for various sections of the training set are given in Table 1. The σ_i values for the energies vary from 0.1 to 1.0 depending on the magnitude of the training set entries $x_{i,TS}$. Altogether, there are approximately 3000 data points in the training set.

Table 1: The approximate values of σ_i used for various sections in the training set.

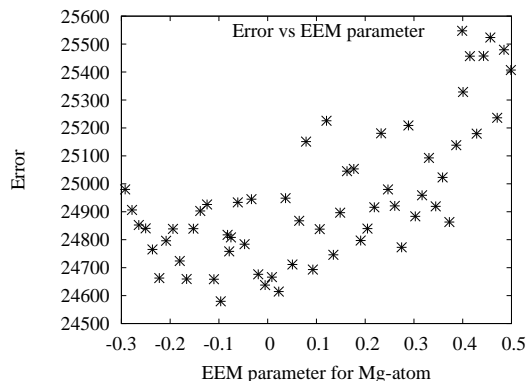
Section in the training set	Approximate σ_i values
Atomic Charges	0.01 (C)
Bond lengths	0.05 (Å)
Valence Angles	3.0 (°)
Energies	0.1 to 1.0 (kcal/mol)

Usually, the optimization of the parameters is done using a successive one-parameter search algorithm¹⁶. According to this method, one particular parameter p in the force field is chosen at a time, and then a parabolic fit is made between three values of the objective function against three values for that parameter: one being the initial value p , and the other two being small perturbations δp in both directions $p \pm \delta p$. The optimum value for this parameter p is then calculated by minimizing this parabola. If the parabola is concave down, the value which corresponds to the smallest error will be selected. The procedure is then repeated by choosing a different parameter and is continued until the error no longer decreases. However, since most of the parameters in the force field are in some way related¹⁶, the procedure has to be repeated several rounds over all the parameters in order to get a converged force field. Another drawback in this procedure is that it will only find a local minimum in the parameter space and a good initial condition for the parameters is necessary for obtaining a good force field. Given the fact that most of the parameters are empirical, a good starting point for these parameters is, very often, difficult to realize. This approach becomes more difficult when the number of parameters to be optimized is large. In addition, there is no guarantee that the error surface has the shape of a parabola in the region of

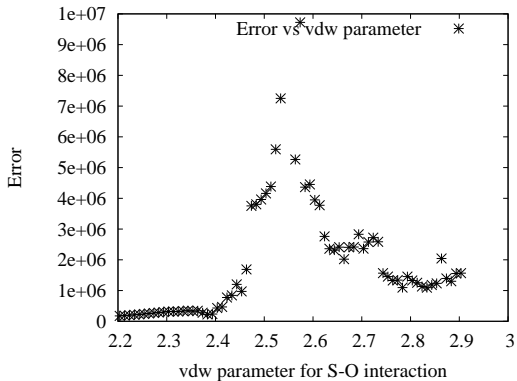
interest and, therefore, the assumption that this procedure will lead to a local minimum is questionable. For instance, Figure 1 shows the shapes of the error surface as functions of few parameters in the $MgSO_4$ -water ReaxFF force field. A continuous parabolic shape for the error surface is observed only for the van der Waals radius of Mg atom, Figure 1(d). The other three are either irregular or rapidly varying. The parabolic-search algorithm for such cases may not provide a good estimation for the parameters and is unsuitable for finding the global minimum in such an error surface, especially when one does not have a suitable initial condition. In this study, we propose to use a more robust optimization scheme to minimize the error function, $Error$ in Eq.5, in a high-dimensional parameter space.



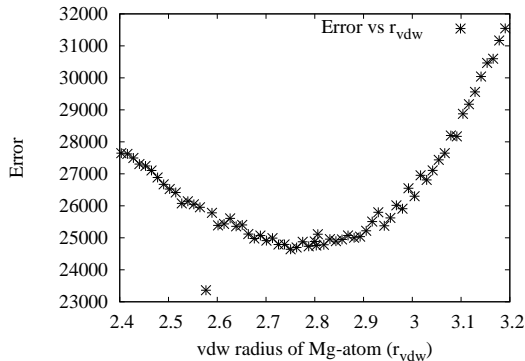
(a) Error vs the σ -bond covalent radius of S-atom



(b) Error vs the EEM parameter of the Mg-atom



(c) Error vs S-O van der Waals parameter



(d) Error vs Mg van der Waals radius

Figure 1: The shapes of a typical ReaxFF error surface for $MgSO_4.H_2O$ system as functions of a few parameters in the force field. All the other parameters were kept constant while sampling the errors. It shows the extent of the irregularity in the energy surface.

Method

A standard mathematical minimization procedure such as the least squares method requires that the error function must be continuous and differentiable in the parameter space. Besides, such a method may only converge to a local minimum depending on the starting point. Owing to the mathematical complexity of the potential¹⁶, chances are high that the accumulation of numerical differentiation errors might create instabilities in the solution procedure. Therefore, a simple and robust Metropolis Monte-Carlo minimization procedure is used to optimize the ReaxFF force field for a magnesium sulfate-water system in the high-dimensional parameter space.

Metropolis Monte-Carlo algorithm

The Metropolis Monte Carlo (MMC) algorithm¹⁷ has been used extensively in efficiently sampling the configurational space and to compute macroscopic properties of molecular systems. In this method, the particles/atoms are instilled with a random movement at every iteration and the so-proposed configuration is accepted with a probability given by

$$P_a = \min [1, \exp(-\beta\Delta E)] \quad (6)$$

where $\Delta E = E_{new} - E_{old}$ is the difference in energy of the configuration after, E_{new} , and before, E_{old} , the proposed move. The Metropolis scheme samples the states in accordance with the Boltzmann distribution. The acceptance of the trial move is thus governed by the Boltzmann distribution. The Boltzmann factor, $\exp(-\beta E)$ (where $\beta = \frac{1}{k_B T}$), assigns a probability for each trial state according to its energy, E . The states with low energy are more probable compared to the states with high energy. When the temperature, T , of the system is high, the system meanders around the energy surface more freely and vice versa. The stochastic nature of the dynamics in the MMC algorithm makes it a robust method to search over an energy hyper-surface with a large number of minima²⁹. The applicability of

the MMC algorithm for the purpose of molecular docking has been successfully established for molecules of varying sizes, degrees of flexibility and types of interactions³⁰.

Metropolis algorithm for ReaxFF force field optimization with simulated annealing

There are optimization methods developed based on the Monte-Carlo algorithm, which includes the combinations such as MMC and the Lattice Statics Method³¹, Monte Carlo minimization and Thermalization³² etc. One commonly used variant of the MMC algorithm for minimization problems is the Metropolis Monte Carlo-Simulated Annealing (MMC-SA) procedure¹⁸. In this method, the simulation is started at a high temperature with a set of randomly generated values for the parameters to be optimized and then the system is slowly cooled down. So, in the initial stage, the system can freely move around the energy surface and one may assert that, eventually, the system will boil down to the global minimum in the energy surface. The procedure has been applied for many systems to find the global energy minimum.³³⁻³⁵ It is to be noted here that the particles in our system are nothing but the parameters in the ReaxFF force field, and the energy E in Eq.6 is the *Error* in Eq. 5. So Eq.6 may be rewritten as

$$P_a = \min [1, \exp(-\beta \Delta Error)] \quad (7)$$

where $\Delta Error = Error_{new} - Error_{old}$. The temperature T determines the number of accessible error levels in our error surface, and it may be viewed differently from the actual temperature in particle mechanics. The order of T has been chosen based on the magnitude of *Error* for a particular optimization run, so that to maintain a certain minimum acceptance rate. At the beginning of the simulation the temperature is set to a high value and is gradually decreased by a factor of 0.999 at every proposed move. Out of a selected 132 parameters of $Mg - S - H - O$ ReaxFF force field (which include the pair interactions between $Mg - O$ and $S - O$, valence angles between $Mg - O - S$, $O - S - O$, $H - O - S$, van der Waals interaction between most pairs and the $O - H - O$ hydrogen bond interactions apart from the atomic parameters of Mg and S) for optimization, only 5% of them were randomly picked at each move and random propositions are made for each according to $p_i = p_i + \delta p_{i,max} \times (2 \times rand(0, 1) - 1)$. Here $\delta p_{i,max}$ is the maximum allowed movement

for each parameter and $rand(0, 1)$ is a random number from a uniform distribution in the interval $[0, 1]$. Throughout the calculation, a minimum acceptance rate of 30% is maintained by scaling $\delta p_{i,max}$ by a factor of 1.03 (i.e. when the acceptance rate is more than 30%, $\delta p_{i,max}$ is scaled up by 3% and when it is less than 30%, $\delta p_{i,max}$ is scaled down by 3%). Figure 2 shows the decrease of $Error$ as the system is cooled down.

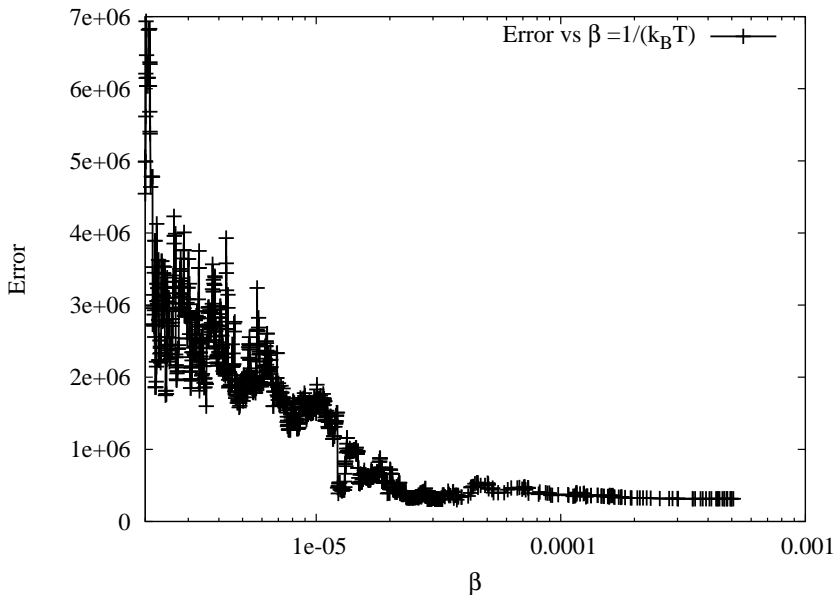


Figure 2: An example for the decreasing $Error$ (Eq.5) with increasing β ($= \frac{1}{k_B T}$) taken from one of the MMC-SA run for the ReaxFF force field optimization.

In the initial stages, relatively large fluctuations were allowed due to high temperature, and in the final stages only small fluctuations were allowed. Each parameters in the force field is given an upper and lower bound, based on the values used in a set of eleven force fields provided in the ReaxFF package of scm³⁶, to ensure that the parameters do not attain any physically meaningless values. These bounds for the parameters are large enough so that it includes the region where force fields for a number of elements (such as H,O,C,N,S,F,Pt,Cl,Ni,Si,V,Zn etc.) and for the the molecules formed by their combinations can be derived out of it. Roughly 25% of the data in the training set is rendered ineffective from the optimization procedure, and is used to test the transferability by cross-validating³⁷

the force field after optimization. A low weighting factor, $(100 \times \sigma_i)$, refer Eq.5, is set for these data compared to the other equivalent data in the training set. In addition, the equations of state for magnesium sulfate penta-hydrate crystal is left out of the training set, to test the transferability of the force field for data points completely outside what is given in the training set. In order to compare the influence of the starting point between the MMC-SA and the parabolic-search algorithm, the optimization is also done using the latter method with random starting points for the selected set of parameters. Multiple sweeps were performed over all the parameters during the parabolic search method.

Results

Usually, the ReaxFF force fields are optimized against quantum chemical data such as the equations of states, binding energies for chemical reactions, specific reaction path ways, dissociation curves etc^{3-5,12} for the respective systems. For optimizing the force field for $MgSO_4 \cdot xH_2O$ system, we have included the DFT optimized structures of various hydrates of $MgSO_4 \cdot xH_2O$, where $x = \{0, 1, \dots, 6\}$, in the training set. The equations of state for $MgSO_4$, $MgSO_4 \cdot 4H_2O$, $MgSO_4 \cdot 7H_2O$ -periodic crystals as a function of volume, were also added. In addition, the binding energy curves of the water molecules to the $MgSO_4$ (100) surface and various proton transfer pathways observed in the DFT optimizations were also included. The DFT calculations are done in ADF (Amsterdam Density Functional) package³⁶ installed in SARA super computers (NCF-Amsterdam)³⁸, using the PW91-Generalized Gradient Approximation (GGA) exchange correlation functional.

Comparison between Parabolic-search and MMC algorithm

In order to assess the performance of Metropolis Monte Carlo scheme to optimize the ReaxFF force field, a comparison is made between the single parameter parabolic-search (ParSearch) method and the MMC algorithm. Since a proper initial condition was not known for our system, we have used five random starting force fields. These five random force fields were then optimized using the MMC method at a low temperature as well as the parabolic-search method. The ratio of $Error_{MMC}$ to $Error_{ParSearch}$ as a function of iteration is shown in

Figure 3.

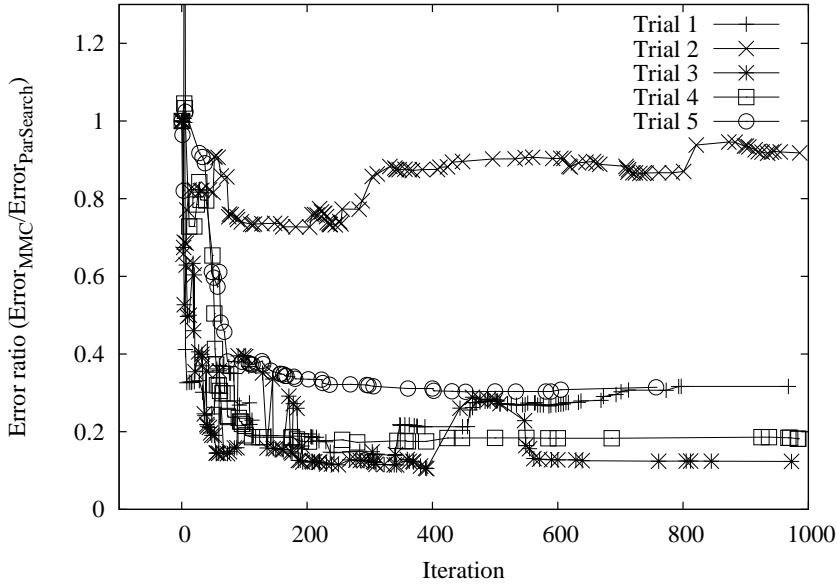


Figure 3: Comparisons between the Parabolic-search method and MMC algorithm in optimizing a ReaxFF force field. The error ratio($Error_{MMC}/Error_{ParSearch}$) as a function of iteration for five different random starting force fields.

All the curves in the figure starts at unity. A value of less than unity for this ratio implies that the $Error_{MMC}$ is lower than $Error_{ParSearch}$ at that iteration and vice versa. A value of greater than unity is found for the beginning of just one simulation and then all the curves lies below unity throughout. In all the five trials, MMC method finds a lower minimum compared to the parabolic-search method. Apart from the the second trial, the error ratios between the MMC and the parabolic-search mostly remain significantly lower than unity, suggesting that the MMC method works better when starting from a random initial force field. In the second trial, we assume that the parabolic-search method had a good initial condition. Nevertheless, MMC method still arrived at a better force field with lower error in this case as well. Figure 4 shows the absolute values of the errors in the MMC and the

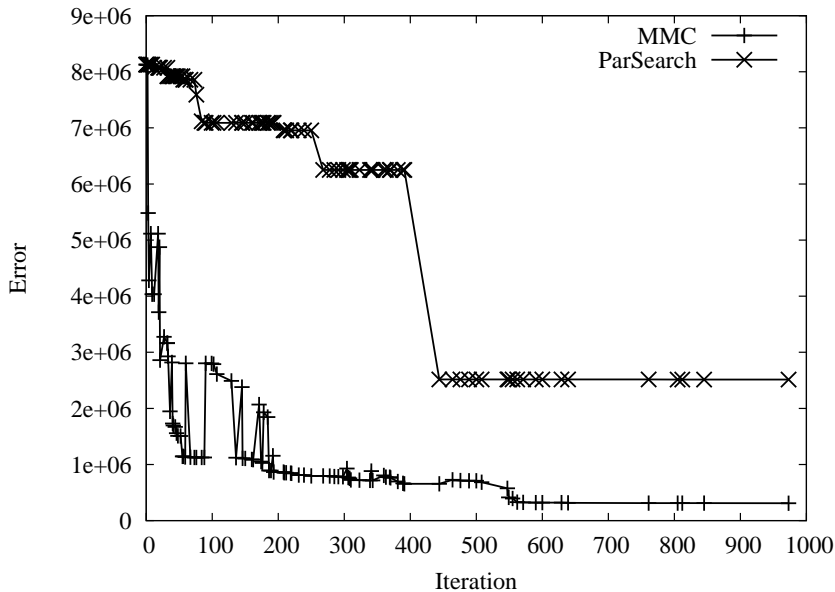


Figure 4: Error vs iteration for an MMC run as well as the parabolic-search method. The figure shows how fast the MMC method reduces the error in the force field.

parabolic-search method for trial 3. This shows that the MMC method reduces the error in the force field faster than the parabolic-search algorithm.

In addition, the initial and the final errors for all these five sets of calculations are given in Table 2. A careful look at the five final errors from the MMC method reveals that all the five trials led to force fields with comparable magnitude in its errors. The ratio of the standard deviation to the average error for the parabolic-search method is 0.6 and the same for MMC method is 0.21 suggesting that spread of the final errors in MMC is much lower than that of the parabolic-search method. This shows the capability of the MMC method to arrive at the global minimum from any random starting point in the parameter space. These results also show the advantage of using MMC algorithm to find the optimum force field for ReaxFF, especially when one does not have a good initial condition. It appears from the results in Figure 3 and from Table 2 that the parabolic-search method often results in a local minimum and this must be avoided in order to get the best force field.

Thus, we argue that the MMC algorithm with Simulated Annealing technique must yield the optimum force field for such a minimization procedure. Now, since the success of the parabolic-search method is dependent on the initial condition, we agree that an extensive

Table 2: Initial and final errors of five simulations shown in Figure 3. The average $\langle Error \rangle$ and the ratio of standard deviations σ of the final errors to $\langle Error \rangle$ from all the trials are also given.

Trial	$Error_{initial} \times 1E - 6$	$Error_{final} \times 1E - 6$	
		ParSearch	MMC
1	4.3	0.9	0.31
2	1.1	0.4	0.44
3	8.1	2.5	0.31
4	2.0	1.4	0.26
5	4.3	1.3	0.42
$\langle Error \rangle$		1.3	0.35
$\frac{std(Error)}{\langle Error \rangle}$		0.6	0.21

comparison between the MMC-SA algorithm and the parabolic-search algorithm may not be appropriate. However, we have shown a limited extent of comparison between the best force fields obtained by both the methods in the following sections.

Energies of $MgSO_4$ hydrates molecules

The ReaxFF optimized structures of $MgSO_4$ hydrates and their (relative) energies after the parameterization are given in Figure 5. The energies are compared with the absolute energies obtained from DFT calculations (QM). The degree of hydration of the molecules increases from left to right in the figure.

The optimized structures shown in Figure 5 are in excellent agreement with the DFT optimized structures. Since the reference states for ReaxFF and DFT are different, all the ReaxFF energies are shifted by a constant value equal to the difference in energy between ReaxFF and DFT of the 1st geometry, $MgSO_4$. This data point, from now on we call the calibration point, is indicated by an arrow in Figures 5 and 7-12. As it can be seen, the

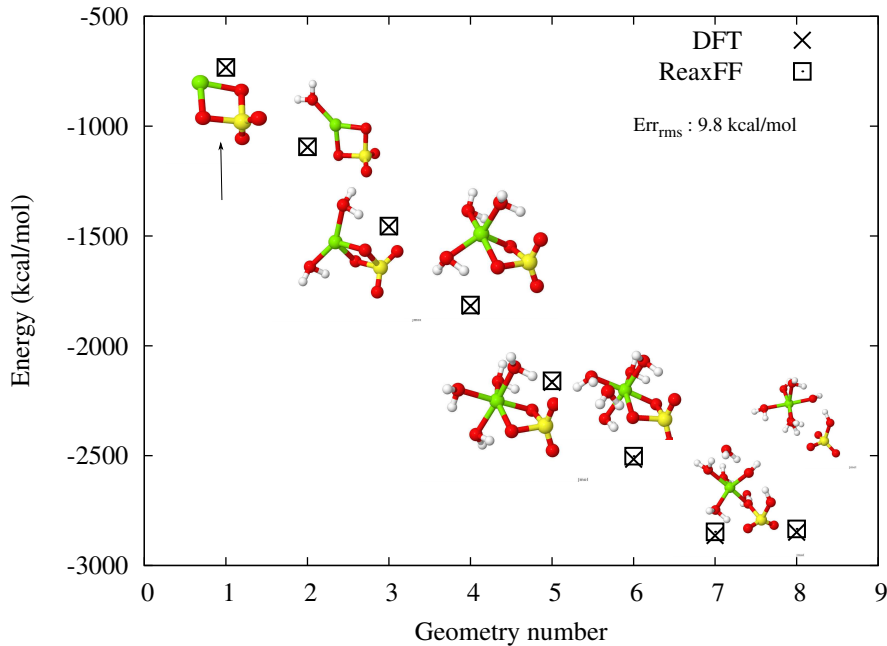


Figure 5: Energies of hydrated molecules of $MgSO_4$ from the optimized ReaxFF force field and the DFT results. The arrow shows the calibration point to correct for the reference state.

agreement in the energies of the molecular structures is good. The energies exhibit an rms error of 9.8 kcal/mol. The force field after the single parameter search method had an rms error of 30 kcal/mol for the molecular structures of the hydrates of $MgSO_4$. Geometry number 7 and 8 are the two isomers of the $MgSO_4 \cdot 6H_2O$ molecule.

The optimized structures of the $MgSO_4 \cdot 6H_2O$ molecule (geometry 7 in Figure 5) both from ReaxFF as well as from DFT are shown in Figure 6 for comparison. The individual bond lengths are also shown in the figure. Although there is significant deviation in one of the $S - O$ bond lengths, the average bond lengths of $Mg - O$ pairs and $S - O$ pairs in the $MgSO_4 \cdot 6H_2O$ molecule, which are shown in Table 3, show good agreement. This agreement in the average bond length is an evidence for the preservation of the coordination around the S atom during the structural change. An important fact to be noticed is the spontaneous proton transfer from one of the co-ordinated water molecules to the SO_4 tetrahedra, in accordance with the DFT structures reported³⁹, as shown by dotted lines in Figure 6. It is worth mentioning that ReaxFF reproduces this spontaneous proton transfer correctly in

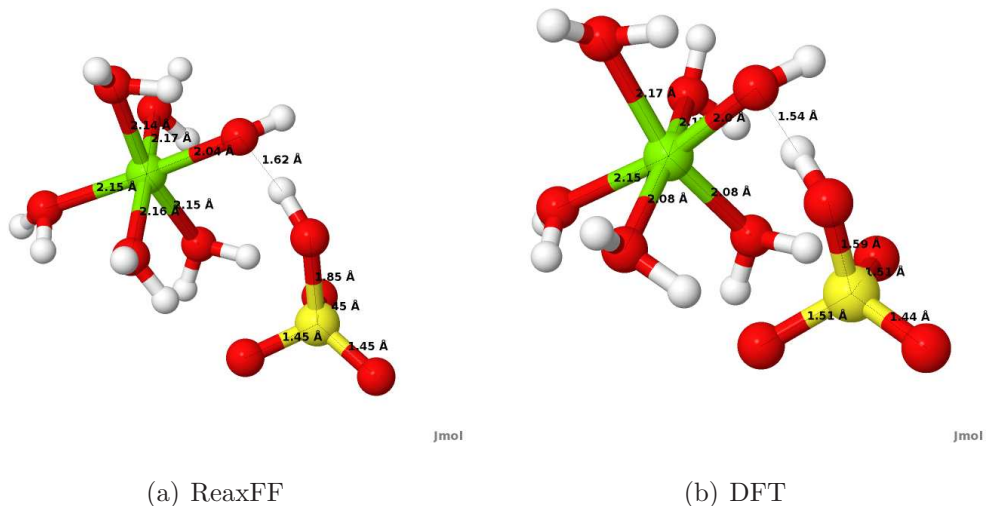


Figure 6: Optimized structures of the $MgSO_4.6H_2O$ molecule from the ReaxFF force field as well as from DFT. The optimized force field reproduced the proton transfer observed in the hexa-hydrate which was not observed before the Metropolis minimization.

both the isomers of $MgSO_4.6H_2O$. This effect was not noticed in the force field obtained after performing a parabolic-search optimization method with a random starting points for a selected set of parameters. The equilibrium hydrogen bond, $O \cdots H$, lengths after the proton transfer are 1.62 Å and 1.54 Å for ReaxFF and DFT, respectively.

Table 3: Average bond length comparison between the DFT and the ReaxFF optimized(both using MMC-SA and parabolic-search method) structures of the $MgSO_4.6H_2O$ molecule

Bonds	DFT	ReaxFF(MMC-SA)	ReaxFF (Parbolic-search)
$Mg - O$ bond (Å)	2.10	2.13	2.11
$S - O$ bond (Å)	1.51	1.55	1.48

Proton transfer reaction in $MgSO_4.6H_2O$ molecule

In order to reproduce the proton transfer in $MgSO_4.6H_2O$ using ReaxFF, the energies calculated along the proton transfer path using DFT were also added to the training-set before starting the optimization. The variation of the potential energy along the proton transfer

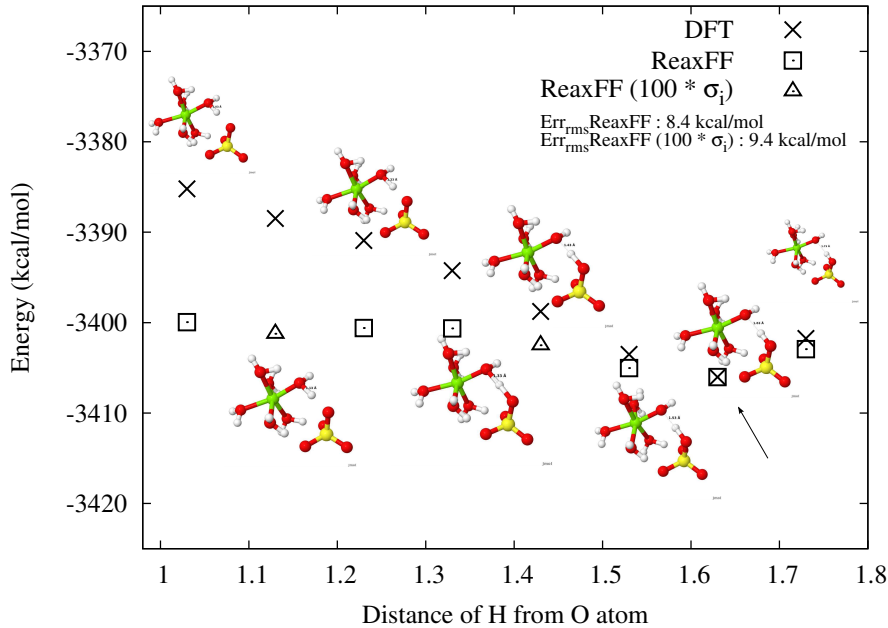


Figure 7: Proton transfer reaction observed in $MgSO_4 \cdot 6H_2O$ molecule is being reproduced with the ReaxFF force field. From left to right, the hydrogen atom moves from the $[Mg \cdot (H_2O)_6]^{2+}$ fragment to SO_4 tetrahedra. The arrow shows the calibration point to correct for the reference state.

path is shown in Figure 7. The hydrogen atom is placed at different positions along the proton transfer pathway and single point calculations are performed to sample the potential energy surface. The energy is decreasing from left to right. The DFT results confirm the absence of any barrier in the proton transfer process as reported in the literature³⁹. Now, it may be pointed out that the equilibrium $O \cdots H$ distance in the ReaxFF structure, Figure 6(a), is 0.08 \AA higher than the DFT optimized structure, Figure 6(b) given in the training set. The reason for this deviation is that the structure in Figure 6(b) is optimized using the PW91 functional and the proton transfer pathway in Figure 7 is sampled, by performing single point calculations, using the B3LYP functional. It is observed that the ReaxFF optimization resulted in reproducing the proton transfer pathway better compared to the structural properties of Figure 6(b). This is evident from the proton transfer pathway, Figure 7, where the DFT and ReaxFF have a minimum around 1.62 \AA and not at 1.54 \AA .

The agreement between the ReaxFF and the DFT calculated energies is fairly good with an rms deviation of less than 10 kcal/mol (≈ 27 kcal/mol before Metropolis minimization). Along with the optimized ReaxFF data points, the low-weightage data points (with large σ_i values) are also plotted. These data points appear to follow the trend of the remaining ReaxFF data points after optimization, and in this sense transferability is respected.

Equations of state for hydrated crystals of magnesium sulfate

Equations of state (EOS) are used to characterize the energy required to change the volume of a solid with respect to its equilibrium volume. A general expression for the equations of state can be derived using the expression for free energy, $dA = -SdT - PdV$. Differentiating A with respect to V at constant T thus results in

$$\left(\frac{\partial A}{\partial V}\right)_T = -P \quad (8)$$

DFT calculations assume the Born-Oppenheimer approximation⁴⁰ and thus the nuclear kinetic energy is zero (i.e $T = 0\text{K}$). This implies that the free energy $A = E$, the internal energy. Therefore, Eq.8 is rewritten as,

$$\left(\frac{\partial E}{\partial V}\right)_{T=0} = -P \quad (9)$$

Integrating Eq.9, we get the general expression for the equations of state,

$$E = E_0 - \int PdV \quad (10)$$

The pressure P is connected to the bulk modulus B via the equation, $B = -V \left(\frac{\partial P}{\partial V}\right)_T$. Given an expression for the bulk modulus of the material, it is thus possible to obtain an expression for energy E as function of volume V . There are several equations of state for solids proposed in the literature^{41,42}. Equations of state data, $E = E(V)$, for hydrated crystals of $MgSO_4$ (anhydrate, tetra-hydrate, penta-hydrate and hepta-hydrate), are generated using DFT in order to optimize the ReaxFF force field to model the various stages of hydration of the crystal. Atomic coordinates and the unit cell data for these hydrates are obtained from

experimental measurements^{43–46}. The unit cell dimensions and the atomic coordinates are further optimized at PW91-TZP level theory. A quadratic tetrahedron method is used for integrating over the Gamma-points in the first Brillouin zone. Owing to the large unit cells, sampling just the Gamma-points is found to be sufficient for the present systems.

Hydrated magnesium sulfate crystals

Crystalline magnesium sulfate realizes an orthorhombic unit cell with four $MgSO_4$ molecules occupied in it⁴³. The volume of the optimized geometry is uniformly varied (by keeping the ratio between the sides constant) changed gradually, and at every instance, the atomic coordinates are optimized to find the equations of state for the $MgSO_4$ crystal at 0K, Figure 8. Similarly, the equations of state for tetra-hydrate, Figure 10, penta-hydrate, Figure 12, and hepta-hydrate, Figure 9, are generated.

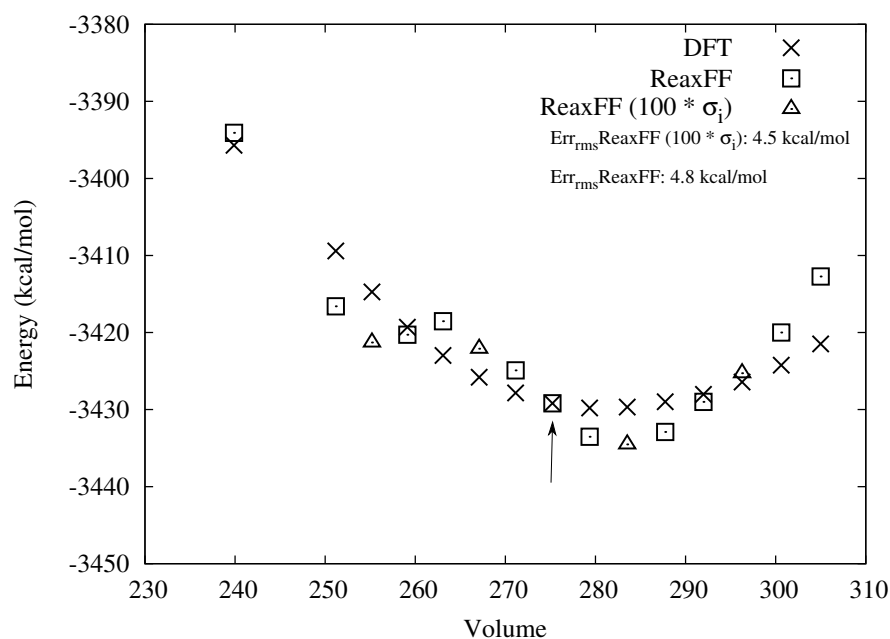


Figure 8: Equations of state for $MgSO_4$ unit cell. Results from the optimized force field are compared with DFT results. The low-weightage data points ($100 \times \sigma_i$) are also indicated. The arrow shows the calibration point to correct for the reference state.

The equations of state data are then included in the training set. As mentioned in the case of proton transfer reaction in magnesium sulfate hexa-hydrate, a set of randomly picked

data in the training set is given a 100 times lower weight ($100 \times \sigma$) compared to the other data in the training set. After optimization of the force field using the Metropolis algorithm, the resulting energies of the geometries are compared with the obtained DFT data in Figures 8-11.

The data for the equations of state for $MgSO_4$ is shown in Figure 8. The ReaxFF optimized data and the DFT results show good agreement between each other with an rms error of less than 5 kcal/mol for the data points which were used for optimization. The low-weightage data points also match well with the DFT results. Before the Metropolis minimization, the overall rms error for the EOS of $MgSO_4$ was approximately 70 kcal/mol.

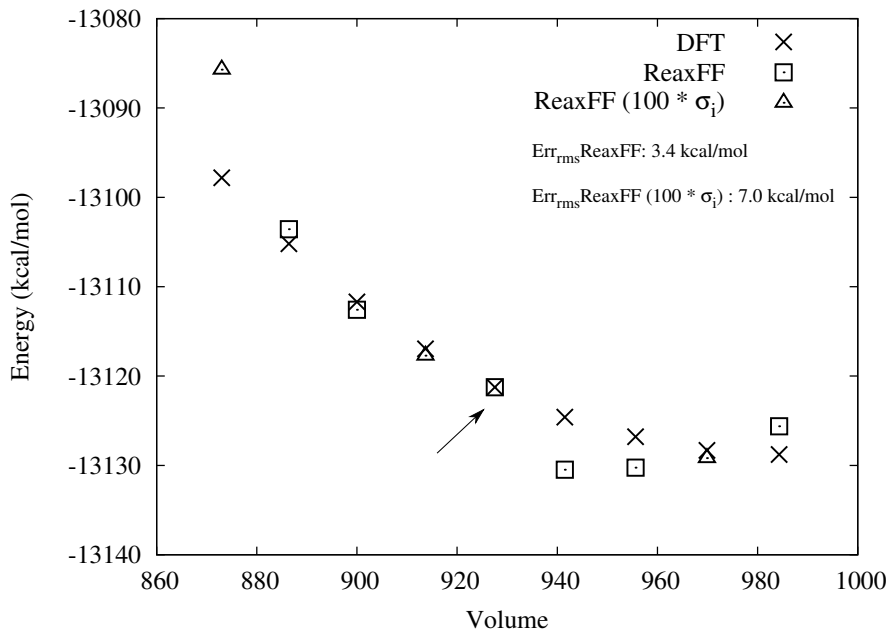


Figure 9: Equations of state for $MgSO_4 \cdot 7H_2O$ unit cell. Results from the optimized force field are compared with DFT results. The low-weightage data points ($100 \times \sigma_i$) also compare well with the DFT data points. The arrow shows the calibration point to correct for the reference state.

The unit cell of magnesium sulfate tetra-hydrate contains four molecules, thus in total 72 atoms in a non-orthorhombic unit cell. Similarly, the hepta hydrate orthorhombic unit cell contains 108 atoms. The equations of state data obtained from DFT as well as from the optimized ReaxFF force field are shown in Figures 9 and 10. The hepta-hydrate data show good agreement with the DFT results with an rms error of 3.4 kcal/mol for the optimized

data, and predicted the energy of the data points with low-weightage with an rms error of 7.0 kcal/mol. The deviation in the energies before the Metropolis minimization was 80 kcal/mol.

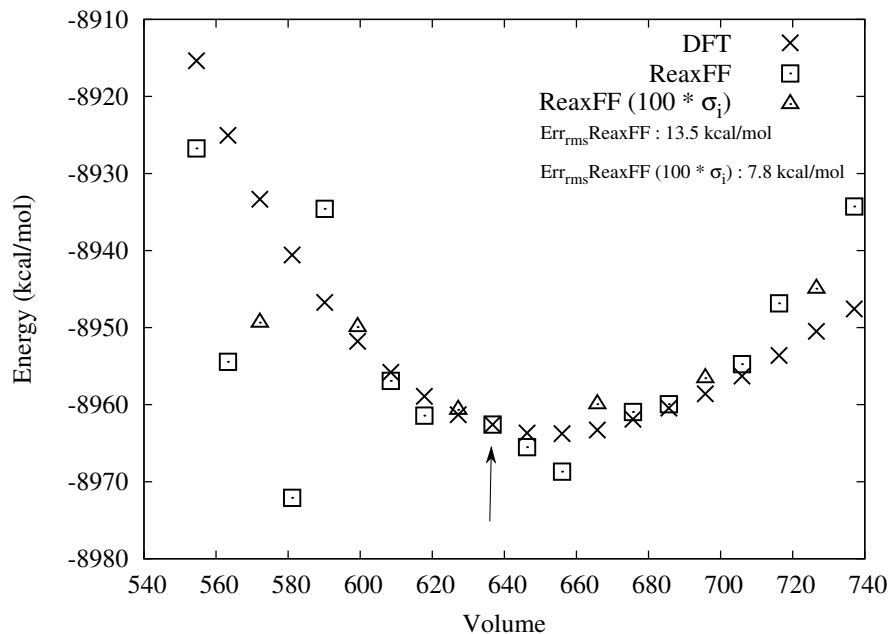


Figure 10: Equations of state for $MgSO_4 \cdot 4H_2O$ unit cell. Results from the optimized force field are compared with DFT results. Significant deviation is observed in the compressed crystal. The low-weightage data points ($100 \times \sigma_i$) compare well with the DFT data as well. The arrow shows the calibration point to correct for the reference state.

Similarly, the comparison between ReaxFF and DFT for the equations of state data for the tetra-hydrated magnesium sulfate is shown in Figure 10. Although the ReaxFF data points follow the correct trend, the compressed crystal (volume $< 600 \text{ \AA}^3$) shows a significant deviation resulting in an rms error of 13.5 kcal/mol (The rms error using parabolic-search optimization was approximately 90 kcal/mol). Nevertheless, the transferability of the force field within the equations of state data set appears to be as good as in the previous cases. The low-weightage data points in all these cases appear to follow the trend of the optimized geometries. This shows the good transferability of the ReaxFF force field for the regions close to the data used for training.

Water binding energy of $MgSO_4$ crystal

Energy as a function of water binding on the (100) surface of $MgSO_4$ crystal is calculated using DFT. A periodic unit cell of $MgSO_4$ with a 10\AA vacuum space in x-direction is used for the calculation. A water molecule is placed at different locations in the vacuum space, and then a geometry optimization is performed after constraining the distance between the O atom in the water molecule and one of the Mg atoms at the (100) surface of the crystal. Figure 11 shows the variation of energy as the water molecule approaches the slab surface (from right to left). ReaxFF reproduces the correct trend of the energy surface after optimization. Even the low-weightage data follows the DFT results.

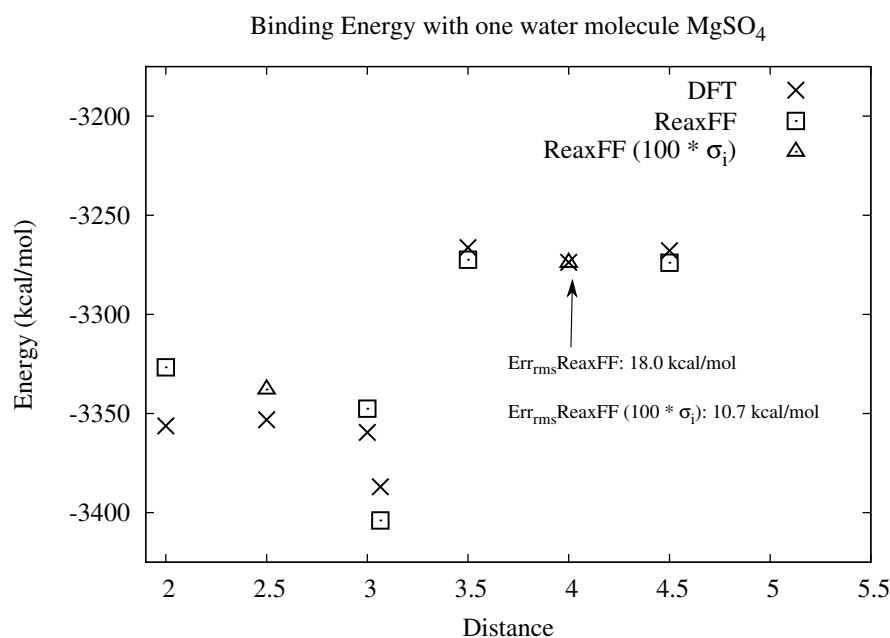


Figure 11: Energy of water binding at (100) surface of $MgSO_4$ crystal. Results from the optimized force field are compared with DFT results. The agreement between DFT and ReaxFF is within reasonable accuracy limit. The low-weightage data points ($100 \times \sigma_i$) compare well with the DFT data as well. The arrow shows the calibration point to correct for the reference state.

Prediction of the equations of state for magnesium sulfate penta-hydrate

To assess the extent of the transferability of the force field, a comparison is made between the equations of state data of magnesium sulfate penta-hydrate with the predictions from the optimized ReaxFF force field. This set of data was not included in the training set, and thus can be marked as a validation of the ability of the force field to predict the energy surface beyond the sampled data set. The non-orthorhombic unit cell of $MgSO_4 \cdot 5H_2O$ contains 42 atoms. The resulting comparison between the DFT and ReaxFF predictions is shown in Figure 12.

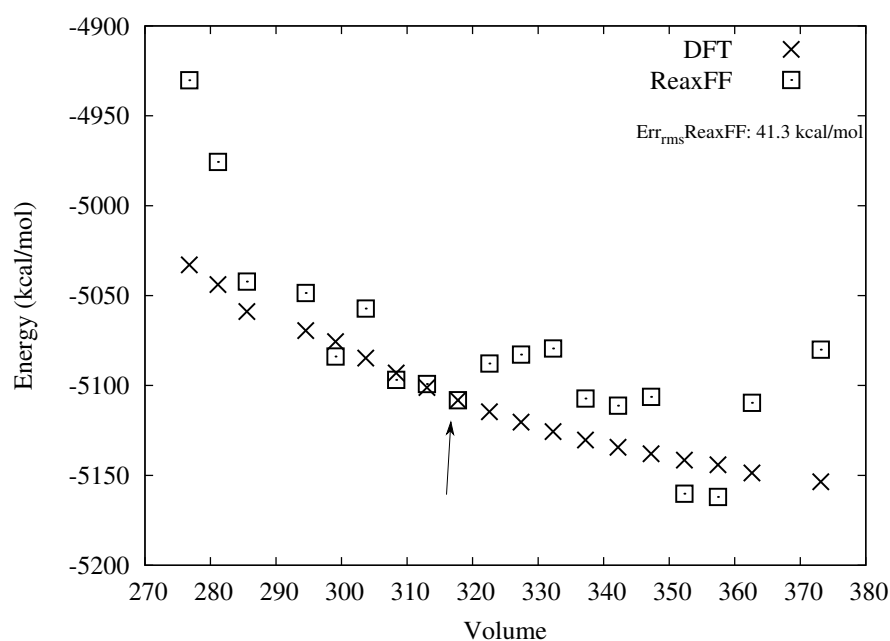


Figure 12: Equations of state for $MgSO_4$ penta hydrate. This data was not included in the training set. The comparison shows the transferability of the ReaxFF force field to a data which was not included in the training set. The arrow shows the calibration point to correct for the reference state.

Since the equation of state for the penta-hydrate was not included in the training set, one would not expect the rms error for the penta-hydrate, Figure 12, to be lower than the average rms error obtained in the previous results. The trend in the energy surface for penta-hydrate

is captured by the ReaxFF calculation, although the rms error of 41.3 kcal/mol is relatively large, resulting in an inferior level of transferability of the force field for penta-hydrate. The rms error observed for this case with the force field obtained after the parabolic-search optimization procedure was found to be 51.5 kcal/mol.

Discussion

Not many attempts have been made in the past to find a suitable optimization scheme to parameterize the Reactive Force Fields. From a conceptual point of view, it appears that the Metropolis scheme is more robust compared to the single parameter parabolic-search algorithm in searching a global minimum for the present problem. The main advantage of the Metropolis Monte-Carlo scheme with Simulated Annealing (MMC-SA) is that it can climb over local maxima and will not be stuck to local minima in the parameter space. A comparison with the single parameter search algorithm is difficult to illustrate here as the success of the single parameter search is entirely dependent on the starting point in the parameter space. Nevertheless, a comparison of the rms errors for various data sets in the training set between the single parameter search and Metropolis is given in Table 4. We admit that this inferior comparison of the parabolic-search against the new Metropolis scheme might be due to a poor starting point in the parameter space. Therefore, if one has a good guess for the parameters to start with, then it is possible that the single parameter search will lead to more or less to the same global minimum. Again, since the success of the Metropolis scheme is not as much dependent on the initial conditions, it is proposed to be an appropriate scheme to parameterize the empirical ReaxFF force field, without having to require much prior knowledge about the approximate values for the parameters.

The optimized force field acts as an approximate model which reproduces the information given in the training set. Naturally, any deviation in the training set data from reality will be reproduced in the optimized force field as well. One should thus keep in mind that the force field can never be better than the training set data. Riley et.al.²⁸ has tabulated the magnitude of various errors from DFT calculation, and for various energies. They have reported that DFT shows an average error in the conformational energies of approximately

Table 4: Comparison of the rms errors (kcal/mol) for the ReaxFF optimized data

Data set	Figure	Parabolic-search	Metropolis	Metropolis (low-weightage)
$MgSO_4.xH_2O$	5	30	9.8	
Proton transfer	7	27	8.4	9.4
EOS- $MgSO_4$	8	70	4.8	4.5
EOS- $MgSO_4.7H_2O$	9	80	3.4	7.0
EOS- $MgSO_4.4H_2O$	10	90	13.5	7.8
Binding Energy	11	22	18.0	10.7
EOS- $MgSO_4.5H_2O$	12	51.5	41.3	

15 kcal/mol. In addition, these errors vary from system to system. Therefore, the average rms variation observed in the present optimized force field may be regarded as satisfactory. Errors can also arise from choosing different models (xc-functionals or basis set) in DFT for various data sets in the training set. One such instance is the case for the proton transfer reaction in the magnesium sulfate hexahydrate molecule as presented in Figure 7, where the B3LYP functional was used for sampling the proton transfer pathway and the PW91 functional was used to optimize the hexahydrate molecule as presented in Figure 6(a), resulting in a deviation in equilibrium $O \cdots H$ distance of 0.08 \AA in the hexahydrate molecule. Such inconsistencies may thus be appropriately taken into account by adjusting the weighting factors σ_i in the training set.

Transferability of the ReaxFF force field for a system, other than what it has been optimized for, is a question which needs to be addressed to assess the applicability of the force field. The general trend of the data points with low-weighting factors as presented in Figures 7-11, validate the excellent transferability of the force field for points close to the

training set data in the potential energy surface. In all the cases, the rms errors for the low-weightage data points are comparable to the of data points.

For the case of the equations of state for the penta-hydrate crystal, Figure 12, which was not included in the training set, ReaxFF captures the right trend in the data although the rms error is relatively large. This shows that the transferability of the ReaxFF force field, for data points beyond which it has been optimized for, should be carefully examined before it may be used for molecular dynamics simulations. At this juncture, the strategy applied by van Duin et.al.¹², in accordance with Wood et.al.⁴⁷, to randomly sample the relevant configurational space from an MD simulation, seems to be the best suitable approach for effectively sampling the potential energy surface. In this method, starting with an approximate force field, an MD simulation is performed on the desired chemical processes and the configurational space is sampled. The DFT energies of such samples are used in the training set to better capture the potential energy surface of the reaction coordinates. This method ensures that the sampling of the potential energy surface is made closer to the relevant regions in the configurational space, and thus results in a good force field applicable in those areas of the configurational space.

We have maintained an acceptance rate of 30% throughout the calculation by scaling the $\delta p_{i,max}$. If we were to maintain lower acceptance rate, e.g. 10%, then for a particular system, at a given temperature, $\delta p_{i,max}$ will be larger in comparison with the case with 30%. This implies that the aggressiveness of random propositions increases if one is working with a lower target acceptance rate and vice versa. Therefore, when one maintains a low acceptance rate, the chances for resulting in a better approximation for the parameters increases due to the aggressive propositions. However, highly aggressive moves in our system often results in unacceptably large *Error* values, resulting in worthless computational effort (due to lower acceptance rate). In essence, a combination of a slow rate of cooling, appropriate perturbation scaling and a low target acceptance rate will eventually get us closer to the global minimum in the parameter space. On the other hand, this will increase the computational time, and therefore a balance needs to be found in choosing the right values for these settings.

Conclusion

A Metropolis Monte Carlo Method with Simulated Annealing (MMC-SA) method is used to optimize the ReaxFF force field against quantum chemical data of hydrated magnesium sulfate systems. This method is more robust compared to the traditional single parameter parabolic-search algorithm. The performance of the MMC method and the parabolic-search method for optimizing the ReaxFF force field is compared in Figure 3. From the figure, it is clear that the MMC method perform much better than the parabolic-search method in optimizing the force field when starting from a random initial force field. A good agreement is obtained between the results from the optimized force field and DFT data for the structural and configurational energies of the hydrated molecules of magnesium sulfate. The force field reproduces the proton transfer phenomena observed in the hexahydrated molecule. In addition, the equations of state of anhydrate, tetra-hydrate, penta-hydrate and hepta-hydrates were also reproduced by the optimized force field within satisfactory accuracy limit. Using the Metropolis Monte Carlo algorithm to optimize the ReaxFF force field avails us the benefit of searching for the global minimum in a high dimensional parameter space. Compared to the single parameter parabolic-search algorithm¹⁶, the Metropolis algorithm is robust and can climb over local maxima, and, therefore, does not depend as much on the initial conditions. In addition, a set of several parameters can be selected simultaneously for each trial unlike in the parabolic-search method where only one parameter at a time can be optimized. Table 4 also shows the improvement of the MMC-SA method over the single-parameter parabolic-search algorithm. The simulated annealing method provides a better control on the parameterization procedure.

The transferability of the force field within the close range of the training set data is proven to be excellent. The agreement is found to be less accurate in the case of a data set farther away from the training set (penta-hydrate). The choice of the data set used for training is thus important in reproducing the correct behavior of the system in MD simulations, and it is asserted that the ReaxFF force field is not indefinitely transferable for a set of atoms and one should make a careful judgement in building the appropriate training set. The method used by van Duin et.al.¹² for randomly sampling the configurational space,

seems to be an appropriate choice for building the training set. Further investigation is needed to find more effective methods in efficiently sampling the configurational space and in optimizing the parameters.

Acknowledgement

We would like to thank Dr. Adri van Duin, for the fruitful discussions and guidance which eventually led to the development of this Metropolis algorithm for force field optimization. We also would like to thank National Computing Facility-SARA for providing us with the necessary computational resources for this project. We are grateful to Scientific Computing and Modelling (SCM) for their computational packages ADF, BAND and ReaxFF. The current implementation is made as an add-on package to the ReaxFF binaries of SCM.

References

1. A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, *J. Phys. Chem. A* **105**, 9396 (2001).
2. K. Chenoweth, A. C. van Duin, and W. A. Goddard, *J. Phys. Chem. A* **112**, 1040 (2007).
3. K. Chenoweth, S. Cheung, A. C. T. van Duin, W. A. Goddard, and E. M. kober kober kober, *J. Am. Chem. Soc.* **127**, 7192 (2005).
4. A. C. T. van Duin, A. Strachan, S. Stewman, Q. Zhang, X. Xu, and G. W. A., *J. Phys. Chem. A* **107**, 3803 (2003).
5. S. Cheung, W. Q. Deng, A. C. T. van Duin, and W. A. Goddard, *J. Phys. Chem. A* **109**, 851 (2005).
6. S. S. Han, A. C. T. van Duin, W. A. Goddard, and H. M. Lee, *J. Phys. Chem. A* **109**, 4575 (2005).
7. P. Valentini, T. E. Schwartzentruber, and L. Cozmuta, *J. Chem. Phys.* **133**, 084703 (2010).
8. M. Aryanpour, A. C. T. van Duin, and J. D. Kubicki, *J. Phys. Chem. A* **114**, 6298 (2010).
9. J. A. Keith, D. Fantauzzi, and T. Jacob, *Phys. Rev. B* **81**, 235404 (2010).
10. K. Joshi, A. C. T. van Duin, and T. Jacob, *J. Mat. Chem.* **20**, 10431 (2010).
11. M. R. LaBrosse, J. K. Johnson, and A. C. T. van Duin, *J. Phys. Chem. A* **114**, 5855 (2010).
12. A. C. T. van Duin, V. S. Bryantsev, M. S. Diallo, W. A. Goddard, O. Rahaman, D. J. Doren, D. Raymand, and K. Hermansson, *J. Phys. Chem. A* **114**, 9507 (2010).
13. O. Rahaman, A. C. T. van Duin, V. S. Bryantsev, J. E. Mueller, S. D. Solares, W. A. Goddard, and D. J. Doren, *J. Phys. Chem. A* **114**, 3556 (2010).

14. V. van Essen, H. Zondag, J. C. Gores, L. Bleijendaal, M. Bakker, R. Schuitema, W. van Helden, Z. He, and C. C. M. Rindt, *Journal of Solar Energy Engineering* **131**, 041014 (2009).
15. K. E. N'Tsoukpoe, H. Liu, N. L. Pierrés, and L. Luo, *Renewable and Sustainable Energy Reviews* **13**, 2385 (2009).
16. A. C. T. van Duin, J. M. A. Baas, and B. van de Graaf, *J. Chem. Soc. Faraday Trans.* **90**, 2881 (1994).
17. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
18. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220**, 671 (1983).
19. J. E. Jones, *Proc. R. Soc. Lond.* **106**, 463 (1924).
20. A. J. Markvoort, P. A. J. Hilbers, and S. V. Nedeia, *Phys. Rev. E* **71**, 066702 (2005).
21. D. W. Brenner, *Phys. Rev. B* **42** (9458-9471).
22. J. Tersoff, *Phys. Rev. Lett* **61**, 2879 (1988).
23. D. W. Brenner, O. A. Shenderova, and J. A. Harrison, *J. Phys. Cond. Matt.* **14**, 783802 (2002).
24. S. J. Stuart, A. B. Tutein, and J. A. Harrison, *J. Chem. Phys.* **112**, 6472 (2000).
25. M. S. Daw and M. I. Baskes, *Phys. Rev. B* **29**, 6443 (1984).
26. R. Car and M. Parrinello, *Phys. Rev. Lett* **55**, 2471 (1985).
27. W. J. Mortier, S. K. Ghosh, and S. Shankar, *J. Am. Chem. Soc.* **108**, 4315 (1986).
28. K. E. Riley, B. T. O. Holt, and J. Kenneth M. Merz, *J. Chem. Theory Comput.* **3**, 407 (2007).
29. C. T. Klein, B. Mayer, G. Köhler, and P. Wolschann, *J. Comp. Chem.* **19**, 1470 (1998).

30. D. J. Diller and C. L. M. J. Verlinde, *J. Comp. Chem.* **20**, 1740 (1999).
31. S. Pongsai, *J. Comp. Chem.* **13**, 1980 (2010).
32. A. Caffisch, P. Niederer, and M. Anliker, *PROTEINS: Structure, Function, and Genetics* **14**, 102 (1992).
33. B. Li, P. Sun, Q. Jin, J. Wang, and D. Ding, *J. Mol. Struct. (Theochem)* **391**, 250 (1997).
34. S. R. Wilson and W. Cui, *Biopolymers* **29**, 225 (1990).
35. S. R. Wilson, W. Cui, J. W. Moskowitz, and K. E. Schmidt, *J. Comput. Chem.* **12**, 342 (1991).
36. <http://www.scm.com>.
37. M. Stone, *J. R. Stat. Soc.* **36**, 111 (1974).
38. <https://www.sara.nl>.
39. E. Iype, S. V. Nedeia, C. C. M. Rindt, A. A. van Steenhoven, H. A. Zondag, and A. P. J. Jansen, *J. Phys. Chem. C* **116**, 18584 (2012).
40. D. A. McQuarrie, *Physical Chemistry: A Molecular Approach* (University Science Books, Sausalito, CA, 1997).
41. F. D. Murnaghan, *Proc. Natl. Acad. Sci.* **30**, 244 (1944).
42. B. Francis, *Phys. Rev.* **71**, 809 (1947).
43. P. J. Rentzeperis and C. T. Soldatos, *Acta. Cryst.* **11**, 686 (1958).
44. W. H. Baur, *Acta. Cryst.* **17**, 863 (1964).
45. W. H. Baur and J. L. Rolin, *Acta. Cryst. B* **28**, 1448 (1972).
46. W. H. Baur, *Acta. Cryst.* **17**, 1361 (1964).

47. R. H. Wood, E. M. Yezdimer, S. Sakane, and J. A. Barriocanal, *J. Chem. Phys.* **110**, 1329 (1999).